



**HAL**  
open science

# Finite volumes and mixed Petrov-Galerkin finite elements: the unidimensional problem

François Dubois

► **To cite this version:**

François Dubois. Finite volumes and mixed Petrov-Galerkin finite elements: the unidimensional problem. Numerical Methods for Partial Differential Equations, 2000, 16 (3), p. 335-360. 10.48550/arXiv.1401.0824 . hal-00923745

**HAL Id: hal-00923745**

**<https://hal.science/hal-00923745>**

Submitted on 4 Jan 2014

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Finite volumes and mixed Petrov-Galerkin finite elements : the unidimensional problem.

François Dubois<sup>1,2</sup>

<sup>1</sup> *Department of Mathematics, University Paris Sud,  
Bât. 425, F-91405 Orsay Cedex, France.*

<sup>2</sup> *Conservatoire National des Arts et Métiers, Paris, France,  
Structural Mechanics and Coupled Systems Laboratory.*

francois.dubois@math.u-psud.fr.

04 October 1999 \*

**Abstract.** For Laplace operator in one space dimension, we propose to formulate the heuristic finite volume method with the help of mixed Petrov-Galerkin finite elements. Weighting functions for gradient discretization are parameterized by some function  $\psi : [0, 1] \rightarrow \mathbb{R}$ . We propose for this function  $\psi$  a compatibility interpolation condition and we prove that such a condition is equivalent to the inf-sup property when studying stability of the numerical scheme. In the case of stable scheme and under two distinct hypotheses concerning the regularity of the solution, we demonstrate convergence of the finite volume method in appropriate Hilbert spaces and with optimal order of accuracy.

**Résumé.** Dans le cas de l'opérateur de Laplace à une dimension d'espace, nous proposons de formuler la méthode heuristique des volumes finis à l'aide d'éléments finis mixtes dans une variante Petrov-Galerkin où les fonctions de poids pour la discrétisation du gradient sont paramétrées par une fonction  $\psi : [0, 1] \rightarrow \mathbb{R}$ . Nous proposons pour cette fonction  $\psi$  une condition de compatibilité d'interpolation qui s'avère équivalente à la condition inf-sup pour l'étude de la stabilité du schéma. Dans ce dernier cas et sous deux hypothèses distinctes concernant la régularité de la solution, nous démontrons la convergence de la méthode des volumes finis dans les espaces de Hilbert appropriés et avec un ordre optimal de précision.

**Keywords:** finite volumes, mixed finite elements, Petrov-Galerkin variational formulation, inf-sup condition, Poisson equation.

**AMS (MOS) classification:** 65N30.

---

\* Article published in *Numerical Methods for Partial Differential Equations*, volume 16, issue 3, pages 335-360, May 2000. Edition 03 January 2014.

## 1) Introduction

• We study in this paper the approximation of the homogeneous Dirichlet problem for Poisson equation on the interval  $\Omega = ]0, 1[$  :

$$(1.1) \quad -\Delta u \equiv \frac{d^2u}{dx^2} = f \quad \text{in } \Omega$$

$$(1.2) \quad u = 0 \quad \text{on the boundary } \partial\Omega \text{ of } \Omega$$

with the finite volume method. Following, *e.g.* Patankar [Pa80], this numerical method is defined as follows. Consider a “triangulation”  $\mathcal{T}$  of the domain  $\Omega$  composed with  $(n + 1)$  points :

$$(1.3) \quad \mathcal{T} = \{0 = x_0 < x_1 < x_2 < \cdots < x_{n-1} < x_n = 1\}.$$

The unknowns are the mean values  $u_{j+1/2}$  ( $j = 0, 1, \dots, n - 1$ ) in each element  $K$  of the mesh  $\mathcal{T}$ , with  $K$  of the form  $K_{j+1/2} = ]x_j, x_{j+1}[$  :

$$(1.4) \quad u_{j+1/2} \approx \frac{1}{x_{j+1} - x_j} \int_{x_j}^{x_{j+1}} u(x) dx.$$

