Numerical solution of the advection-diffusion equation coupled with Darcy's law

Margaux Bruliard
Mathieu Rigal

May 18, 2018

M. Bruliard & M. Rigal, Numerical Solution of the Advection-Diffusion Equation coupled with Darcy's law, 2018, Sup'Galilée

M. Bruliard & M. Rigal, Numerical Solution of the Advection-Diffusion Equation coupled with Darcy's law, 2018, Sup'Galilée

Contents

M. Bruliard & M. Rigal, Numerical Solution of the Advection-Diffusion Equation coupled with Darcy's law, 2018, Sup'Galilée

M. Bruliard & M. Rigal, Numerical Solution of the Advection-Diffusion Equation coupled with Darcy's law, 2018, Sup'Galilée

Chapter 1

Introduction

1.1 Motivation

Advection and diffusion are two physical phenomena referring to the transfer of a quantity in some domain. Although these two phenomena represent passive processes affecting the transport, they derive from quite different causes. The next lines emphasize the distinction between these two:

Advection is the macroscopic effect of general motion due to a bulk flow, such as the fluid motion of a river.

Diffusion constitutes a transformation through random motion of particles at a microscopic level, without any bulk flow being required for it.

Various types of quantities can be considered, but most of the time heat, energy or concentration are being dealt with. When one is analyzing the transfer disregarding its causes — meaning taking into account advection and diffusion at the same time — we speak of *convection*.

The study of the aforesaid transfers is of great interest. In fact many substances are moving through the natural environment, impacting us more or less directly. A better comprehension of these processes can lead to a more efficient prediction of hazardous outcomes, as illustrated in the following examples:

1. **Ecosystem Dynamics.** In order to create energy, living organisms need food resources, and more precisely *nutrients* (oxygen, carbon dioxide, phosphorus and nitrogen among others). Thus the growth or decay of a species population in water can be anticipated through the observation of *nutrients* transfer [2, chap. 1]. An example of this is the green algae bloom that occurred on Brittany beaches in 2009, as a response to phytoplankton and zooplankton dynamics which have been altered by agriculture pollution (76 000 tons of nitrogen was discharged into the sea in 2008 according to the Loire-Brittany Water Agency). This resulted in the death of a worker, and several animal bodies have been found near the coast (including dogs, wild boars and horses).
2. **Toxicity.** Many industrial processes involve the production or use of chemical substances, which at some point may lead to an increase of toxic concentration. In the interest of the worker and population security, it is then essential to keep the probability of such increase as low as possible within reasonable budget. Studying transfers also helps better handle a crisis situation following an accident. This was the case with the Fukushima incident that took place after an offshore earthquake and a tsunami on March 11$^{\text{th}}$, 2011. The radiation

leakage into the atmosphere and the sea that came after was considerable, reaching exceptionally high levels in the direct surrounding of the power plant. A simulation of the radioactive dispersion allowed for a greater understanding of the underlying consequences, and more pertinent counter-actions were decided [4].

3. **Global Climate Change.** Unwanted effects of some chemical species on the atmosphere is pushing us consider ways to reduce their concentration.

1.2 Modeling transfers

In our case, we are interested in the transfer of concentration of a constituent in a porous media. Concentration can refer to the molar concentration (amount of a constituent in moles divided by the volume of the mixture, in mol/m^3), or the number concentration (number of entities of the species divided by the volume of the mixture, in 1/m^3). The system of partial differential equations describing this is given below:

$$(1.1) \quad \begin{cases} -\nabla \cdot (K\vec{\nabla} h(C(\mathbf{x}), \mathbf{x})) = f_1 & (1.1a) \\ \dfrac{\partial C}{\partial t} + \nabla \cdot (\mathbf{q}C) = \nabla \cdot (D\nabla C) + f_2 & (1.1b) \\ \mathbf{q} = -K\vec{\nabla} h & (1.1c) \end{cases}$$

where:

- $C : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is the concentration;
- $D \in \mathcal{M}_2(\mathbb{R})$ is the matrix of diffusion coefficients;
- $\mathbf{q} : \mathbb{R} \times \mathbb{R} \to \mathcal{M}_{2,1}(\mathbb{R})$ is the velocity;
- $K : \mathbb{R} \times \mathbb{R} \to \mathcal{M}_2(\mathbb{R})$ is Darcy's permeability;
- f_1 and f_2 are two source functions;

If the pressure h does not depend on the concentration C, then system (1.1) is linear. This is the case whenever a change of concentration does not impact the dynamics of the media. In the opposite case, the assumption that the pressure h is concentration-independent is not valid anymore, and (1.1) becomes nonlinear.

Chapter 2

Numerical solution of the linear problem

In this section, we discuss a numerical method for solving the linear system, which can be done in three main steps. We will first have to determine the pressure given by equation (1.1a) and Dirichlet boundary conditions, which will allow us to compute the velocity through Darcy's law (1.1c). Finally, we will use a flux approach with equation (1.1b) in order to approximate the concentration. It is interesting to note that in the linear case, the pressure can be calculated one time, since it does not depend on the concentration. This is quite convenient given the fact that the first step can be very demanding in terms of number of operations. roadmap

2.1 Domain and time discretization

Before going any further, we must introduce a discretization of the domain $\Omega \subset R^2$, and to simplify the problem we shall choose $\Omega =]0,1[\times]0,1[$. Then we divide this domain in a collection of rectangular cells $(\mathcal{C}_{ij})_{i,j}$ part of a structured mesh. Let $N_x, N_y \in \mathbb{N}$ such that N_x is the number of columns and N_y the number of rows forming the grid , and we note $\Delta x = 1/N_x$ and $\Delta y = 1/N_y$. For each pair $(i,j) \in [\![0, N_x - 1]\!] \times [\![0, N_y - 1]\!]$, we consider the index $k = i + (N_x - 1)j$ in order to have the bijective mapping $\mathcal{C}_k \leftrightarrow \mathcal{C}_{ij}$, as shown in fig. 2.1.

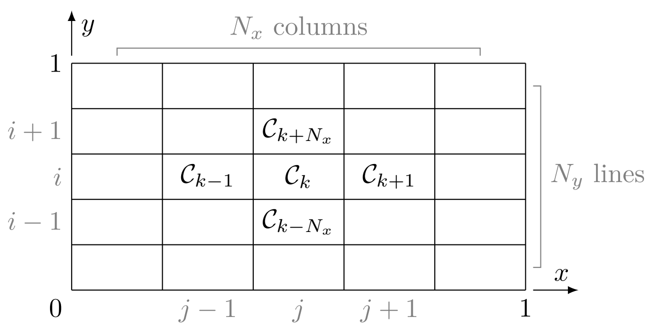

Figure 2.1: Rectangular grid

We define $\mathcal{N}_k$ the index set of all neighbors of cell $\mathcal{C}_k$:

$$\mathcal{N}_k = \left\{ i \in [0, N_x N_y - 1],\ \mathcal{C}_k \cap \mathcal{C}_i \text{ is a segment} \right\}$$

We will call γ_{kl} the frontier between the k^{th} cell and the l^{th} cell, with $l \in \mathcal{N}_k$. Thus we have: $\forall k \in [0, N_x N_y - 1]$ such that $\mathcal{C}_k$ is not a border cell, $\partial \mathcal{C}_k = \bigcup_{l \in \mathcal{N}_k} \gamma_{kl}$.

As for time discretization, we will divide an interval of time $[0, T]$ into N_t sub-intervals, meaning that we will get $N_t + 1$ points of time discretization.

