LECOURTIER Frédérique 16 juin 2025

FEMs

Table des matières

1 Introduction

In the following, we will consider the Poisson problem with Dirichlet condition (homogeneous or non-homogeneous) :
Problem : Find $u : \Omega \to \mathbb{R}^d$ such that

$$\begin{cases} -\Delta u = f, & \text{in } \Omega, \\ u = g, & \text{on } \partial\Omega, \end{cases}$$

with Δ the Laplace operator and $\Omega \subset \mathbb{R}^d$ a smooth bounded open set (and $\partial\Omega$ its boundary).

2 Standard FEM

In this section, we will present the standard finite element method. We'll start by presenting some general notions of functional analysis, then explain the general principle of FEM. Then we'll give a few more details on the method and finish by describing the application to the Poisson problem (with Dirichlet condition). For more information, please refer to [**?**] and [**?**].

2.1 Some notions of functional analysis.

In this section, we'll recall some of the notions of functional analysis that will be used in the next sections. In particular, Lebesgue spaces and Sobolev spaces. Please refer to the book [**?**]. Let's consider Ω a smooth open-set of $\mathbb{R}^d$ $(d = 1, 2, 3)$ with boundary Γ.
We begin here by defining Lebesgue spaces :

Definition 2.1 (Lebesgue spaces)**.** *Lebesgue spaces, denoted L^p, are vector spaces of classes of functions whose exponent power p is integrable in the Lebesgue sense, where p is a strictly positive real number. They are defined by*

$$L^p(\Omega) = \left\{ u : \Omega \to \mathbb{R} | \int_\Omega u^p d\nu < +\infty \right\}$$

In particular, taking $p = 2$, we define the space

$$L^2(\Omega) = \left\{ u : \Omega \to \mathbb{R} | \int_\Omega u^2 d\nu < +\infty \right\}$$

which is the space of integrable square functions.

We also define Sobolev spaces of order 1 and order 2 :

Definition 2.2 (Sobolev spaces)**.** *The Sobolev space of order 1, denoted H^1, is defined by*

$$\begin{aligned} H^1(\Omega) &= \left\{ u \in L^2(\Omega) | \partial_{x_i} u \in L^2(\Omega) \right\} \\ &= \left\{ u \in L^2(\Omega), \nabla u \in L^2(\Omega)^d \right\} \end{aligned}$$

with the scalar product $\langle u, v \rangle_{H^1(\Omega)}$, defined by :

$$\langle u, v \rangle_{H^1(\Omega)} = \int_\Omega uv + \nabla u \cdot \nabla v, \forall u, v \in H^1(\Omega)$$

and the induced norm $|| \cdot ||_{H^1(\Omega)}$.
We also define the space

$$H_0^1(\Omega) = \left\{ u \in H^1(\Omega) | u_{|\Gamma} = 0 \right\}$$

The Sobolev space of order 2, denoted H^2, is defined by

$$H^2(\Omega) = \left\{ u, u', u'' \in L^2(\Omega) \right\}$$

with scalar product $\langle u, v \rangle_{H^2(\Omega)}$, defined by :

$$\langle u, v \rangle_{H^2(\Omega)} = \int_\Omega uv + u'v' + u''v'', \forall u, v \in H^1(\Omega)$$

and the induced norm $|| \cdot ||_{H^2(\Omega)}$.

Remarque. *In view of these definitions, we can see that*

$$||u||^2_{H^1(\Omega)} = ||u||^2_{L^2(\Omega)} + |u|^2_{H^1\Omega)}$$

with $|u|_{H^1\Omega)} = ||\nabla u||_{L^2(\Omega)}$ called H^1 semi-norm.
We also note that

$$||u||^2_{H^2(\Omega)} = ||u||^2_{L^2(\Omega)} + |u|^2_{H^1(\Omega)} + |u|^2_{H^2(\Omega)}$$

with $|u|_{H^2\Omega)} = ||\nabla^2 u||_{L^2(\Omega)}$ called H^2 semi-norm.

Remarque. *In the following, we will note $|| \cdot ||_{0,\Omega}$ the L^2 norm on Ω, $|| \cdot ||_{1,\Omega}$ the H^1 norm on Ω and $|| \cdot ||_{2,\Omega}$ the H^2 norm on Ω. We will also note $| \cdot |_{1,\Omega}$ the H^1 semi-norm on Ω and $| \cdot |_{2,\Omega}$ the H^2 semi-norm on Ω.*

2.2 General principle of the method

Let's consider a domain Ω whose boundary is denoted $\partial\Omega$. We seek to determine a function u defined on Ω, solution of a partial differential equation (PDE) for given boundary conditions.
The general approach of the finite element method is to write down the variational formulation of this PDE, thus giving us a problem of the following type :
Variational Problem :

$$\text{Find } u \in V \text{ such that } a(u,v) = l(v),\ \forall v \in V$$

where V is a Hilbert space, a is a bilinear form and l is a linear form.
To do this, we multiply the PDE by a test function $v \in V$, then integrate over $L^2(\Omega)$.
The idea of FEM is to use Galerkin's method. We then look for an approximate solution u_h in V_h, a finite-dimensional subspace dependent on a positive parameter h such that

$$V_h \subset V, \quad \dim V_h = N_h < \infty, \quad \forall h > 0.$$

The variational problem can then be approached by :
Approach Problem :

$$\text{Find } u_h \in V_h \text{ such that } a(u_h, v_h) = l(v_h),\ \forall v_h \in V.$$