From these  $n$  values, the method proposes an heuristic evaluation of the gradient  $p = \text{grad } u = \frac{du}{dx}$  at vertex  $x_j$  with the help of finite differences :

$$(1.5) \quad p_j = \frac{1}{h_j}(u_{j+1/2} - u_{j-1/2}), \quad j = 0, 1, \dots, n$$

$$(1.6) \quad u_{-1/2} = u_{n+1/2} = 0$$

to take into account the boundary condition (1.2) ; the length  $h_{j+1/2}$  of interval  $]x_j, x_{j+1}[$  is defined by

$$(1.7) \quad h_{j+1/2} = x_{j+1} - x_j$$

and distance  $h_j$  between the centers of two cells  $K_{j-1/2}$  and  $K_{j+1/2}$  satisfy the relations

$$(1.8) \quad \begin{cases} h_0 = \frac{1}{2} h_{1/2} \\ h_j = \frac{1}{2}(h_{j-1/2} + h_{j+1/2}), \quad j = 1, \dots, n - 1 \\ h_n = \frac{1}{2} h_{n-1/2}. \end{cases}$$

When  $p_j$  is known at vertex  $x_j$ , an integration of the “conservation law”  $\text{div}(p) + f \equiv \frac{dp}{dx} + f = 0$  over the interval  $K_{j+1/2}$  takes the following form

$$(1.9) \quad \frac{1}{h_{j+1/2}}(p_{j+1} - p_j) + \frac{1}{h_{j+1/2}} \int_{x_j}^{x_{j+1}} f(x) dx = 0, \quad j = 0, \dots, n - 1$$

and defines  $n$  equations that “closes” the problem. This method is very popular, gives the classical three point finite difference scheme

$$(1.10) \quad \frac{1}{h}(-u_{j-1/2} + 2u_{j+1/2} - u_{j+3/2}) = \frac{1}{h} \int_{x_j}^{x_{j+1}} f(x) dx, \quad j = 0, \dots, n - 1$$

for uniform meshes ( $h_{j+1/2} \equiv h$  for each  $j$ ), but the numerical analysis is difficult in the general case. First tentative was due to Gallouët [Ga92] and weak star topology in

space  $L^\infty(\Omega)$  has been necessary to take into account the possibility for meshes to “jump” abruptly from one value  $h_{j-1/2}$  to an other  $h_{j+1/2}$ .

• On the other hand, the mixed finite element method proposed by Raviart and Thomas [RT77] introduces approximate discrete finite element spaces. Let  $\mathcal{T}$  be a mesh given at relation (1.3) and  $P_1$  be the space of polynomials of total degree  $\leq 1$ . We set

$$(1.11) \quad U_{\mathcal{T}} = \{u : \Omega \mapsto \mathbb{R}, \forall K \in \mathcal{T}, u|_K \in \mathbb{R}\}$$

$$(1.12) \quad P_{\mathcal{T}} = \{p : \bar{\Omega} \mapsto \mathbb{R}, p \text{ continuous on } \Omega, \forall K \in \mathcal{T}, p|_K \in P_1\}.$$

The mixed finite element method consists in solving the problem (1.13)-(1.15) with

$$(1.13) \quad u_{\mathcal{T}} \in U_{\mathcal{T}}, p_{\mathcal{T}} \in P_{\mathcal{T}}$$

$$(1.14) \quad (p_{\mathcal{T}}, q) + (u_{\mathcal{T}}, \operatorname{div} q) = 0, \quad \forall q \in P_{\mathcal{T}}$$

$$(1.15) \quad (\operatorname{div} p_{\mathcal{T}}, v) + (f, v) = 0, \quad \forall v \in U_{\mathcal{T}}.$$

