THEORETICAL AND NUMERICAL ISSUES OF INCOMPRESSIBLE FLUID FLOWS

Appendix: Variational approximation of linear problems, FEM

Instructor: Pascal Frey

Sorbonne Université, CNRS
4, place Jussieu, Paris

University of Tehran, 2018
Section A.1

Variational approximation
We introduce a linear continuous functional \( \ell : V \to \mathbb{R} \), i.e. an element \( \ell \) of the topological dual space \( V' = \mathcal{L}(V, \mathbb{R}) \) of \( V \), endowed with the dual norm:

\[
\| \ell \|_{V'} = \sup_{v \in V \atop v \neq 0} \frac{\ell(v)}{\| v \|}.
\]

One of the most remarkable result in Hilbert spaces theory, states that the dual space \( V' \) of continuous linear functionals on \( V \) can be identified with \( V \) via the inner product \((\cdot, \cdot)\) in \( V \):

**Theorem 1 (Riesz’ representation theorem).** Let \( \ell \in V' \) be a continuous linear functional on \( V \). Then there is a unique element \( u \) of \( V \) for which

\[
\ell(v) = (u, v), \quad \text{for all } v \in V.
\]

In addition,

\[
\| \ell \| = \| u \|.
\]
The first application of the Riesz representation theorem is to determine the existence of the adjoint of a linear transformation.

- Assume $V$ and $W$ are Hilbert spaces and $L \in \mathcal{L}(V, W)$ is a linear operator. We use theorem 1 to define a new operator $L' : W \rightarrow V$, called the adjoint of $L$.

- Given $w \in W$, define a linear function $\ell_w \in V'$ by
  \[ \ell_w(v) = (Lv, w)_W, \quad \forall v \in V. \]

- Riesz' representation theorem states that there is a unique element (the adjoint), denoted $L'(w) \in V$ such that
  \[ \ell_w(v) = (Lv, w)_W = (v, L'(w))_V, \quad \forall v \in V, w \in W. \]

- In the particular case $V = W$ and $L = L'$, $L$ is called a self-adjoint operator. When $L$ is a self-adjoint operator from $\mathbb{R}^n$ to $\mathbb{R}^n$, it is represented by a symmetric matrix in $\mathbb{R}^{n \times n}$. Equations of the form $Lv = w$ involving a self-adjoint operator occur in many physical settings.
Linear spaces and linear operators provide a convenient setting for the analysis of such problems that are also interesting for the analysis of nonlinear operators.

- We consider the following formulation:

\[
\text{Find } u \in V, \text{ such that } a(u, v) = \ell(v), \quad \forall v \in V.
\] (1)

- The Lax-Milgram lemma provides an answer about the existence and uniqueness of the solution of such problem.

- Consider a linear operator \( A : V \rightarrow V' \) and a bilinear form \( a : V \times V \rightarrow \mathbb{R} \) such that

\[
(Au, v) = a(u, v), \quad \forall u, v \in V.
\] (2)

**Theorem 2.** Relation (2) provides a one-to-one correspondence between linear continuous operators \( A : V \rightarrow V' \) and continuous bilinear forms \( a(\cdot, \cdot) \) defined on \( V \times V \).
Many properties of the bilinear form $a(\cdot, \cdot)$ can be deduced from those of the linear operator $A$, and conversely, e.g.:

- $a$ is bounded (i.e. $a(u, v) \leq M \|u\| \|v\|$, for all $u, v \in V$) if and only if $A$ is bounded (i.e. $\|Au\| \leq M \|v\|$, for all $v \in V$).

- $a$ is positive (i.e. $(a(v, v) \geq 0$, for all $v \in V, v \neq 0$) if and only if $A$ is positive (i.e. $(Av, v) \geq 0$, for all $v \in V$).

- $a$ is strictly positive (i.e. $a(v, v) > 0$, for all $v \in V, v \neq 0$) if and only if $A$ is strictly positive (i.e. $(Av, v) > 0$, for all $v \in V, v \neq 0$).

- $a$ is strongly positive or $V$-elliptic (i.e. $a(v, v) \geq \alpha \|v\|^2$, for all $v \in V, \alpha > 0$) if and only if $A$ is strongly positive (i.e. $(Av, v) \geq \alpha \|v\|^2$, for all $v \in V$).