As V_h is of finite dimension, we can consider a basis $(\varphi_1, \ldots, \varphi_{N_h})$ of V_h and thus decompose u_h on this basis as :

$$u_h = \sum_{i=1}^{N_h} u_i \varphi_i \tag{1}$$

By bilinearity of a, the approached problem is then rewritten as

$$\text{Find } u_1, \ldots, u_{N_h} \text{ such that } \sum_{i=1}^{N_h} u_i a(\varphi_i, v_h) = l(v_h),\ \forall v_h \in V$$

which is equivalent to

$$\text{Find } u_1, \ldots, u_{N_h} \text{ such that } \sum_{i=1}^{N_h} u_i a(\varphi_i, \varphi_j) = l(\varphi_j),\ \forall j \in \{1, \ldots, N_h\}$$

Thus, to find an approximation to the solution of the PDE, we simply solve the following linear system :

$$AU = b$$

with

$$A = (a(\varphi_i, \varphi_j))_{1 \le i,j \le N_h}, \quad U = (u_i)_{1 \le i \le N_h} \quad \text{and} \quad b = (l(\varphi_j))_{1 \le j \le N_h}$$

Remarque. *To impose Dirichlet boundary conditions, we can use one of 2 methods. The elimination method consists in modifying the rows associated with the boundary nodes in the finite element matrix. More precisely, we set the rows to 0 except 1 on the diagonal and the value of the condition on the second member. In other words, we simply write the value of the degrees of freedom at the Dirichlet boundary. The penalization method consists in modifying the matrix and the second member as follows :*

$$A_{i,i} := A_{i,i} + \frac{1}{\epsilon}$$
$$f_i := f_i + \frac{1}{\epsilon} g_i$$

with $\epsilon > 0$ and i is a boundary nodes.

2.3 Some details on FEM

After having seen the general principle of FEM, it remains to define the V_h spaces and the $\{\varphi_i\}$ basis functions.

Remarque. *The choice of V_h space is fundamental to have an efficient method that gives a good approximation u_h of u. In particular, the choice of the $\{\varphi_i\}$ basis of V_h influences the structure of the A matrix in terms of its sparsity and its condition number but also affects the quality of the approximation.*

To do this, we'll need several notions, which will be detailed in the following sections. First, we'll need to generate a **mesh** of our Ω domain. This will enable us to solve the PDE discretely at selected points. This is where the notion of **finite Lagrange elements** comes in. The properties of these elements, particularly in terms of their **affine family of finite elements**, is a key point of the method, which will enable us to bring each element of the mesh back to a **reference element** by using a **geometric transformation**. To describe these steps, we'll need to know 2 basic concepts : the **unisolvence** principle and the definitions of the **polynomial spaces** used ($\mathbb{P}_k$ and $\mathbb{Q}_k$).

2.3.1 Unisolvance

Definition 2.3. *Let $\Sigma = \{a_1, \dots, a_N\}$ be a set of N distinct points of $\mathbb{R}^n$. Let P be a finite-dimensional vector space of $\mathbb{R}^n$ functions taking values in $\mathbb{R}$. We say that Σ is P-unisolvent if and only if for all real $\alpha_1, \dots, \alpha_N$, there exists a unique element p of P such that $p(a_i) = \alpha_i, i = 1, \dots, N$. This means that the function*

$$\begin{aligned} L \ : \ & P \to \mathbb{R}^N \\ & p \mapsto (p(a_1), \dots, p(a_N)) \end{aligned}$$

is bijective.

Remarque. *In practice, to show that Σ is P-unisolvent, we simply check that $\dim P = card\ (\Sigma)$ and then prove the injectivity or surjectivity of L. The injectivity of L is demonstrated by showing that the only function of P that annuls on all points of Σ is the null function. The surjectivity of L is shown by identifying a family $p_1, \dots, p_N$ of elements of P such that $p_i(a_j) = \delta_{ij}$. Given real $\alpha_1, \dots, \alpha_N$, the function $p = \sum_{i=1}^{N} \alpha_i p_i$ then verifies $p(a_j) = \alpha_j, j = 1 \dots, N$.*

Remarque. *We call local basis functions of element K the N functions $p_1, \dots, p_N$ of P such that*

$$p_i(a_j) = \delta_{ij}, \quad 1 \leq i, j \leq N$$

2.3.2 Polynomial space

Let $\mathbb{P}_k$ be the vector space of polynomials of total degree less than or equal to k.

- In $\mathbb{R}$: $\mathbb{P}_k = \text{Vect}\{1, X, \dots, X^k\}$ and $\dim \mathbb{P}_k = k + 1$
- In $\mathbb{R}^2$: $\mathbb{P}_k = \text{Vect}\{X^i Y^j, 0 \leq i + j \leq k\}$ and $\dim \mathbb{P}_k = \frac{(k+1)(k+2)}{2}$
- In $\mathbb{R}^3$: $\mathbb{P}_k = \text{Vect}\{1, X^i Y^j Z^l, 0 \leq i + j + l \leq k\}$ and $\dim \mathbb{P}_k = \frac{(k+1)(k+2)(k+3)}{6}$

Let $\mathbb{Q}_k$ be the vector space of polynomials of degree less than or equal to k with respect to each variable.