2.2 A finite volume approach for the pressure field

2.2.1 The diamond scheme

pressure approximation

Now that the domain is discretized, we start by approximating the pressure h . To achieve this, we use a finite volume approach that will lead us to the *diamond scheme*. Because we are considering Dirichlet boundary conditions, the values of h on the border of the domain are already known. For this reason, in what follows we restrict ourselves to indices k or (i, j) that are not referring to a border cell. Put in other words, we solve the problem on the mesh exluding its first and last rows and columns, which will be referred to as the *truncated mesh* of $N_x - 2$ columns and $N_y - 2$ rows. With this in mind, we first integrate equation (1.1a) over a cell $\mathcal{C}_k$ for a fixed k:

$$\begin{aligned} &\int_{\mathcal{C}_k} f_1 dx = -\int_{\mathcal{C}_k} \nabla \cdot (K \vec{\nabla} h) dx = -\int_{\partial \mathcal{C}_k} K \vec{\nabla} h \cdot \mathbf{n}\, d\sigma \\ \Longrightarrow\ &|\mathcal{C}_k| f_{1,\mathcal{C}_k} = -\sum_{l \in \mathcal{N}_k} \int_{\gamma_{kl}} K \vec{\nabla} h \cdot \mathbf{n}\, d\sigma \qquad \text{with } f_{1,\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|} \int_{\mathcal{C}_k} f_1 dx \end{aligned} \tag{2.1}$$

The intermediate step here is to approximate $\vec{\nabla} h = (\partial_x h, \partial_y h)^T$. In this regard, we do the following assumption: the pressure gradient is a constant on every

diamond cell

diamond cell $\mathcal{D}_{kl}$, where $\mathcal{D}_{kl}$ is a quadrilateral cell such that $\gamma_{kl} \subset \mathcal{D}_{kl}$ (see fig. 2.2). From this new perspective, finding the pressure gradient on γ_{kl} $\forall l \in \mathcal{N}_k$ can be completed by finding its value on $\mathcal{D}_{kl}$ $\forall l \in \mathcal{N}_k$.

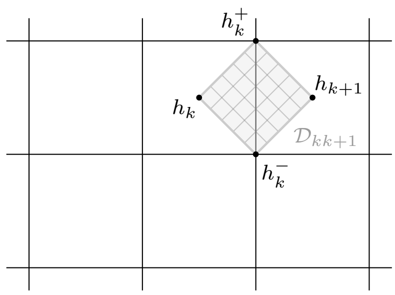

Figure 2.2: Diamond cell centered on the border γ_{kk+1}

Figure 2.2 illustrates the diamond cell centered around γ_{kk+1}. The values h_k^- and

h_k^+ can be approximated at the second order as described below:

$$\begin{cases} h_k = h_k^- - \dfrac{\Delta x}{2}\partial_x h_k^- + \dfrac{\Delta y}{2}\partial_y h_k^- + \mathcal{O}(\Delta x^2 + \Delta y^2) \\ h_{k+1} = h_k^- + \dfrac{\Delta x}{2}\partial_x h_k^- + \dfrac{\Delta y}{2}\partial_y h_k^- + \mathcal{O}(\Delta x^2 + \Delta y^2) \\ h_{k-N_x} = h_k^- - \dfrac{\Delta x}{2}\partial_x h_k^- - \dfrac{\Delta y}{2}\partial_y h_k^- + \mathcal{O}(\Delta x^2 + \Delta y^2) \\ h_{k+1-N_x} = h_k^- + \dfrac{\Delta x}{2}\partial_x h_k^- - \dfrac{\Delta y}{2}\partial_y h_k^- + \mathcal{O}(\Delta x^2 + \Delta y^2) \end{cases}$$

$$\Longrightarrow h_k^- = \frac{h_{k-N_x} + h_{k+1-N_x} + h_k + h_{k+1}}{4} + \mathcal{O}(\Delta x^2 + \Delta y^2) \tag{2.2}$$

Following the same idea we have:

$$h_k^+ = \frac{h_k + h_{k+1} + h_{k+N_x} + h_{k+1+N_x}}{4} + \mathcal{O}(\Delta x^2 + \Delta y^2) \tag{2.3}$$

Lastly, we need to characterize the frontier $\partial\mathcal{D}_{kl}$ of $\mathcal{D}_{kl}$. We define the segments $(d_i)_{1\le i\le 4}$ such that $\partial\mathcal{D}_{kl} = \bigcup_{1\le i\le 4} d_i$, d_1 being the lower right diamond border, d_2 the upper right, d_3 the upper left and d_4 the lower left. It is clear that $\forall 1 \le i \le 4,\ |d_i| = \frac{1}{2}\sqrt{\Delta x^2 + \Delta y^2}$, and the associated unit normal vectors are:

$$\mathbf{n}_{d_1} = \frac{1}{2|d_1|}\begin{pmatrix}\Delta y \\ -\Delta x\end{pmatrix} = -\mathbf{n}_{d_3}, \quad \mathbf{n}_{d_2} = \frac{1}{2|d_1|}\begin{pmatrix}\Delta y \\ \Delta x\end{pmatrix} = -\mathbf{n}_{d_4}$$

We carry on by integrating $\vec{\nabla} h$ over $\mathcal{D}_{kl}$:

$$\int_{\mathcal{D}_{kl}} \vec{\nabla} h dx = \int_{\partial\mathcal{D}_{kl}} h\mathbf{n}_{kl} d\sigma \tag{2.4}$$

Knowing that $\vec{\nabla} h$ is a constant in $\mathcal{D}_{kl}$ we can write:

$$|\mathcal{D}_{kl}|\vec{\nabla} h_{|\mathcal{D}_{kl}} = \int_{\partial\mathcal{D}_{kl}} h\mathbf{n}_d d\sigma \Longrightarrow \vec{\nabla} h_{|\gamma_{kl}} = \frac{1}{|\mathcal{D}_{kl}|}\int_{\partial\mathcal{D}_{kl}} h\mathbf{n}_d d\sigma \tag{2.5}$$

Going back to equation (2.1) we have:

$$\begin{aligned} -|\mathcal{C}_k| f_{1,\mathcal{C}_k} &= \sum_{l\in\mathcal{N}_k} \int_{\gamma_{kl}} K\vec{\nabla} h \cdot \mathbf{n} d\sigma \\ &= \sum_{l\in\mathcal{N}_k} |\gamma_{kl}| K\vec{\nabla} h_{|\gamma_{kl}} \cdot \mathbf{n}_{kl} \\ &= \sum_{l\in\mathcal{N}_k} \frac{|\gamma_{kl}|}{|\mathcal{D}_{kl}|}\left(K\int_{\partial\mathcal{D}_{kl}} h\mathbf{n}_d d\sigma\right)\cdot \mathbf{n}_{kl} \end{aligned} \tag{2.6}$$

Then we rewrite the integrals as below:

$$\int_{\partial\mathcal{D}_{kk+1}} h\mathbf{n}_d d\sigma$$

$$= \frac{|d_1|}{2}(h_k^- + h_{k+1})\mathbf{n}_{d_1} + \frac{|d_2|}{2}(h_{k+1} + h_k^+)\mathbf{n}_{d_2} + \frac{|d_3|}{2}(h_k^+ + h_k)\mathbf{n}_{d_3}$$

$$
\begin{aligned}
&+ \frac{|d_4|}{2}(h_k + h_k^-)\mathbf{n}_{d_4} + (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y) \\
&= \frac{|d_1|}{2}\Big(h_k^-(\mathbf{n}_{d1} + \mathbf{n}_{d4}) + h_k(\mathbf{n}_{d3} + \mathbf{n}_{d4}) + h_{k+1}(\mathbf{n}_{d1} + \mathbf{n}_{d2}) \\
&\qquad + h_k^+(\mathbf{n}_{d2} + \mathbf{n}_{d3})\Big) + (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y) \\
&= \frac{1}{2}\left(-h_k^-\begin{pmatrix}0\\ \Delta x\end{pmatrix} - h_k\begin{pmatrix}\Delta y\\ 0\end{pmatrix} + h_{k+1}\begin{pmatrix}\Delta y\\ 0\end{pmatrix} + h_k^+\begin{pmatrix}0\\ \Delta x\end{pmatrix}\right) \\
&+ (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y) \\
&= \frac{1}{8}\begin{pmatrix}0\\ \Delta x\end{pmatrix}(h_{k+N_x} + h_{k+1+N_x} - h_{k-N_x} - h_{k+1-N_x}) \\
&+ \frac{1}{2}\begin{pmatrix}\Delta y\\ 0\end{pmatrix}(h_{k+1} - h_k) + (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y)
\end{aligned} \tag{2.7}
$$