- $a$ is symmetric (i.e. $a(u, v) = a(v, u)$ for all $u, v \in V$) if and only if $A$ is symmetric (i.e. $(Au, v) = (Av, u)$ for all $u, v \in V$).
**Theorem 3.** Suppose $K \subset V$ is a non-empty, closed and convex subset of the Hilbert space $V$, $a(\cdot, \cdot)$ is a bilinear, symmetric, bounded and $V$-elliptic form on $V$ and $\ell \in V'$. Let

$$J(v) = \frac{1}{2}a(v, v) - \ell(v), \quad v \in V.$$ 

Then, there exists a unique minimizer $u \in K$ such that

$$J(u) = \inf_{v \in K} J(v),$$

which is also the unique solution of the variational inequality

$$u \in K, \quad a(u, v - u) \geq \ell(v - u), \quad \forall v \in K,$$

or equivalently, if $K$ is a subspace of $V$

$$u \in K, \quad a(u, v) = \ell(v), \quad \forall v \in K.$$  

**Remark 1.** When the bilinear form $a(\cdot, \cdot)$ is not symmetric, there is no longer a corresponding minimization problem.
**Definition 1.** The problem (1) is said to be well-posed if there is a unique solution to this problem and if the following linear stability property is satisfied:

\[
\text{there exists } c > 0; \text{ for all } \ell \in V', \quad \| u \|_V \leq c \| \ell \|_{V'}. \]

The next result will allows to attest the well-posedness character of a variational problem.

---

**Lemma 1 (Lax-Milgram).** Let $V$ be a Hilbert space. Suppose $a(\cdot, \cdot)$ is bounded, $V$-elliptic bilinear form on $V$, $\ell \in V'$. Then, there is a unique solution to the problem

\[
\begin{aligned}
    \ u \in V, \quad a(u, v) &= \ell(v), \quad \forall v \in V. \\
\end{aligned}
\]

Furthermore, the following estimate holds:

\[
\text{for all } \ell \in V', \quad \| u \|_V \leq \frac{1}{\alpha} \| \ell \|_{V'}. 
\]
**Corollary 1.** Under the same hypothesis of theorem 1 and assuming that the bilinear form $a(\cdot, \cdot)$ is symmetric, then the solution $u$ to problem (6) is characterized by the property

$$u \in V, \quad J(u) = \frac{1}{2}a(u, u) - \ell(u) = \inf_{v \in V} \left( \frac{1}{2}a(v, v) - \ell(v) \right).$$

**Remark 2.** (i) When the bilinear form $a(\cdot, \cdot)$ is symmetric, the problem (6) corresponds to the minimization of a quadratic functional on a Hilbert space $V$, which is the abstract formulation of numerous problems in calculus of variations. This explains why this problem is called a variational problem.

(ii) When the bilinear form $a(\cdot, \cdot)$ is symmetric and $V$-elliptic, the Lax-Milgram theorem indicates that the optimization problem $\inf_{v \in V} J(v)$ has a unique solution. The $V$-ellipticity of $a$ can be seen as a property of strong convexity of the functional $J$.

(iii) In many physical applications, the functional $J$ corresponds to an energy term (e.g. the deformation of an elastic membrane).
We consider the variational problem (1)

\[ \text{Find } u \in V, \text{ such that } a(u, v) = \ell(v), \quad \forall v \in V. \]  

and assume that all hypothesis of Lax-Milgram theorem 1 are satisfied: \( a(\cdot, \cdot) \) is a bounded, \( V \)-elliptic bilinear form on \( V \), \( \ell \in V' \) is a continuous linear form.

In order to obtain a numerical approximation of the solution \( u \), we will consider replacing this problem by a discrete problem posed in a functional space of finite dimension.

General principle:
1. introduce an approximation: sequence of subspaces of increasing finite dimension.
2. the resulting discrete problem is then more easy to solve than the initial one.
3. finally, we consider the limit case when the dimension of the subspaces tends toward infinity to construct a solution to the initial problem.
Consider a family \( \{ V_h \} \) of closed subspaces of the infinite dimensional Hilbert space \( V \), for a given parameter \( h \) related to the discretization of the domain and intended to tend towards zero in the analysis.

The notion of internal approximation is justified by the choice of subspaces \( V_h \subset V \), for all \( h \).