- In $\mathbb{R}$: $\mathbb{Q}_k = \mathbb{P}_k$.
- In $\mathbb{R}^2$: $\mathbb{Q}_k = \text{Vect}\{X^i Y^j, 0 \leq i, j \leq k\}$ and $\dim \mathbb{Q}_k = (k + 1)^2$
- In $\mathbb{R}^3$: $\mathbb{Q}_k = \text{Vect}\{1, X^i Y^j Z^l, 0 \leq i, j, l \leq k\}$ and $\dim \mathbb{Q}_k = (k + 1)^3$

Remarque. *In practice, we will use the $\mathbb{P}^k$ family for triangles/tetrahedra and $\mathbb{Q}_k$ for quadrilaterals.*

2.3.3 Finite Lagrange Element

The most classic and simplest type of finite element is the Lagrange finite element.

Definition 2.4 (Lagrange Finite Element)**.** *A finite Lagrange element is a triplet (K, Σ, P) such that*

- *K is a geometric element of $\mathbb{R}^n$ ($n = 1, 2$ or 3), compact, connected and of non-empty interior.*
- *$\Sigma = \{a_1, \dots, a_N\}$ is a finite set of N distinct points of K.*
- *P is a finite-dimensional vector space of real functions defined on K and such that Σ is P-unisolvent (so $\dim P = N$).*

Example. *Let K be the segment $[a_1, a_2]$. Let's show that $\Sigma = \{a_1, a_2\}$ is P-unisolvent for $P = \mathbb{P}^1$. Since $\{1, x\}$ is a base of $\mathbb{P}^1$, we have $\dim P = card\ \Sigma = 2$.*
Moreover, we can write $p_i = \alpha_i x + \beta_i, i = 1, 2$. Thus

$$\begin{cases} p_1(a_1) = 1 \\ p_1(a_2) = 0 \end{cases} \iff \begin{cases} \alpha_1 a_1 + \beta_1 = 1 \\ \alpha_1 a_2 + \beta_1 = 0 \end{cases} \iff \begin{cases} \alpha_1 = \dfrac{1}{a_1 - a_2} \\ \beta_1 = -\dfrac{a_2}{a_1 - a_2} \end{cases}$$

and

$$\begin{cases} p_2(a_1) = 0 \\ p_2(a_2) = 1 \end{cases} \iff \begin{cases} \alpha_2 a_1 + \beta_2 = 0 \\ \alpha_2 a_2 + \beta_2 = 1 \end{cases} \iff \begin{cases} \alpha_1 = \dfrac{1}{a_2 - a_1} \\ \beta_1 = -\dfrac{a_1}{a_2 - a_1} \end{cases}$$

Thus

$$p_1(x) = \frac{x - a_2}{a_1 - a_2} \quad and \quad p_2(x) = \frac{x - a_1}{a_2 - a_1}$$

We deduce the surjectivity of L and Σ is $\mathbb{P}^1$-unisolvent.
Thus (K, Σ, P) is a Lagrange Finite Element.

Definition 2.5. *Two finite elements $(\hat{K}, \hat{\Sigma}, \hat{P})$ and (K, Σ, P) are affine-equivalent if and only if there exists an irreversible affine function F such that*

- $K = F(\hat{K})$
- $a_i = F(\hat{a}_i), i = 1, \ldots, N$
- $P = \{\hat{p} \circ F^{-1}, \hat{p} \in \hat{P}\}$.

*We then call an **affine family of finite elements** a family of finite elements, all affine-equivalent to the same element $(\hat{K}, \hat{\Sigma}, \hat{P})$, called the **reference element**.*

Remarque. *Let $(\hat{K}, \hat{\Sigma}, \hat{P})$ and (K, Σ, P) be two affine-equivalent finite elements, via an F transformation. Let $\hat{p}_i$ be the local basis functions on $\hat{K}$. Then the local basis functions on K are $p_i = \hat{p}_i \circ F^{-1}$.*

Remarque. *In practice, working with an affine family of finite elements means that all integral calculations can be reduced to calculations on the reference element.*
The reference elements in 1D, 2D triangular and 3D tetrahedral are :

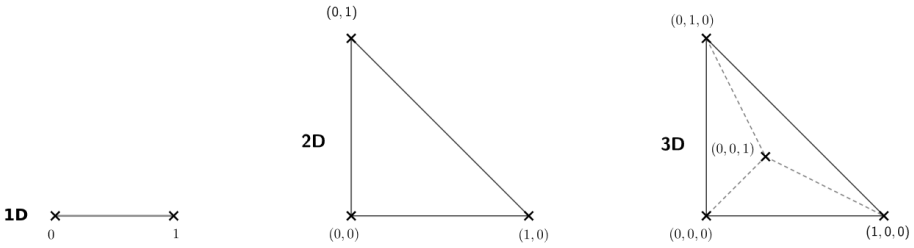

FIGURE 1 – Example of reference Elements.

2.3.4 Mesh

In 1D, the construction of a mesh consists in creating a subdivision of the interval $[a, b]$. We can extend this definition in 2D and 3D by considering that a mesh is formed by a family of elements $\mathcal{T}_h = \{K_1, \ldots, K_{N_e}\}$ (see Fig 2) where N_e is the number of elements.
In 2D, these elements can be triangles or rectangles. In 3D, they can be tetrahedrons, parallelepipeds or prisms.