We do the same for the three other borders:

$$
\begin{aligned}
\int_{\partial\mathcal{D}_{kk+N_x}} h\mathbf{n}_d d\sigma &= \frac{1}{8}\begin{pmatrix}\Delta y\\ 0\end{pmatrix}(h_{k+1+N_x} + h_{k+1} - h_{k-1+N_x} - h_{k-1}) \\
&+ \frac{1}{2}\begin{pmatrix}0\\ \Delta x\end{pmatrix}(h_{k+N_x} - h_k) + (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y)
\end{aligned} \tag{2.8}
$$

$$
\begin{aligned}
\int_{\partial\mathcal{D}_{kk-1}} h\mathbf{n}_d d\sigma &= \frac{1}{8}\begin{pmatrix}0\\ \Delta x\end{pmatrix}(h_{k-1+N_x} + h_{k+N_x} - h_{k-1-N_x} - h_{k-N_x}) \\
&+ \frac{1}{2}\begin{pmatrix}\Delta y\\ 0\end{pmatrix}(h_k - h_{k-1}) + (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y)
\end{aligned} \tag{2.9}
$$

$$
\begin{aligned}
\int_{\partial\mathcal{D}_{kk-N_x}} h\mathbf{n}_d d\sigma &= \frac{1}{8}\begin{pmatrix}\Delta y\\ 0\end{pmatrix}(h_{k+1} + h_{k+1-N_x} - h_{k-1} - h_{k-1-N_x}) \\
&+ \frac{1}{2}\begin{pmatrix}0\\ \Delta x\end{pmatrix}(h_k - h_{k-N_x}) + (\Delta x + \Delta y)\vec{\varepsilon}(\Delta x, \Delta y)
\end{aligned} \tag{2.10}
$$

Substituting equations (2.7) to (2.10) into (2.6) we get a nine points finite volume scheme:

$$
\begin{aligned}
-8|\mathcal{C}_k|^2 f_{1,\mathcal{C}_k} = &-\Delta x\Big[K_{21}\Delta y(h_{k+1} + h_{k+1-N_x} - h_{k-1} - h_{k-1-N_x}) \\
&\qquad + 4K_{22}\Delta x(h_k - h_{k-N_x})\Big] \\
&- \Delta y\Big[4K_{11}\Delta y(h_k - h_{k-1}) \\
&\qquad + K_{12}\Delta x(h_{k-1+N_x} + h_{k+N_x} - h_{k-1-N_x} - h_{k-N_x})\Big] \\
&+ \Delta y\Big[4K_{11}\Delta y(h_{k+1} - h_k) \\
&\qquad + K_{12}\Delta x(h_{k+N_x} + h_{k+1+N_x} - h_{k-N_x} - h_{k+1-N_x})\Big] \\
&+ \Delta x\Big[K_{21}\Delta y(h_{k+1+N_x} + h_{k+1} - h_{k-1+N_x} - h_{k-1}) \\
&\qquad + 4K_{22}\Delta x(h_{k+N_x} - h_k)\Big]
\end{aligned}
$$

$$
\begin{aligned}
& + \mathcal{O}(\Delta x^2 + \Delta y^2) \\
-8|\mathcal{C}_k|^2 f_{1,\mathcal{C}_k} = \; & (K_{12}+K_{21})\Delta x \Delta y h_{k-1-N_x} + 4K_{22}\Delta x^2 h_{k-N_x} \\
& - (K_{12}+K_{21})\Delta x \Delta y h_{k+1-N_x} + 4K_{11}\Delta y^2 h_{k-1} \\
& - 8(K_{22}\Delta x^2 + K_{11}\Delta y^2) h_k + 4K_{11}\Delta y^2 h_{k+1} \\
& - (K_{12}+K_{21})\Delta x \Delta y h_{k-1+N_x} + 4K_{22}\Delta x^2 h_{k+N_x} \\
& + (K_{12}+K_{21})\Delta x \Delta y h_{k+1+N_x} \\
& + \mathcal{O}(\Delta x^2 + \Delta y^2)
\end{aligned} \tag{2.11}
$$

Finally we give the matrix form of (2.11):

$$
\mathrm{R}\mathbf{h} = \mathbf{b} \tag{2.12}
$$

The rigidity matrix R and the vectors $\mathbf{h}, \mathbf{b}$ are defined below:

$$
\mathrm{R} = \begin{pmatrix}
\mathrm{A} & \mathrm{B}^T & & & \\
\mathrm{B} & \mathrm{A} & \mathrm{B}^T & & \\
& \ddots & \ddots & \ddots & \\
& & \mathrm{B} & \mathrm{A} & \mathrm{B}^T \\
& & & \mathrm{B} & \mathrm{A}
\end{pmatrix}
\qquad
\begin{aligned}
\mathbf{h} &= \big(h_k\big)_{\substack{0 \le k \le N_x N_y - 1 \\ \mathcal{C}_k \text{ not a border cell}}} \\
\mathbf{b} &= \big(-8\Delta x^2 \Delta y^2 f_{1,\mathcal{C}_k}\big)_{\substack{0 \le k \le N_x N_y - 1 \\ \mathcal{C}_k \text{ not a border cell}}}
\end{aligned}
$$

$$
\mathrm{A} = \begin{pmatrix}
\alpha & \beta & & & \\
\beta & \alpha & \beta & & \\
& \ddots & \ddots & \ddots & \\
& & \beta & \alpha & \beta \\
& & & \beta & \alpha
\end{pmatrix}
\quad \text{with} \quad
\begin{cases}
\alpha = -8(K_{22}\Delta x^2 + K_{11}\Delta y^2) \\
\beta = 4K_{11}\Delta y^2
\end{cases}
$$

$$
\mathrm{B} = \begin{pmatrix}
\gamma & -\delta & & & \\
\delta & \gamma & -\delta & & \\
& \ddots & \ddots & \ddots & \\
& & \delta & \gamma & -\delta \\
& & & \delta & \gamma
\end{pmatrix}
\quad \text{with} \quad
\begin{cases}
\gamma = 4K_{22}\Delta x^2 \\
\delta = (K_{12}+K_{21})\Delta x \Delta y
\end{cases}
$$

In virtue of what has been stated regarding the Dirichlet boundary conditions, the problem is being solved on a mesh of size $(Nx-2)(N_y-2)$, and thus $\mathrm{R} \in \mathcal{M}_{(N_x-2)(N_y-2)}(\mathbb{R})$, $\mathrm{A}, \mathrm{B} \in \mathcal{M}_{N_x-2}(\mathbb{R})$, and $\mathbf{b}, \mathbf{h} \in \mathcal{M}_{(N_x-2)(N_y-2),1}(\mathbb{R})$. Since we excluded the boundary values of $\mathbf{h}$, they don't appear in the system $\mathrm{R}\mathbf{h} = \mathbf{b}$ which is true if $h_{|\partial\Omega} = 0$. Therefore, expression (2.12) is appropriate only when considering the homogeneous Dirichlet conditions, the non-homogeneous case being treated in subsection 2.2.2.