We assume that for every \( v \in V \), there exists an element \( r_h v \in V_h \) such that
\[
\lim_{h \to 0} \| r_h v - v \| = 0.
\]

The bilinear form \( a(\cdot, \cdot) \) and the linear form \( \ell(\cdot) \) are defined on \( V_h \times V_h \) and \( V_h \), respectively and the problem (1) can be approximated by the following discrete problem:

\[
\text{Find } u_h \in V_h, \text{ such that } a(u_h, v_h) = \ell(v_h), \quad \forall v_h \in V_h.
\]

In this case, the discrete problem is called a Galerkin approximation.
**Theorem 4.** Under the hypothesis of theorem 1, problem (9) has a unique solution $u_h$ in $V_h$ and the following estimate holds for all $\ell \in V'$:

$$\|u_h\|_V \leq \frac{1}{\alpha} \|\ell\|_{V'}.$$

**Remark 3.** Usually in applications, it is necessary to replace, using numerical integration, the bilinear form $a(\cdot, \cdot)$ and the linear form $\ell(\cdot)$ respectively by a bilinear form $a_h(\cdot, \cdot)$ and a linear form $\ell_h(\cdot)$.

Furthermore, for obvious reasons, we have the following result:

**Corollary 2.** Under the hypothesis of theorem 4 and assuming the bilinear form $a(\cdot, \cdot)$ is symmetric, then the solution $u_h$ to problem (9) is characterized by the property:

$$u_h \in V_h, \quad J(u_h) \leq J(v_h), \quad \text{for all } v_h \in V_h,$$

(10)

where the functional $J(\cdot)$ is defined as previously by (7).
Since we have replaced a variational problem by an approximation of this problem, it is interesting to question the error related to the substitution of $V$ by a subspace $V_h$.

**Proposition 1.** Under the previous hypothesis, we have the orthogonality identity:

$$\forall v_h \in V_h, \quad a(u - u_h, v_h) = 0.$$  

If $a(\cdot, \cdot)$ is a continuous, $V$-elliptic and symmetric bilinear form on $V \times V$ of ellipticity constant $\alpha$, then it defines an inner product and an energy norm associated to it as:

$$\|u\|_e = (a(u, u))^{1/2}, \quad \text{for all } v \in V$$  \hfill (11)

which is equivalent to the natural norm of $V$:

$$\alpha^{1/2} \|u\|_V \leq \|u\|_e \leq \|A\|^{1/2}\|u\|_V, \quad \text{for all } u \in V.$$

The approximate solution $u_h$ is the orthogonal projection for the inner product of $a(\cdot, \cdot)$ of the solution $u$ onto the subspace $V_h$. 


The following result helps to understand why the variational approximation method is so interesting.

**Lemma 2 (CÉA).** Let \( u \) (resp. \( u_h \)) be the solution of problem (1) (resp. (9)). We have the error estimate

\[
\| u_h - u \| \leq \frac{\| A \|}{\alpha} \inf_{v_h \in V_h} \| v_h - u \|,
\]

If the bilinear form \( a(\cdot, \cdot) \) is symmetric, then the previous estimate becomes

\[
\| u_h - u \| \leq \left( \frac{\| A \|}{\alpha} \right)^{1/2} \inf_{v_h \in V_h} \| v_h - u \|.
\]

It provides an **optimal estimate** of the error between the exact solution \( u \) of the problem (1) and the approximate solution \( u_h \) of the problem (9).
This lemma shows that the evaluation of the error is equivalent to the evaluation of the quantity $\inf_{v_h \in V_h} \|v_h - u\|$, which consists in evaluating the distance in $V$ between the solution $u$ of the problem (1) and the subspace $V_h$ of $V$.

It is useful to obtain error estimates.

**Theorem 5.** Suppose the hypothesis on $V$, $a$ and $\ell$ are those of the previous sections. Suppose that there exists a subspace $\mathcal{V} \subset V$, dense in $V$ and a linear mapping $r_h : \mathcal{V} \rightarrow V_h$ such that

$$\lim_{h \to 0} \|r_h v - v\| = 0,$$

for all $v \in \mathcal{V}$, then, the approximation method converges, i.e.

$$\lim_{h \to 0} \|u_h - u\| = 0,$$

where $u$ (resp. $u_h$) is the solution of problem (1) (resp. (9)).
Section A.2

Finite elements
When dealing with $\mathbb{P}_1$ finite elements, the approximation space $V_h$ is a space of continuous affine functions on each simplex. Globally, the functions of $V_h$ are uniquely defined by the values at the vertices of the triangulation.