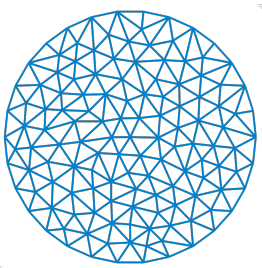

FIGURE 2 – Example of a triangular mesh on a circles.

Remarque. *Note that it's important to have a certain geometric quality in the mesh, as this can influence the accuracy of the approximation. For example, if we're using triangles as 2D elements, it's preferable that all the elements in the mesh are not too flattened.*

2.3.5 Construction of V_h space

Geometric transformation : A mesh is generated by

- A reference element noted $\hat{K}$.
- A family of geometric transformations mapping $\hat{K}$ to the elements $K_1, \dots, K_{N_e}$. Thus, for a cell $K \in \mathcal{T}_h$, we denote T_K the geometric transformation mapping $\hat{K}$ to K :

$$T_K : \hat{K} \to K$$

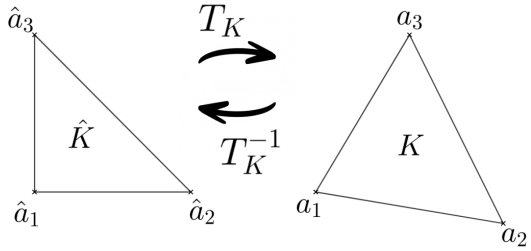

FIGURE 3 – Geometric transformation applied to a triangle.

Let $(\hat{K}, \hat{\Sigma}, \hat{P})$ be the finite reference element with

- the degrees of freedom of the reference element $\hat{K}$: $\hat{\Sigma} = \{\hat{a}_1, \dots, \hat{a}_{n_f}\}$ with n_f the number of degrees of freedom.
- the local basis functions of $\hat{K}$: $\{\hat{\psi}_1, \dots, \hat{\psi}_{n_f}\}$ (also called form functions)

So for each $K \in \mathcal{T}_h$, we consider a tuple $\{a_{K,1}, \dots, a_{K,n_f}\}$ (degrees of freedom) and the associated geometric transformation is defined by :

$$T_K : \hat{x} \mapsto \sum_{i=1}^{n_f} a_{K,i} \hat{\psi}_i(\hat{x})$$

In particular, we have

$$T_K(\hat{a}_i) = a_{K,i}, \quad i = 1, \dots, n_f$$

Remarque. *In particular, if the form functions are affine, the geometric transformations will be too. This is an interesting property, as the gradient of these geometric transformations will be constant.*

Remarque. *In the following, we will assume that these transformations are C^1-diffeomorphisms (i.e. the transformation and its inverse are C^1 and bijective).*

Construction of the basis (φ_i) of V_h :
For each $K \in \mathcal{T}_h$, let (K, Σ, P) be an finite element with

- the degrees of freedom of the element K : $\Sigma = \{a_{K,i} = T_K(\hat{a}_i), i = 1, \dots, n_f\}$

- the local basis functions of K : $\{\psi_{K,i} = \hat{\psi}_i \circ T_K^{-1}, i = 1, \ldots, n_f\}$ (because $(\hat{K}, \hat{\Sigma}, \hat{P})$ and (K, Σ, P) are affine-equivalent).

By noting $\{a_1, \ldots, a_{N_f}\} = \bigcup_{K \in \mathcal{T}_h} \{a_{K,1}, \ldots, a_{K,n_f}\}$ with N_f the total number of degrees of freedom (over all the geometry), we have

$$\forall j \in \{1, \ldots, N_f\}, \quad \varphi_{j|K} = \psi_{K,a_{K,j}}$$

The ϕ_j functions are then in the space of piece-wise affine continuous functions, defined by

$$P^k_{C,h} = \{v_h \in C^0(\bar{\Omega}), \forall K \in \mathcal{T}_h, v_{h|K} \in \mathbb{P}_k\} \subset H^1(\Omega)$$

In fact, the functions $\{\varphi_1, \ldots, \varphi_{N_f}\}$ form a basis of $P^k_{C,h}$ and so we can choose $V_h = P^k_{C,h}$.

2.4 Application to the Poisson problem

Weak formulation :
We want to apply the standard FEM method to the Poisson problem with Dirichlet boundary condition under consideration. Let's start by writing the variational formulation of the problem. For the moment, we have the following strong formulation of the problem :

$$-\Delta u = f \text{ on } \Omega$$

Multiplying by a test function $v \in H^1_0(\Omega)$ and integrating over Ω, we obtain

$$-\int_\Omega \Delta u v = \int_\Omega f v.$$

By integration by parts, we have

$$-\int_\Omega \Delta u v = \int_\Omega \nabla u \cdot \nabla v - \int_\Gamma \frac{\partial u}{\partial n} v.$$

This leads to the following weak formulation

$$\text{Find } u \in H^1_0(\Omega) \text{ such that } a(u, v) = l(v), \ \forall v \in H^1_0(\Omega)$$

with

$$\begin{cases} a(u, v) = \displaystyle\int_\Omega \nabla u \cdot \nabla v \\ l(v) = \displaystyle\int_\Omega f v \end{cases}$$

because $v \in H^1_0(\Omega)$.