Since we would like to solve (2.12), we give a condition under which the rigidity matrix R is nonsingular. If $K_{11} \neq 0$, $K_{22} \neq 0$ and $K_{12} = K_{21} = 0$ are satisfied, we can prove that R is a *weakly chained diagonally dominant* matrix, and thus is inversible. In fact for each line i in R we verify that $|\mathrm{R}_{ii}| - \sum_{j \neq i} |\mathrm{R}_{ij}| \geq 0$. Furthermore, for each row that is not strictly diagonally dominant there exist a walk in the directed graph of R ending at a strictly diagonally dominant row. This can be seen by considering the diagonal, upper-diagonal and lower-diagonal of R which contain only non-zero coefficients. A more general result gives us existence and uniqueness of the solution h — and thus the nonsingularity of R — when K is positive-definite (see appendix A.1).

In the more general case where the permeability K has different values depending on the position, the rigidity and second member matrix are more complex to express. Without detailing it, from (2.1) we would have found:

$$\begin{aligned}
-8|\mathcal{C}_k|^2 f_{1,\mathcal{C}_k} = \quad & (\tilde{K}_{12}^{k-1} + \tilde{K}_{21}^{k-N_x})\Delta x\Delta y\, h_{k-1-N_x} \\
& + (\tilde{K}_{12}^{k-1}\Delta x\Delta y - \tilde{K}_{12}^{k+1}\Delta x\Delta y + 4\tilde{K}_{22}^{k-N_x}\Delta x^2)h_{k-N_x} \\
& - (\tilde{K}_{12}^{k+1} + \tilde{K}_{21}^{k-N_x})\Delta x\Delta y\, h_{k+1-N_x} \\
& + (\tilde{K}_{21}^{k-N_x}\Delta x\Delta y - \tilde{K}_{21}^{k+N_x}\Delta x\Delta y + 4\tilde{K}_{11}^{k-1}\Delta y^2)h_{k-1} \\
& - 4(\tilde{K}_{11}^{k-1}\Delta y^2 + \tilde{K}_{11}^{k+1}\Delta y^2 + \tilde{K}_{22}^{k-N_x}\Delta x^2 + \tilde{K}_{22}^{k+N_x}\Delta x^2)h_k \\
& + (\tilde{K}_{21}^{k+N_x}\Delta x\Delta y - \tilde{K}_{21}^{k-N_x}\Delta x\Delta y + 4\tilde{K}_{11}^{k+1}\Delta y^2)h_{k+1} \\
& - (\tilde{K}_{12}^{k-1} + \tilde{K}_{21}^{k+N_x})\Delta x\Delta y h_{k-1+N_x} \\
& + (\tilde{K}_{12}^{k+1}\Delta x\Delta y - \tilde{K}_{12}^{k-1}\Delta x\Delta y + 4\tilde{K}_{22}^{k+N_x}\Delta x^2)h_{k+N_x} \\
& + (\tilde{K}_{12}^{k+1} + \tilde{K}_{21}^{k+N_x})\Delta x\Delta y h_{k+1+N_x} \\
& + \mathcal{O}(\Delta x^2 + \Delta y^2) \qquad (2.13)
\end{aligned}$$

with $\tilde{K}_{ij}^{l}$ being the mean between coefficients K_{ij} respectively at the center of cells $\mathcal{C}_k$ and $\mathcal{C}_l$: $K_{ij}^{l} = \frac{1}{2}(K_{ij}^{k} + K_{ij}^{l})$.

2.2.2 Non-homogeneous Dirichlet boundary conditions

upper left corner

Taking into account non-homogeneous Dirichlet conditions can be achieved by modifying the second member vector $\mathbf{b}$. As a matter of fact, if we take for example the upper left corner of our truncated mesh, we see that the nine points diamond scheme (2.11) cannot be computed directly, since it would require some additional cells. More specifically five points are out of the sub-mesh previously defined, and don't relate to any of the unkown $(h_k)_k$ (see fig. 2.3).

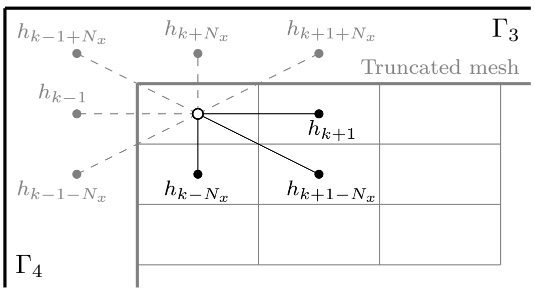

Figure 2.3: Upper left cell $\mathcal{C}_{(N_x-2)(N_y-2)}$. Gray nodes refer to "virtual" points that would exist if there was one more line and one more column in the truncated mesh. In reality, these "virtual" points are replaced by the known Dirichlet boundary conditions.

adapting the diamond scheme

If we note Γ_1 the lower border of the domain, Γ_2 the right border, Γ_3 the upper border, Γ_4 the left border , and $h_{\Gamma_1}, ..., h_{\Gamma_4}$ the corresponding Dirichlet boundary conditions, the nine points scheme could be replaced at the level of $\mathcal{C}_{(N_x-2)(N_y-2)}$ by:

$$\begin{aligned}
-8|\mathcal{C}_k|^2 f_{1,\mathcal{C}_k} = \quad & \delta h_{\Gamma_4}(1-2\Delta y) + \gamma h_{k-N_x} - \delta h_{k+1-N_x} + \beta h_{\Gamma_4}(1-\Delta y) \\
& + \alpha h_k + \beta h_{k+1} - \delta h_{\Gamma_4}(1) + \gamma h_{\Gamma_3}(\Delta x) + \delta h_{\Gamma_3}(2\Delta x)
\end{aligned}$$

$$+ \mathcal{O}(\Delta x^2 + \Delta y^2) \tag{2.14}$$

And thus the new line in the matrix system is:

$$\begin{aligned}
&\sum_{j=1}^{N_x} \mathrm{R}_{N_y j} h_j = \tilde{b}_{N_y} \\
&\text{with } \tilde{b}_{N_y} = \quad b_{N_y} - \delta h_{\Gamma_4}(1 - 2\Delta y) - \beta h_{\Gamma_4}(1 - \Delta y) + \delta h_{\Gamma_4}(1) \\
&\qquad\qquad - \gamma h_{\Gamma_3}(\Delta x) - \delta h_{\Gamma_3}(2\Delta x)
\end{aligned} \tag{2.15}$$

2.3 Computing the velocity

Up to this point, we managed to approximate the pressure field across the mesh. We continue by computing the velocity $\mathbf{q}$ thanks to Darcy's law (1.1c). Using Taylor expansions, we know that for each non-boreder cell $\mathcal{C}_k$ the pressure gradient verifies:

$$\vec{\nabla} h_{|\gamma_{kl}} = \begin{pmatrix} (h_l - h_k)/\Delta x \\ (h_{k-1}^+ - h_{k-1}^-)/\Delta y \end{pmatrix} + \vec{\varepsilon}(\Delta x, \Delta y) \quad \text{if } l = k-1$$

$$\vec{\nabla} h_{|\gamma_{kl}} = \begin{pmatrix} (h_l - h_k)/\Delta x \\ (h_k^+ - h_k^-)/\Delta y \end{pmatrix} + \vec{\varepsilon}(\Delta x, \Delta y) \quad \text{if } l = k+1$$

$$\vec{\nabla} h_{|\gamma_{kl}} = \begin{pmatrix} (h_k^- - h_{k-1}^-)/\Delta x \\ (h_l - h_k)/\Delta y \end{pmatrix} + \vec{\varepsilon}(\Delta x, \Delta y) \quad \text{if } l = k - N_x$$

$$\vec{\nabla} h_{|\gamma_{kl}} = \begin{pmatrix} (h_k^+ - h_{k-1}^+)/\Delta x \\ (h_l - h_k)/\Delta y \end{pmatrix} + \vec{\varepsilon}(\Delta x, \Delta y) \quad \text{if } l = k + N_x$$