The construction of a basis of $V_h$ uses the technique introduced by Lagrange (1736-1813). The basis functions are the functions $w_l$ of $V_h$ defined at the $N$ mesh nodes by the following conditions:

$$w_l(x_j, y_j) = \delta_{lj}.$$ 

Notice that these functions have a support reduced to the set of triangles having the point $(x_l, y_l)$ as a vertex. In this basis, a function of $V_h$ can be decomposed as follows:

$$v_h(x, y) = \sum_l v_l w_l(x, y).$$
$P^1$ FINITE ELEMENTS

In each simplex $K$, we denote by $\lambda_i$ the barycentric coordinates, that are the restrictions of the basis functions of $V_h$ to the triangle $K$, associated with the three vertices $a_i$ of $K$.

$\lambda_i$ is a polynomial of degree 1, equal to 1 at point $a_i$ and equal to 0 at $a_2, a_3$:

$$\lambda_1(x, y) = a_0 + a_1x + a_2y$$

Hence, $\lambda_i$ can be determined by solving the following $3 \times 3$ linear system:

$$\begin{cases}
\lambda_1(x, y) = a_0 + a_1x_1 + a_2y_1 = 1 \\
\lambda_2(x, y) = a_0 + a_1x_2 + a_2y_2 = 0 \\
\lambda_3(x, y) = a_0 + a_1x_3 + a_2y_3 = 0
\end{cases}$$
The determinant of this system
\[
\begin{vmatrix}
1 & x_1 & y_1 \\
1 & x_1 & y_1 \\
1 & x_1 & y_1 \\
1 & x_1 & y_1 \\
\end{vmatrix} = 2 \text{ area}(K)
\]
is non null only if the 3 points are not aligned.

By solving this system and similar systems for \(\lambda_2, \lambda_3\), we obtain the formulas:

\[
\lambda_1(x, y) = \frac{x_2y_3 - x_3y_2 + x(y_2 - y_3) + y(x_3 - x_2)}{2\text{area}(K)}
\]

\[
\lambda_2(x, y) = \frac{x_3y_1 - x_1y_3 + x(y_3 - y_1) + y(x_1 - x_3)}{2\text{area}(K)}
\]

\[
\lambda_3(x, y) = \frac{x_1y_2 - x_2y_1 + x(y_1 - y_2) + y(x_2 - x_1)}{2\text{area}(K)}
\]
From the previous expressions of $\lambda_i$, we deduce the expressions of the gradients:

$$\frac{\partial \lambda_1}{\partial x} = \frac{y_2 - y_3}{2\text{area}(K)}$$

$$\frac{\partial \lambda_1}{\partial y} = \frac{x_3 - x_2}{2\text{area}(K)}$$

$$\frac{\partial \lambda_2}{\partial x} = \frac{y_3 - y_1}{2\text{area}(K)}$$

$$\frac{\partial \lambda_2}{\partial y} = \frac{x_1 - x_3}{2\text{area}(K)}$$

$$\frac{\partial \lambda_3}{\partial x} = \frac{y_1 - y_2}{2\text{area}(K)}$$

$$\frac{\partial \lambda_3}{\partial y} = \frac{x_2 - x_1}{2\text{area}(K)}$$

The three functions $\lambda_i$ are also called area coordinates since they represent at a point $M(x, y)$ the algebraic ratio between the areas of triangle $Mai_ia_j$ and of triangle $K$.
With $\mathbb{P}_1$ finite elements, in each simplex, the solution is approximated using:

$$u(x, y) \approx \sum_{i=1,3} u_i \lambda_i(x, y)$$

The gradient of $u$ is thus represented in each simplex as:

$$\nabla u \approx \sum_{i=1,3} u_i \nabla \lambda_i$$

The gradient of $\mathbb{P}_1$ functions is constant in each element.
Consider a polygonal bounded domain $\Omega \subset \mathbb{R}^2$ with Lipschitz boundary $\partial \Omega$. We consider the following problem:

$$
\begin{align*}
-\Delta u &= f \quad \forall x \in \Omega \\
\frac{\partial u}{\partial n} + ku &= g \quad \forall x \in \partial \Omega
\end{align*}
$$