Uniqueness of the solution :
An important result of the FEM method is the following theorem, which shows the uniqueness of the solution :

Proposition 2.1 (Lax-Milgram)**.** *Let a be a continuous, coercive bilinear form on V and l a continuous, linear form on V. Then the variational problem has a unique solution $u \in V$.*
Moreover, if the bilinear form is symmetrical, u is a solution to the following minimization problem :

$$J(u) = \min_{v \in V} J(v), \quad J(v) = \frac{1}{2}a(v,v) - l(v)$$

Let's show that the Poisson problem with Dirichlet boundary condition has a unique weak solution $u \in H_0^1(\Omega)$.

- It's easy to see that a is a bilinear (and symmetrical) form.
  Let's show that a is continuous. Let $u, v \in H_0^1(\Omega)$, then

$$\begin{aligned} |a(u,v)| &= \left| \int_\Omega \nabla u \cdot \nabla v \right| = |\langle u, v \rangle_{H^1(\Omega)} \\ &\leq ||u||_{H^1(\Omega)} ||v||_{H^1(\Omega)} \quad \text{by Cauchy-Schwarz} \end{aligned}$$

  Let's show that a is coercive. Let $u \in H_0^1(\Omega)$, then

$$\begin{aligned} a(u,u) &= \int_\Omega \nabla u \cdot \nabla u = \int_\Omega |\nabla u|^2 \\ &= \frac{1}{2}\int_\Omega |\nabla u|^2 + \frac{1}{2}\int_\Omega |\nabla u|^2 \\ &\geq \frac{1}{2}\alpha \int_\Omega u^2 + \frac{1}{2}\int_\Omega |\nabla u|^2 \quad \text{by Poincaré} \\ &\geq \alpha \int_\Omega u^2 + |\nabla u|^2 = \alpha ||u||_{H^1(\Omega)} \end{aligned}$$

- It is easy to see that l is a linear form.
  Let's show that l is continuous. Let $v \in H_0^1(\Omega)$, then

$$\begin{aligned} |l(v)| &= \left| \int_\Omega fv \right| = |\langle f, v \rangle_{L^2(\Omega)}| \\ &\leq ||f||_{L^2(\Omega)} ||v||_{L^2(\Omega)} \quad \text{by Cauchy-Schwarz} \\ &\leq ||f||_{L^2(\Omega)} ||v||_{H^1(\Omega)} \end{aligned}$$

By the Lax-Milgram theorem, we deduce that the Poisson problem with Dirichlet boundary condition has a unique weak solution $u \in H_0^1(\Omega)$.

3 ϕ-FEM

In this section, we will present the ϕ-FEM method. We will first present fictitious domain methods (Section 3.1). Next, we will give a general presentation of the method with a description of the spaces required (Section 3.2), followed by a description of the ϕ-FEM direct method (Section 3.3) and a description of the ϕ-FEM dual method (Section 3.4). Finally, we will give some details on the stabilization terms (Section 3.5).

3.1 Fictitious domain methods

The method we are interested in, called the ϕ-FEM method, is a fictitious domain method, i.e. it does not require a mesh conforming to the real boundary. In the context of augmented surgery, fictitious domain methods presents a considerable advantage in comparison to standard FEM approaches. During real-time simulation, the geometry (in our specific context, an organ such as the liver, for example) can deform over time. Methods such as standard FEM, which requires a mesh fitted to the boundary, necessitate a complete remeshing of the geometry at each time step (Figure 4). Unlike this type of method, fictitious domain methods requires only the generation of a single mesh : the mesh of a fictitious domain containing the entire geometry (Figure 5).

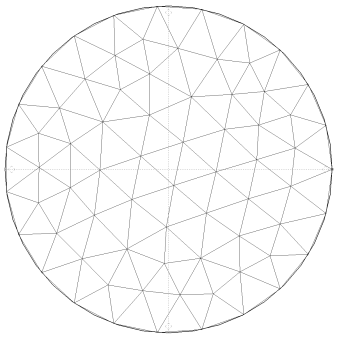

FIGURE 4 – Standard FEM mesh example.

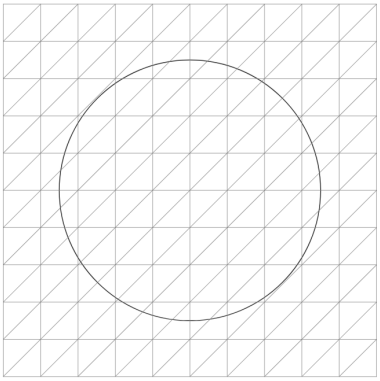

FIGURE 5 – Fictitious domain methods mesh example.

Application to the ϕ-FEM method :
In the case of the ϕ-FEM Method, as the boundary of the geometry is represented by a level-set function ϕ, only this function will change over time, which is a real time-saver.
For the purposes of this internship, the geometries considered are not organs (such as the liver), because these are complex geometries. We are considering simpler geometries such as circles or squares.
It is also important to note that the ϕ-FEM method has a considerable advantage : by constructing a fictitious mesh around the domain, we can generate a Cartesian mesh. This type of mesh can easily be represented by matrices, in the same way as images, hence the possibility of teaching these ϕ-FEM solutions to an FNO who generally works on images. A paper in progress presents results with the combination of ϕ-FEM and an FNO on more complex geometries, notably ellipses.