If $i \in \{0, N_x - 1\}$ or $j \in \{0, N_y - 1\}$, we will apply Neumann boundary conditions as indicated below:

$$\begin{aligned}
\partial_x h(x_0, y_j) &= \frac{h(x_1, y_j) - h(x_0, y_j)}{\Delta x} + \mathcal{O}(\Delta x) \\
\partial_x h(x_{N_x-1}, y_j) &= \frac{h(x_{N_x-1}, y_j) - h(x_{N_x-2}, y_j)}{\Delta x} + \mathcal{O}(\Delta x) \\
\partial_y h(x_i, y_0) &= \frac{h(x_i, y_1) - h(x_i, y_0)}{\Delta y} + \mathcal{O}(\Delta y) \\
\partial_y h(x_y, y_{N_y-1}) &= \frac{h(x_i, y_{N_y-1}) - h(x_i, y_{N_y-2})}{\Delta y} + \mathcal{O}(\Delta y)
\end{aligned}$$

At last the velocity is deduced from the matrix product:

$$\mathbf{q} = -K\vec{\nabla} h$$

2.4 Computing the concentration

2.4.1 A flux approach

We assume to know the concentration state at time $t_n \in [0,T]$. In order to approximate it at time $t_{n+1} = t_n + \Delta t$ we integrate (1.1b) (with $D = 0$ and $f_2 = 0$) over a given cell $\mathcal{C}_k$ between times t_n and t_{n+1}:

$$\int_{t_n}^{t_{n+1}} \int_{\mathcal{C}_k} \frac{\partial C}{\partial t} dx dt + \int_{t_n}^{t_{n+1}} \int_{\mathcal{C}_k} \nabla \cdot (\mathbf{q} C) dx dt = 0$$

$$\implies |\mathcal{C}_k|(C_k^{n+1} - C_k^n) = -\int_{t_n}^{t_{n+1}} \int_{\partial \mathcal{C}_k} (\mathbf{q} \cdot \mathbf{n}) C d\sigma dt$$

$$\implies C_k^{n+1} = C_k^n - \frac{\Delta t}{|\mathcal{C}_k|} \sum_{l \in \mathcal{N}_k} |\gamma_{kl}| \, \mathbf{q}_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \, \phi_{\gamma_{kl}}(C^n, \mathbf{q})$$

with $\phi_{\gamma_{kl}}$ the flux term corresponding to border γ_{kl}, which is defined as:

$$\forall k, \ \forall l \in \mathcal{N}_k, \ \phi_{\gamma_{kl}}(C^n, \mathbf{q}) = \begin{cases} C_k^n & \text{if } \mathbf{q}_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \geq 0 \\ C_{k-N_x}^n & \text{if } \mathbf{q}_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \leq 0 \text{ and } l = k - N_x \\ C_{k-1}^n & \text{if } \mathbf{q}_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \leq 0 \text{ and } l = k - 1 \\ C_{k+1}^n & \text{if } \mathbf{q}_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \leq 0 \text{ and } l = k + 1 \\ C_{k+N_x}^n & \text{if } \mathbf{q}_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \leq 0 \text{ and } l = k + N_x \end{cases}$$

Then we can write:

$$\begin{aligned} C_k^{n+1} = C_k^n - & \frac{\Delta t}{\Delta x} (q_{x \, |\gamma_{kk+1}} \phi_{\gamma_{kk+1}}(C^n, \mathbf{q}) - q_{x \, |\gamma_{kk-1}} \phi_{\gamma_{kk-1}}(C^n, \mathbf{q})) \\ - & \frac{\Delta t}{\Delta y} (q_{y \, |\gamma_{kk+N_x}} \phi_{\gamma_{kk+N_x}}(C^n, \mathbf{q}) - q_{y \, |\gamma_{kk-N_x}} \phi_{\gamma_{kk-N_x}}(C^n, \mathbf{q})) \end{aligned} \tag{2.16}$$

Since in our case fluxes are modeled by the displacement of the concentration in a cell to its direct neighbors, the step of time has to verify the following condition:

$$\Delta t \leq \min \left(\frac{\Delta x}{\max(|q_x|)}, \frac{\Delta y}{\max(|q_y|)} \right) \tag{2.17}$$

As a matter of fact this condition ensures that at each iteration the transport of the concentration will not be greater than the width or height of a cell, which constitutes a constraint inherent to our scheme.

In order to take charge of the diffusion phenomenon represented by $\nabla \cdot (D \vec{\nabla} C)$ in the continuous equation, we detail its discretization:

$$\begin{aligned} \frac{1}{|\mathcal{C}_k|} \int_{t_n}^{t_{n+1}} \int_{\mathcal{C}_k} \nabla \cdot (D \vec{\nabla} C) \, dx &= \frac{1}{|\mathcal{C}_k|} \int_{t_n}^{t_{n+1}} \int_{\partial \mathcal{C}_k} D \vec{\nabla} C \cdot \mathbf{n}_{|\gamma_{kl}} \, d\sigma \\ &\approx \frac{\Delta t}{|\mathcal{C}_k|} \sum_{l \in \mathcal{N}_k} |\gamma_{kl}| D \vec{\nabla} C_{|\gamma_{kl}} \cdot \mathbf{n}_{|\gamma_{kl}} \end{aligned}$$

The concentration gradient is approximated exactly the same way tan what has been done for the pressure gradient. One should then add the above quantity to the right hand side of the scheme (2.16):

$$C_k^{n+1} = C_k^n - \frac{\Delta t}{\Delta x} (q_{x \, |\gamma_{kk+1}} \phi_{\gamma_{kk+1}}(C^n, \mathbf{q}) - q_{x \, |\gamma_{kk-1}} \phi_{\gamma_{kk-1}}(C^n, \mathbf{q}))$$

$$-\frac{\Delta t}{\Delta y}(q_{y\,|\gamma_{kk+N_x}}\phi_{\gamma_{kk+N_x}}(C^n,\mathbf{q}) - q_{y\,|\gamma_{kk-N_x}}\phi_{\gamma_{kk-N_x}}(C^n,\mathbf{q}))$$
$$+\frac{\Delta t}{|\mathcal{C}_k|}\sum_{l\in\mathcal{N}_k}|\gamma_{kl}|D\vec{\nabla}C_{|\gamma_{kl}}\cdot\mathbf{n}_{|\gamma_{kl}} \tag{2.18}$$

2.4.2 Modified equation and error behavior

Now on we try to get an insight on the error induced by this scheme. First we write the corresponding modified equation in the case where there is no diffusion. This modified equation is solved with a better accuracy by the studied scheme than the original one. The terms appearing in the new equation will inform us on the behavior of the error, as shown below:

error behavior

$$\partial_t C + \frac{\Delta t}{2}\partial^2_{tt}C + \nabla\cdot(\mathbf{q}C) - \frac{\Delta x}{2}\partial^2_{xx}(q_xC) - \frac{\Delta y}{2}\partial^2_{yy}(q_yC) = 0$$
$$\Longrightarrow \partial_t C + \nabla\cdot(\mathbf{q}C) = -\frac{\Delta t}{2}\partial^2_{tt}C + \frac{\Delta x}{2}\partial^2_{xx}(q_xC) + \frac{\Delta y}{2}\partial^2_{yy}(q_yC)$$