(14)

The formulation of this problem is written in the space $V$ of the function $s\in H^1(\Omega)$. We obtain the variational problem:

$$
\begin{align*}
\text{find } u \in H^1(\Omega) \text{ such that } \forall v \in H^1(\Omega) \\
\int_\Omega \nabla u \nabla v \, dx + \int_{\partial \Omega} kuv \, d\gamma &= \int_\Omega fv \, dx + \int_{\partial \Omega} gv \, d\gamma
\end{align*}
$$
The discrete problem is written in the space $V_h$ of the continuous functions, piecewise affine in each triangle. Let $I$ denotes the set of mesh node indices.

The solution $u_h$ is written in the basis of the $w_j$ for $j \in I$ according to:

$$u_h(x, y) = \sum_{j \in I} u_j w_j(x, y).$$

We obtain the discrete problem in $V_h$:

$$\begin{cases}
\text{find the values } u_j \text{ for } j \in I, \text{ such that } \\
\sum_{j \in I} \left( \int_{\Omega} \nabla w_j \nabla w_i \, dx + \int_{\partial\Omega} k w_j w_i \, d\gamma \right) = \int_{\Omega} f w_i \, dx + \int_{\partial\Omega} g w_i \, d\gamma
\end{cases}$$

We obtain a linear system of $N_I$ equations where $N_I$ is the number of mesh nodes:

$$Au = F$$

where $A$ is the matrix of coefficients

$$a_{i,j} = \int_{\Omega} \nabla w_j \nabla w_i \, dx + \int_{\partial\Omega} k w_j w_i \, d\gamma$$
The evaluation of the coefficients $a_{i,j}$ is realized by a procedure of assembly of the contributions of any of the mesh triangles $K$ in $\mathcal{T}_h$. For example, for the stiffness matrix:

$$K_{i,j} = \int_\Omega \nabla w_j \nabla w_i dx = \sum_k \int_{K_k} \nabla w_j \nabla w_i dx$$

In each $K = (a_1, a_2, a_3)$ corresponding to the nodes $x_i, x_j, x_k$, the sole basis functions that are not null are the functions $w_i, w_j, w_k$. Their restrictions to the simplex $K$ are the barycentric coordinates $\lambda_i$. Hence, the elementary matrix associated to triangle $K$ is a $3 \times 3$ matrix of coefficients:

$$\text{elem}(K_{i,j}) = \int_{K_k} \nabla \lambda_j \nabla \lambda_i dx \quad \forall i, j = 1, 3$$

With $\mathbb{P}^1$ elements, these integrals are easy to compute (as the functions are constant). It is sufficient to multiply the values by the area of the simplex.
The elementary mass matrix can be obtained by using the following integration formulas (that are exact):

\[
\int_K \lambda_i dx = \frac{\text{area}(K)}{3}
\]

\[
\int_K \lambda_i^2 dx = \frac{\text{area}(K)}{6}
\]

\[
\int_K \lambda_i \lambda_j dx = \frac{\text{area}(K)}{12} \quad \text{if } i \neq j.
\]

This yields

\[
\text{elem}M_{K_k} = \text{area}(K_k) \begin{pmatrix}
\frac{1}{6} & \frac{1}{12} & \frac{1}{12} \\
\ldots & \frac{1}{6} & \frac{1}{12} \\
\ldots & \ldots & \frac{1}{6}
\end{pmatrix}
\]
The gradients of $\mathbb{P}^1$ functions are constant in each element. To obtain a $\mathbb{P}^1$ gradient at the mesh vertices, we proceed using a $L^2$ projection.

Let $\bar{G}$ denote the gradient at the element and $G$ be the $\mathbb{P}^1$ gradient. We write that $G$ is the projection of $\bar{G}$ according to:

$$(G, w_i) = (\bar{G}, w_i) \quad \forall i$$

By developing $G$ in the basis $w_j : G = \sum G_j w_j$ et by computing the integrals using an approximation scheme (exact on $\mathbb{P}^1$), we obtain the formula:

$$G_j = \frac{\sum \text{area}(K_i) \bar{G}_i}{\sum \text{area}(K_i)}$$

where the sum is considered over the set of triangles sharing the vertex $j$. 

P^1 GRADIENT RECONSTRUCTION