3.2 General presentation of the ϕ-FEM method

In this section, we will present the ϕ-FEM method. We consider the case of the Poisson problem with homogeneous Dirichlet boundary conditions [**?**].

$$\begin{cases} -\Delta u = f, & \text{in } \Omega, \\ u = g, & \text{on } \partial\Omega, \end{cases} \tag{2}$$

where the domain Ω and its boundary Γ are given by a level-set function ϕ such that

$$\Omega = \{\phi < 0\} \quad \text{and} \quad \Gamma = \{\phi = 0\}.$$

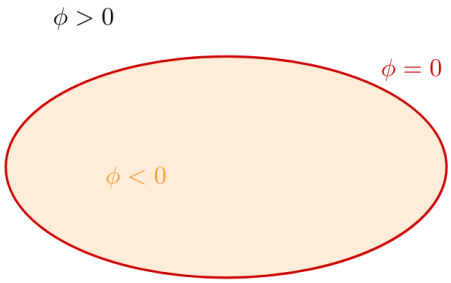

FIGURE 6 – Definition of the level-set function.

Remarque. *For more details on mesh assumptions, convergence results and finite element matrix condition number, please refer to [?]. ϕ-FEM schemes for the Poisson problem with Neumann or mixed (Dirichlet and Neumann) conditions are presented in [?, ?]. The ϕ-FEM scheme can also be found for other PDEs, including linear elasticity [?, Chapter 2], the heat equation [?, Chapter 5] and the Stokes problem [?].*

Example. *If Ω is a circle of center A of coordinates (x_A, y_A) and radius r, a level-set function can be defined by*

$$\phi(x,y) = -r^2 + (x-x_A)^2 + (y-y_A)^2.$$

If Ω is an ellipse with center A of coordinates (x_A, y_A) and parameters (a,b), a level-set function can be defined by

$$\phi(x,y) = -1 + \frac{(x-x_A)^2}{a^2} + \frac{(y-y_A)^2}{b^2}.$$

We assume that Ω is inside a domain $\mathcal{O}$ and we introduce a simple quasi-uniform mesh $\mathcal{T}_h^{\mathcal{O}}$ on $\mathcal{O}$ (Figure 7).
We introduce now an approximation $\phi_h \in V_{h,\mathcal{O}}^{(l)}$ of ϕ given by $\phi_h = I_{h,\mathcal{O}}^{(l)}(\phi)$ where $I_{h,\mathcal{O}}^{(l)}$ is the standard Lagrange interpolation operator on

$$V_{h,\mathcal{O}}^{(l)} = \left\{ v_h \in H^1(\mathcal{O}) : v_{h|_T} \in \mathbb{P}_l(T) \; \forall T \in \mathcal{T}_h^{\mathcal{O}} \right\}$$

and we denote by $\Gamma_h = \{\phi_h = 0\}$, the approximate boundary of Γ (Figure 8).
We will consider $\mathcal{T}_h$ a sub-mesh of $\mathcal{T}_h^{\mathcal{O}}$ obtained by removing the elements located entirely outside Ω (Figure 8). To be more specific, $\mathcal{T}_h$ is defined by

$$\mathcal{T}_h = \left\{ T \in \mathcal{T}_h^{\mathcal{O}} : T \cap \{\phi_h < 0\} \neq \emptyset \right\}.$$

We denote by Ω_h the domain covered by the $\mathcal{T}_h$ mesh (Ω_h will be slightly larger than Ω) and $\partial\Omega_h$ its boundary (Figure 8). The domain Ω_h is defined by

$$\Omega_h = \left(\cup_{T \in \mathcal{T}_h} T \right)^O.$$

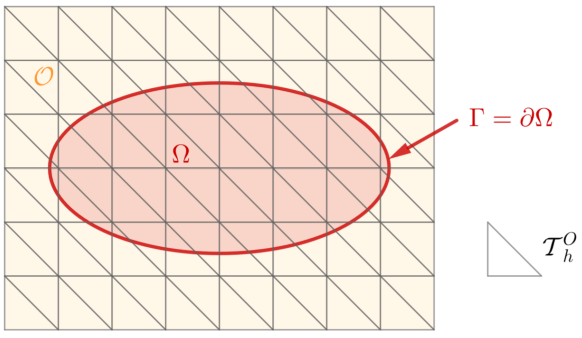

FIGURE 7 – Fictitious domain.

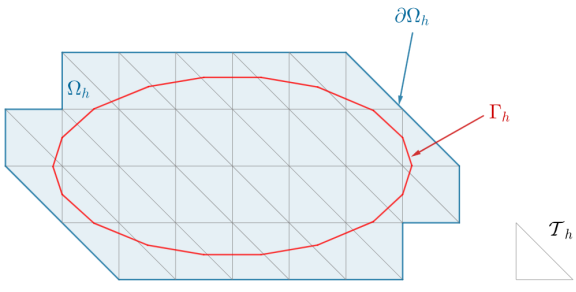

FIGURE 8 – Domain considered.