If we suppose that $\mathbf{q}$ does not depend on the position, we get:

$$\partial^2_{tt}C + q_x\partial^2_{tx}C + q_y\partial^2_{ty}C = -\frac{\Delta t}{2}\partial^3_{ttt}C + q_x\frac{\Delta x}{2}\partial^3_{txx}C + q_y\frac{\Delta y}{2}\partial^3_{tyy}C$$
$$q_x\partial^2_{xt}C + q_x^2\partial^2_{xx}C + q_xq_y\partial^2_{xy}C = -q_x\frac{\Delta t}{2}\partial^3_{xtt}C + q_x^2\frac{\Delta x}{2}\partial^3_{xxx}C + q_xq_y\frac{\Delta y}{2}\partial^3_{xyy}C$$
$$q_y\partial^2_{yt}C + q_xq_y\partial^2_{yx}C + q_y^2\partial^2_{yy}C = -q_y\frac{\Delta t}{2}\partial^3_{ytt}C + q_xq_y\frac{\Delta x}{2}\partial^3_{yxx}C + q_y^2\frac{\Delta y}{2}\partial^3_{yyy}C$$

By subtracting the second and third lines to the first one we find:

$$\partial^2_{tt}C = q_x^2\partial^2_{xx}C + 2q_xq_y\partial^2_{xy}C + q_y^2\partial^2_{yy}C + \mathcal{O}(\Delta t,\Delta x,\Delta y)$$

Substituting $\partial^2_{tt}C$ in the modified equation holds:

$$\partial_t C + \nabla\cdot(\mathbf{q}C) = -\frac{\Delta t}{2}(q_x^2\partial^2_{xx}C + 2q_xq_y\partial^2_{xy}C + q_y^2\partial^2_{yy}C)$$
$$+ q_x\frac{\Delta x}{2}\partial^2_{xx}C + q_y\frac{\Delta y}{2}\partial^2_{yy}C + \mathcal{O}(\Delta t^2,\Delta t\Delta x,\Delta t\Delta y)$$

$$\partial_t C + \nabla\cdot(\mathbf{q}C) = (\Delta x - q_x\Delta t)\frac{q_x}{2}\partial^2_{xx}C + (\Delta y - q_y\Delta t)\frac{q_y}{2}\partial^2_{yy}C$$
$$- q_xq_y\Delta t\partial^2_{xy}C + \mathcal{O}(\Delta t^2,\Delta t\Delta x,\Delta t\Delta y) \tag{2.19}$$

In the above equation, the first two right terms are diffusive ones only if their respective coefficients are positive, meaning that if $\Delta t \le \min\left(\frac{\Delta x}{|q_x|},\frac{\Delta y}{|q_y|}\right)$ the scheme will effectively diffuse the concentration through the domain. If in the contrary $\Delta t > \min\left(\frac{\Delta x}{|q_x|},\frac{\Delta y}{|q_y|}\right)$, the solution will condense toward its center prior to the direction in which the last inequality is realized.

isotropic term

The third term can be interpreted as an-isotropic diffusion in the direction normal to the flow, and anti-diffusion in the diretion of the flow . Since $\Delta t > 0$, this term never disappears, and its impact can be seen through the elongated shape of the concentration perpendicularly to the velocity, and flattened in the velocity direction (see fig. 2.4).

an-isotropic term

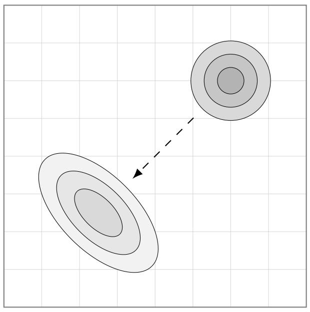

Figure 2.4: An-isotropic diffusion

2.4.3 Von Neumann analysis and the CFL condition

Let suppose that the solution takes the form $C_k^n = \phi_n\, e^{i(l\Delta x + m\Delta y)}$, with l, m two integers verifying $k = l + (m-1)N_x$. We are looking for a condition that will ensure us the stability of the scheme. This is the case whenever the modulus of C_k^n decreases as n increases, so the condition $|\phi_{n+1}| < |\phi_n|$ is satisfying. We split our problem in two part: first we will get such a condition only considering the advection, and then only by considering the diffusion. Injecting the new form of C_k^n in (2.16) leads to:

$$\begin{aligned}
&\frac{\phi_{n+1}-\phi_n}{\Delta t} e^{i(l\Delta x + m\Delta y)} \\
&\quad = + \frac{\phi_n}{\Delta x} e^{im\Delta y}\Big[\quad |q_{x\,|\gamma_{kk+1}}|(-e^{il\Delta x}\mathbb{1}_{q_x>0} + e^{i(l+1)\Delta x}\mathbb{1}_{q_x\le 0}) \\
&\qquad\qquad + |q_{x\,|\gamma_{kk-1}}|(e^{i(l-1)\Delta x}\mathbb{1}_{q_x>0} - e^{il\Delta x}\mathbb{1}_{q_x\le 0})\Big] \\
&\qquad + \frac{\phi_n}{\Delta y} e^{il\Delta x}\Big[\quad |q_{y\,|\gamma_{kk+N_x}}|(-e^{im\Delta y}\mathbb{1}_{q_y>0} + e^{i(m+1)\Delta y}\mathbb{1}_{q_y\le 0}) \\
&\qquad\qquad + |q_{y\,|\gamma_{kk-N_x}}|(e^{i(m-1)\Delta y}\mathbb{1}_{q_y>0} - e^{im\Delta y}\mathbb{1}_{q_y\le 0})\Big]
\end{aligned}$$

Multiplying both sides by $\Delta t / e^{i(l\Delta x + m\Delta y)}$ we find:

$$\begin{aligned}
\phi_{n+1} = \phi_n\Big[1 &+ \frac{\Delta t}{\Delta x}\,|q_{x\,|\gamma_{kk+1}}|(-\mathbb{1}_{q_x>0} + e^{i\Delta x}\mathbb{1}_{q_x\le 0}) \\
&+ \frac{\Delta t}{\Delta x}\,|q_{x\,|\gamma_{kk-1}}|(e^{-i\Delta x}\mathbb{1}_{q_x>0} - \mathbb{1}_{q_x\le 0}) \\
&+ \frac{\Delta t}{\Delta y}\,|q_{y\,|\gamma_{kk+N_x}}|(-\mathbb{1}_{q_y>0} + e^{i\Delta y}\mathbb{1}_{q_y\le 0}) \\
&+ \frac{\Delta t}{\Delta y}\,|q_{y\,|\gamma_{kk-N_x}}|(e^{-i\Delta y}\mathbb{1}_{q_y>0} - \mathbb{1}_{q_y\le 0})\Big]
\end{aligned}$$

As in the previous section, we make the assumption that $\mathbf{q}$ does not depend on the position. To further simplify the problem we shall suppose that $q_x, q_y > 0$ with no loss of generality. We then express the amplification factor A as in [3]:

$$A = \frac{\phi_{n+1}}{\phi_n} = \alpha + i\beta, \quad \text{with} \begin{cases} \alpha = 1 - Z_x(1-\cos(\Delta x)) - Z_y(1-\cos(\Delta y)) \\ \beta = -Z_x\sin(\Delta x) - Z_y\sin(\Delta y) \\ Z_x = q_x\frac{\Delta t}{\Delta x},\ Z_y = q_y\frac{\Delta t}{\Delta y} \end{cases}$$

If $Z_x + Z_y > 1$, we can find values of Δx and Δy for which the magnitude of A is strictly greater than 1, meaning that the scheme is unstable. We can for instance take $0 < \Delta x = \Delta y \leq 1$ and Δt great enough so that $Z_x + Z_y > 1$:

$$A = (1 - (Z_x + Z_y)(1 - \cos(\Delta x))) - i(Z_x + Z_y)\sin(\Delta x)$$

$$\begin{aligned}
\implies |A|^2 &= 1 - 2(Z_x + Z_y)(1 - \cos(\Delta x)) + (Z_x + Z_y)^2(1 - \cos(\Delta x))^2 \\
&\quad + (Z_x + Z_y)^2 \sin^2(\Delta x) \\
&= 1 - 2(Z_x + Z_y)(1 - \cos(\Delta x)) + 2(Z_x + Z_y)^2 - 2(Z_x + Z_y)^2 \cos(\Delta x) \\
&> 1 - 2(Z_x + Z_y)^2(1 - \cos(\Delta x)) + 2(Z_x + Z_y)^2 - 2(Z_x + Z_y)^2 \cos(\Delta x) \\
&> 1
\end{aligned}$$