Now, we can introduce $\mathcal{T}_h^\Gamma \subset \mathcal{T}_h$ (Figure 9) which contains the mesh elements cut by the approximate boundary $\Gamma_h = \{\phi_h = 0\}$, i.e.

$$\mathcal{T}_h^\Gamma = \{ T \in \mathcal{T}_h : T \cap \Gamma_h \neq \emptyset \},$$

and $\mathcal{F}_h^\Gamma$ (Figure 10) which collects the interior facets of the mesh $\mathcal{T}_h$ either cut by Γ_h or belonging to a cut mesh element

$$\mathcal{F}_h^\Gamma = \{ E \text{ (an internal facet of } \mathcal{T}_h) \text{ such that } \exists T \in \mathcal{T}_h : T \cap \Gamma_h \neq \emptyset \text{ and } E \in \partial T \}.$$

We denote by Ω_h^Γ the domain covered by the $\mathcal{T}_h^\Gamma$ mesh (Figure 9) and also defined by

$$\Omega_h^\Gamma = \left(\cup_{T\in\mathcal{T}_h^\Gamma} T\right)^O.$$

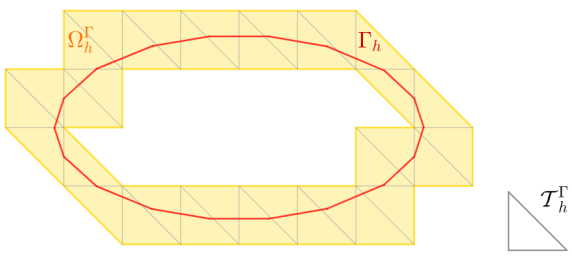

FIGURE 9 – Boundary cells.

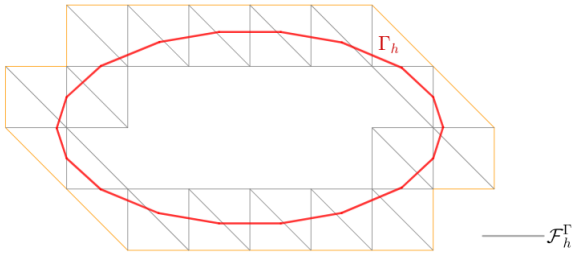

FIGURE 10 – Boundary edges.

3.3 Description of the ϕ-FEM direct method

As with standard FEM, the general idea behind ϕ-FEM is to find a weak solution (i.e. a solution to the variational problem) to the considered problem (2). The main difference lies in the spaces considered. In fact, we are no longer looking to solve the problem on Ω (of boundary Γ) but on Ω_h (of boundary $\partial\Omega_h$). Since our boundary conditions are defined on Γ, we don't have a direct condition on the $\partial\Omega_h$ boundary, so we will have to add terms to the variational formulation of the problem, called stabilization terms.
Let's first consider the homogeneous case, then assuming that the source term f is currently well-defined on Ω_h and that the solution u can be extended on Ω_h such that $-\Delta u = f$ on Ω_h, we can introduce a new unknown $w \in H^1(\Omega_h)$ such that $u = \phi w$ and the boundary condition on Γ is satisfied (since $\phi = 0$ on Γ). After an integration by parts, we have

$$\int_{\Omega_h} \nabla(\phi w)\cdot\nabla(\phi v) - \int_{\partial\Omega_h} \frac{\partial}{\partial n}(\phi w)\phi v = \int_{\Omega_h} f\phi v, \quad \forall v \in H^1(\Omega_h).$$

Remarque. *Note that Ω_h is constructed using ϕ_h and therefore implicitly depends on ϕ.*

Given an approximation ϕ_h of ϕ on the mesh $\mathcal{T}_h$, as defined in Section 3.2, and a finite element space V_h on $\mathcal{T}_h$, we can then search for $w_h \in V_h$ such that

$$a_h(w_h, v_h) = l_h(v_h), \quad \forall v_h \in V_h.$$

The bilinear form a_h and the linear form l_h are defined by

$$a_h(w,v) = \int_{\Omega_h} \nabla(\phi_h w)\cdot\nabla(\phi_h v) - \int_{\partial\Omega_h} \frac{\partial}{\partial n}(\phi_h w)\phi_h v + G_h(w,v)$$

and

$$l_h(v) = \int_{\Omega_h} f\phi_h v + G_h^{rhs}(v)$$

with

$$G_h(w,v) = \sigma h \sum_{E\in\mathcal{F}_h^\Gamma} \int_E \left[\frac{\partial}{\partial n}(\phi_h w)\right]\left[\frac{\partial}{\partial n}(\phi_h v)\right] + \sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T \Delta(\phi_h w)\Delta(\phi_h v)$$

and

$$G_h^{rhs}(v) = -\sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T f\Delta(\phi_h v).$$

with σ an independent parameter of h, which we'll call the stabilization parameter.
We can consider the finite element space $V_h = V_h^{(k)}$ with

$$V_h^{(k)} = \left\{v_h \in H^1(\Omega_h) : v_{h|_T} \in \mathbb{P}_k(T)\ \forall T \in \mathcal{T}_h\right\}.$$