If in the contrary $Z_x + Z_y \leq 1$ we prove that the scheme is stable. Noting $\mathbf{e}_x$, $\mathbf{e}_y$ and $\mathbf{x}$ the vectors defined as:

$$\mathbf{e}_x = \begin{pmatrix} \cos(\Delta x) \\ -\sin(\Delta x) \end{pmatrix}, \quad \mathbf{e}_y = \begin{pmatrix} \cos(\Delta y) \\ -\sin(\Delta y) \end{pmatrix}$$
$$\mathbf{x} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = Z_x \mathbf{e}_x + Z_y \mathbf{e}_y + \begin{pmatrix} 1 - Z_x - Z_y \\ 0 \end{pmatrix}$$

we can write the magnitude of A:

$$|A| = ||\mathbf{x}|| \leq Z_x + Z_y + |1 - Z_x - Z_y| = 1$$

Finally we obtain the scheme stability under the Courant–Friedrichs–Lewy condition (CFL):

$$\Delta t \leq \frac{\Delta x \Delta y}{\Delta y \max|q_x| + \Delta x \max|q_y|} \tag{2.20}$$

One can easily verify that this inequality is more restrictive than the previous condition (2.17) inherent to the flux-scheme, and assures us the isotropic diffusion established with the modified equation (2.19).

Now that we know a sufficient condition to have a stable scheme in an advection setting, we search a similar condition for a purely diffusive problem. We start from scheme (2.18) with all flux terms set to zero:

$$\begin{aligned}
&\frac{|\mathcal{C}_k|}{\Delta t}(\phi_{n+1} - \phi_n)e^{i(l\Delta x + m\Delta y)} = \sum_{s \in \mathcal{N}_k} |\gamma_{ks}| D\vec{\nabla} C_{|\gamma_{ks}} \cdot \mathbf{n}_{|\gamma_{ks}} \\
&= -\Delta x \Big(D_{21} \frac{C_{k+1}^n + C_{k+1-N_x}^n - C_{k-1}^n - C_{k-1-N_x}^n}{4\Delta x} + D_{22} \frac{C_k^n - C_{k-1}^n}{\Delta y} \Big) \\
&\quad + \Delta y \Big(D_{11} \frac{C_{k+1}^n - C_k^n}{\Delta x} + D_{12} \frac{C_{k+N_x}^n + C_{k+1+N_x}^n - C_{k-N_x}^n - C_{k+1-N_x}^n}{4\Delta y} \Big) \\
&\quad + \Delta x \Big(D_{21} \frac{C_{k+1}^n + C_{k+1+N_x}^n - C_{k-1}^n - C_{k-1+N_x}^n}{4\Delta x} + D_{22} \frac{C_{k+N_x}^n - C_k^n}{\Delta y} \Big) \\
&\quad - \Delta y \Big(D_{11} \frac{C_k^n - C_{k-1}^n}{\Delta x} + D_{12} \frac{C_{k+N_x}^n + C_{k-1+N_x}^n - C_{k-N_x}^n - C_{k-1-N_x}^n}{4\Delta y} \Big) \\
&= \frac{D_{12} + D_{21}}{4}(C_{k+1+N_x}^n + C_{k-1-N_x}^n - C_{k-1+N_x}^n - C_{k+1-N_x}) \\
&\quad + D_{22}\Delta x \frac{C_{k+N_x}^n - 2C_k^n + C_{k-N_x}^n}{\Delta y} + D_{11}\Delta y \frac{C_{k+1}^n - 2C_k^n + C_{k-1}^n}{\Delta x}
\end{aligned}$$

Dividing by $e^{i(l\Delta x+m\Delta y)}$ and factorizing by ϕ_n leads to:

$$\begin{aligned}
&\phi_{n+1}-\phi_n \\
&\quad=\frac{\Delta t}{\Delta x\Delta y}\,\phi_n\Big[D_{11}\frac{\Delta y}{\Delta x}(e^{i\Delta x}+e^{-i\Delta x}-2)+D_{22}\frac{\Delta x}{\Delta y}(e^{i\Delta y}+e^{-i\Delta y}-2)\\
&\qquad\qquad+\frac{D_{12}+D_{21}}{4}(e^{i(\Delta x+\Delta y)}+e^{-i(\Delta x+\Delta y)})\\
&\qquad\qquad-\frac{D_{12}+D_{21}}{4}(e^{i(-\Delta x+\Delta y)}+e^{i(\Delta x-\Delta y)})\Big]\\
&\phi_{n+1}=\phi_n\Big[1+2D_{11}\frac{\Delta t}{\Delta x^2}(\cos\Delta x-1)+2D_{22}\frac{\Delta t}{\Delta y^2}(\cos\Delta y-1)\\
&\qquad\qquad+\frac{\Delta t}{2\Delta x\Delta y}(D_{12}+D_{21})(\cos(\Delta x+\Delta y)-\cos(\Delta x-\Delta y))\Big]
\end{aligned}$$

By denoting $A=\phi_{n+1}/\phi_n$ the amplification factor, we are searching the conditions under which $|A|\leq 1$:

$$-2\leq\left\{\begin{array}{l}2D_{11}\frac{\Delta t}{\Delta x^2}(\cos\Delta x-1)+2D_{22}\frac{\Delta t}{\Delta y^2}(\cos\Delta y-1)\\ -\frac{\Delta t}{\Delta x\Delta y}(D_{12}+D_{21})\sin\Delta x\sin\Delta y\end{array}\right\}\leq 0$$

If $D_{12}=D_{21}=0$ we simply have:

$$-1\leq D_{11}\frac{\Delta t}{\Delta x^2}(\cos\Delta x-1)+D_{22}\frac{\Delta t}{\Delta y^2}(\cos\Delta y-1)$$

A sufficient condition to satisfy this inequality is to chose Δt such that:

$$2\Delta t\Big(\frac{D_{11}}{\Delta x^2}+\frac{D_{22}}{\Delta y^2}\Big)\leq 1\iff\Delta t\leq\frac{\Delta x^2\Delta y^2}{2D_{11}\Delta y^2+2D_{22}\Delta x^2}\tag{2.21}$$

Finally, in order to ensure the stability of the scheme disregarding the nature of the transport we can chose:

$$\Delta t=\frac{\alpha}{2}\min\left(\frac{\Delta x\Delta y}{\Delta y\max|q_x|+\Delta x\max|q_y|},\ \frac{\Delta x^2\Delta y^2}{2D_{11}\Delta y^2+2D_{22}\Delta x^2}\right)\text{ with }\alpha\in[0,1]$$

Appendix A

Properties of the diamond scheme

A.1 Existence and uniqueness of the solution

In order to prove the existence and uniqueness of the solution for the diamond scheme with a constant permeability matrix K, we introduce two discrete derivation operators d and $\mathbf{g}$, which respectively correspond to discretized vertions of the divergence and the gradient:

$$(\mathbf{g}h)_{k+1/2} = \begin{pmatrix} (h_{k+1} - h_k)/\Delta x \\ (h_k^+ - h_k^-)/\Delta y \end{pmatrix}, \quad (\mathbf{g}h)_{k+N_x/2} = \begin{pmatrix} (h_k^+ - h_{k-1}^+)/\Delta x \\ (h_{k+N_x} - h_k)/\Delta y \end{pmatrix} \tag{A.1}$$

$$(\mathrm{d}\mathbf{w})_k = \frac{w^1_{k+1/2} - w^1_{k-1/2}}{\Delta x} + \frac{w^2_{k+N_x/2} - w^2_{k-N_x/2}}{\Delta y} \quad \forall \mathbf{w} = \begin{pmatrix} w^1 \\ w^2 \end{pmatrix} \tag{A.2}$$

Using these operators we write for a given non-border cell $\mathcal{C}_k$:

$$\begin{aligned}(\mathrm{d}(K\mathbf{g}h))_k = \quad & K_{11}\frac{(\mathbf{g}h)^1_{k+1/2} - (\mathbf{g}h)^1_{k-1/2}}{\Delta x} + K_{12}\frac{(\mathbf{g}h)^2_{k+1/2} - (\mathbf{g}h)^2_{k-1/2}}{\Delta x} \\ & + K_{21}\frac{(\mathbf{g}h)^1_{k+N_x/2} - (\mathbf{g}h)^1_{k-N_x/2}}{\Delta y} + K_{22}\frac{(\mathbf{g}h)^2_{k+N_x/2} - (\mathbf{g}h)^2_{k-N_x/2}}{\Delta y}\end{aligned}$$

Multiplying by the squared surface of $\mathcal{C}_k$ we find:

$$\begin{aligned}|\mathcal{C}_k|^2(\mathrm{d}(K\mathbf{g}h))_k = \quad & K_{11}\Delta y^2(h_{k+1} - 2h_k + h_{k-1}) \\ & + K_{12}\Delta x\Delta y\Big((h_k^+ - h_k^-) - (h_{k-1}^+ - h_{k-1}^-)\Big) \\ & + K_{21}\Delta x\Delta y\Big((h_k^+ - h_{k-1}^+) - (h_k^- - h_{k-1}^-)\Big) \\ & + K_{22}\Delta x^2(h_{k+N_x} - 2h_k + h_{k-N_x})\end{aligned} \tag{A.3}$$

By comparing the right term of (A.3) with equation (2.11) we can deduce the following equality:

$$|\mathcal{C}_k|^2(\mathrm{d}(K\mathbf{g}h))_k = -2|\mathcal{C}_k|^2 f_{1,\mathcal{C}_k} \Longrightarrow (\mathrm{d}(K\mathbf{g}h))_k = -2f_{1,\mathcal{C}_k} \tag{A.4}$$

with $f_{1,\mathcal{C}_k} = \frac{1}{|\mathcal{C}_k|}\int_{\mathcal{C}_k} f_1 dx$.

It is interesting to draw the parallel between the discrete pressure equation and the continuous one:

Continuous equation (1.1a)	Discrete equation (A.4)
$-\nabla \cdot (K\vec{\nabla} h) = f_1$	$-(\mathrm{d}(K\mathbf{g}h))_k = 2f_{1,\mathcal{C}_k} \quad \forall \mathcal{C}_k$ non-border cell

Although these two equations are close, there is a factor two appearing in the right hand side of (A.4). This is due to the internal approximation that $\vec{\nabla} h$ is constant on each diamond cell, which leads us to approximate h on a border d_i of $\mathcal{D}_{kl}$ by taking the mean value of its trace on d_i: $h_{|d_i} \approx \frac{h_{|d_i \cap d_{i-1}} + h_{|d_i \cap d_{i+1}}}{2}$.

Proving the existence and uniqueness of the solution of (2.12) with Dirichlet conditions is equivalent to prove these properties for the new problem:

$$\begin{cases} -(\mathrm{d}(K\mathbf{g}h))_k = 2f_{1,\mathcal{C}_k} & \forall k \text{ such that } \mathcal{C}_k \text{ is not a border cell} \\ h_k = \sigma_k \text{ with } \sigma_k \text{ known} & \forall k \text{ such that } \mathcal{C}_k \text{ is a border cell} \end{cases} \tag{A.5}$$

discrete scalar products

For this purpose we then introduce two discrete scalar products on the truncated mesh denoted $\mathcal{T}$:

$$\langle u, v\rangle_{\mathcal{C}} = \sum_{\mathcal{C}_i \text{ non-border cell}} |\mathcal{C}_i| u_i v_i \tag{A.6}$$

$$\langle \mathbf{u}, \mathbf{v}\rangle_{\mathcal{D}} = \sum_{\mathcal{C}_i \text{ non-border cell}} \left(\sum_{k \in \mathcal{N}_i} |\mathcal{D}_{ik}| \mathbf{u}_{|\mathcal{D}_{ik}} \cdot \mathbf{v}_{|\mathcal{D}_{ik}} \right) \tag{A.7}$$

discrete Green formula

We can show the discrete Green formula:

$$\langle \mathrm{d}\mathbf{u}, v\rangle_{\mathcal{C}} = -\langle \mathbf{u}, \mathbf{g}v\rangle_{\mathcal{D}} + \beta \sum_{\mathcal{C}_i \text{ border cell}} v_i \mathbf{u}_i \cdot \mathbf{n}_i \tag{A.8}$$

with $\beta \in \mathbb{R}$ a constant that can be determined by calculation, and $\mathbf{n}_i$ the normal unit vector of the domain border at the level of cell $\mathcal{C}_i$ if $\mathcal{C}_i$ is not a corner, and the sum of the two corresponding normal unit vector at the level of $\mathcal{C}_i$ if it is a corner.

variational formulation

If in the problem (A.5) we have $f_{1,\mathcal{C}_k} = 0$ for all non border cell $\mathcal{C}_k$ and $\sigma_i = 0$ for all border cell $\mathcal{C}_i$, we necessarily have:

$$\forall (\omega_i)_i,\ -\langle \mathrm{d}(K\mathbf{g}h), \omega\rangle_{\mathcal{C}} = 0 \Longleftrightarrow \langle K\mathbf{g}h, \mathbf{g}\omega\rangle_{\mathcal{D}} = 0$$

Choosing $\omega = h$ and assuming that K is positive-definite for the usual scalar product we get:

$$\langle K\mathbf{g}h, \mathbf{g}h\rangle_{\mathcal{D}} = 0 \Longleftrightarrow (\mathbf{g}h)_{k\pm 1/2} = (\mathbf{g}h)_{k\pm N_x/2} = 0_{\mathbb{R}^2} \text{ for all non-border cell } \mathcal{C}_k$$

Since the Dirichlet conditions are homogeneous, we conclude that $h_k = 0$ for all index k. Therefore if the solution exists it is unique as soon as K is positive-definite. Because the system (A.5) is linear, the uniqueness of the solution is equivalent to its existence *for all second member* f_1.

Finally we have proved that if K is positive-definite, there is existence and uniqueness of the solution of (2.12) with Dirichlet conditions, and thus the rigidity matrix R is nonsingular.

Bibliography

[1] Coudière, Yves and Vila, Jean-Paul and Villedieu, Philippe. *Convergence rate of a finite volume scheme for a two dimensional convection-diffusion problem.* http://www.numdam.org/item?id=M2AN_1999__33_3_493_0. 1999.

[2] Scott A. Socolofsky and Gerhard H. Jirka. *Advective Diffusion Equation.* https://ceprofs.civil.tamu.edu/ssocolofsky/cven489/Book/Book.htm. 2004.

[3] Sophie Depeyre. *Etude de schémas d'ordre élevé en volumes finis pour des problèmes hyperboliques. Applications aux équations de Maxwell, dEuler et aux autres écoulements diphasiques dispersés.* https://pastel.archives-ouvertes.fr/tel-00005613. HAL Id: <tel-00005613>. 1997.

[4] Umberto Rizza and Daniela Buske and Bardo Bodmann and Marco T. Vilhena. *An advection-diffusion model for radioactive substance dispersion released from the Fukushima-Daiichi nuclear power plant.* https://periodicos.ufsm.br/cienciaenatura/article/viewFile/9392/5543. 2000.