Remarque. *Note that* $[\,\cdot\,]$ *is the jump on the interface* E *defined by*

$$\left[\frac{\partial}{\partial n}(\phi_h w)\right] = \nabla(\phi_h w)^+ \cdot n - \nabla(\phi_h w)^- \cdot n$$

with n *is the unit normal vector outside* E*.*

In the case of a non-homogeneous Dirichlet condition, we want to impose $u = g$ on Γ. With the direct method, we must suppose that g is currently given over the entire Ω_h and not just over Γ. We can then write the solution u as

$$u = \phi w + g, \text{ on } \Omega_h.$$

It can then be injected into the weak formulation of the homogeneous problem and we can then search for w_h on Ω_h such that

$$\begin{aligned}\int_{\Omega_h} \nabla(\phi_h w_h)\nabla(\phi_h v_h) - \int_{\partial\Omega_h} \frac{\partial}{\partial n}(\phi_h w_h)\phi_h v_h + G_h(w_h, v_h) &= \int_{\Omega_h} f\phi_h v_h \\ &- \int_{\Omega_h} \nabla g \nabla(\phi_h v_h) + \int_{\partial\Omega_h} \frac{\partial g}{\partial n}\phi_h v_h + G_h^{rhs}(v_h),\ \forall v_h \in \Omega_h\end{aligned}$$

with

$$G_h(w, v) = \sigma h \sum_{E\in\mathcal{F}_h^\Gamma} \int_E \left[\frac{\partial}{\partial n}(\phi_h w)\right]\left[\frac{\partial}{\partial n}(\phi_h v)\right] + \sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T \Delta(\phi_h w)\Delta(\phi_h v)$$

and

$$G_h^{rhs}(v) = -\sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T f\Delta(\phi_h v) - \sigma h \sum_{E\in\mathcal{F}_h^\Gamma} \int_E \left[\frac{\partial g}{\partial n}\right]\left[\frac{\partial}{\partial n}(\phi_h v)\right] - \sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T \Delta g \Delta(\phi_h v)$$

3.4 Description of the ϕ-FEM dual method

The idea here is the same as for the direct method, but with the dual method, we assume that g is defined on Ω_h^Γ and not on Ω_h. We then introduce a new unknown p on Ω_h^Γ in addition to the unknown u on Ω_h and so we aim to impose

$$u = \phi p + g, \text{ on } \Omega_h^\Gamma.$$

So we look for u on Ω_h and p on Ω_h^Γ such that

$$\begin{aligned}\int_{\Omega_h} \nabla u \nabla v - \int_{\partial\Omega_h} \frac{\partial u}{\partial n} v + \frac{\gamma}{h^2} \sum_{T\in\mathcal{T}_h^\Gamma} \int_T \left(u - \frac{1}{h}\phi p\right)\left(v - \frac{1}{h}\phi q\right) + G_h(u, v) &= \int_{\Omega_h} fv \\ &+ \frac{\gamma}{h^2} \sum_{T\in\mathcal{T}_h^\Gamma} \int_T g\left(v - \frac{1}{h}\phi q\right) + G_h^{rhs}(v),\ \forall v \text{ on } \Omega_h,\ q \text{ on } \Omega_h^\Gamma.\end{aligned}$$

with γ an other positive stabilization parameter,

$$G_h(u, v) = \sigma h \sum_{E\in\mathcal{F}_h^\Gamma} \int_E \left[\frac{\partial u}{\partial n}\right]\left[\frac{\partial v}{\partial n}\right] + \sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T \Delta u \Delta v$$

and

$$G_h^{rhs}(v) = -\sigma h^2 \sum_{T\in\mathcal{T}_h^\Gamma} \int_T f\Delta v.$$

Remarque. *The factors* $\frac{1}{h}$ *and* $\frac{1}{h^2}$ *control the condition number of the finite element matrix. For more details, please refer to the article [?].*

Remarque. *In the context of this internship, we won't be concerned with the choice of stabilization parameters* σ *and* γ*. We'll always take* $\sigma = 20$ *and* $\gamma = 1$*, but it's important to note that they can have a significant influence on the results.*

3.5 Some details on the stabilization terms

In this section, we will give some informations on stabilization terms. As introduced previously, the stabilization terms are intended to reduce the errors created by the "fictitious" boundary, but they also have the effect of ensuring the correct condition number of the finite element matrix and permitting to restore the coercivity of the bilinear scheme.
The first term of $G_h(w, v)$ defined by

$$\sigma h \sum_{E \in \mathcal{F}_h^\Gamma} \int_E \left[\frac{\partial}{\partial n}(\phi_h w)\right] \left[\frac{\partial}{\partial n}(\phi_h v)\right]$$

is a first-order stabilization term. This stabilization term is based on [**?**]. It also ensures the continuity of the solution by penalizing gradient jumps.
By subtracting $G_h^{rhs}(v)$ from the second term of $G_h(w, v)$, i.e.

$$\sigma h^2 \sum_{T \in \mathcal{T}_h^\Gamma} \int_T \Delta(\phi_h w)\Delta(\phi_h v) + \sigma h^2 \sum_{T \in \mathcal{T}_h^\Gamma} \int_T f\Delta(\phi_h v),$$

which can be rewritten as

$$\sigma h^2 \sum_{T \in \mathcal{T}_h^\Gamma} \int_T (\Delta(\phi_h w) + f)\, \Delta(\phi_h v),$$

we recognize the strong formulation of the Poisson problem. This second-order stabilization term penalizes the scheme by requiring the solution to verify the strong form on Ω_h^Γ. In fact, this term cancels out if $\phi_h w$ is the exact solution of the Poisson problem under consideration.