When we explicit the basis  $\chi_{j+1/2}$  ( $j = 0, 1, \dots, n-1$ ) of linear space  $U_{\mathcal{T}}$  ( $\chi_{j+1/2}$  is the numerical function equal to 1 in  $K_{j+1/2}$  and equal to 0 elsewhere) and the basis  $\varphi_j$  ( $j = 0, 1, 2, \dots, n$ ) of space  $P_{\mathcal{T}}$  (recall that  $\varphi_j$  belongs to space  $P_{\mathcal{T}}$  and satisfies the Kronecker condition  $\varphi_j(x_k) = \delta_{j,k}$  (for  $j$  and  $k = 0, 1, 2, \dots, n$ ), we introduce vectorial unknowns  $u_{\mathcal{T}}$  and  $p_{\mathcal{T}}$  according to the relations

$$(1.16) \quad u_{\mathcal{T}} = \sum_{j=0}^{n-1} u_{j+1/2} \chi_{j+1/2}$$

$$(1.17) \quad p_{\mathcal{T}} = \sum_{j=0}^n p_j \varphi_j$$

and writing again  $u_{\mathcal{T}}$  (respectively  $p_{\mathcal{T}}$ ) the vector in  $\mathbb{R}^n$  (respectively in  $\mathbb{R}^{n+1}$ ) composed by the numbers  $u_{j+1/2}$  (respectively  $p_j$ ), system (1.14)-(1.15) takes the form

$$(1.18) \quad \begin{cases} M p_{\mathcal{T}} + B^{\mathfrak{t}} u_{\mathcal{T}} = 0 \\ B p_{\mathcal{T}} = -f_{\mathcal{T}} \end{cases}$$

with

$$(1.19) \quad f_{\mathcal{T}} = \sum_{j=0}^{n-1} f_{j+1/2} \chi_{j+1/2} \equiv \sum_{j=0}^{n-1} (f, \chi_{j+1/2}) \chi_{j+1/2}.$$

The notations  $(\bullet, \bullet)$  and  $B^{\mathfrak{t}}$  define respectively the scalar product in  $L^2(\Omega)$  and the transpose of matrix  $B$ . First equation in (1.18) introduces the so-called mass matrix  $M$  and gradient matrix  $B^{\mathfrak{t}}$  according to formulae

$$(1.20) \quad \begin{cases} M_{j,k} = (\varphi_j, \varphi_k), & 0 \leq j \leq n, \quad 0 \leq k \leq n \\ B_{j,l}^{\mathfrak{t}} = (\chi_{l+1/2}, \operatorname{div} \varphi_j), & 0 \leq j \leq n, \quad 0 \leq l \leq n-1. \end{cases}$$

and second equation of (1.18) introduces the divergence matrix  $B$  which is the transpose of the gradient matrix  $B^{\mathfrak{t}}$ . The advantage of mixed formulation is that the numerical analysis is well known [RT77] : the error  $\|u - u_{\mathcal{T}}\|_0 + \|p - p_{\mathcal{T}}\|_1$  is of order 1 when the mesh size  $h_{\mathcal{T}} \equiv \sup_j h_{j+1/2}$  tends to zero when solution  $u$  of problem (1.1)-(1.2) is

sufficiently regular. The main drawback of mixed finite elements is that system (1.18) is more difficult to solve than system (1.5)-(1.9) and for this reason, the finite volume method remains very popular.

- We focus on the details of non nulls terms of tridiagonal mass matrix ; we have

$$(1.21) \quad M_{j,j} = \frac{2}{3} h_j$$

$$(1.22) \quad M_{j,j+1} = M_{j,j-1} = \frac{1}{6} h_{j+1/2}$$

and therefore

$$(1.23) \quad h_j = \sum_{k=0}^n M_{j,k}, \quad j = 0, 1, \dots, n.$$

We remark that equation (1.5) is just obtained by the “mass lumping” of the first equation of system (1.18), replacing this equation by the diagonal matrix  $h_j \delta_{j,k}$ . We refer to Baranger, Maître and Oudin [BMO96] for recent developments of this idea in one and two space dimensions.

- In the following of this article, we show that mixed finite element formulation (1.13)-(1.15) can be adapted in a Petrov-Galerkin way in order to recover both simple numerical analysis in classical Hilbert spaces. Let  $d$  be some integer  $\geq 1$  and  $\Omega$  be a bounded open set in  $\mathbb{R}^d$ . We will denote by  $L^2(\Omega)$  (or  $L^2(0, 1)$  in one space dimension when  $\Omega = ]0,1[$ ) the Hilbert space composed by squarely integrable functions and by  $\| \bullet \|_0$  the associated norm :

$$(1.24) \quad \| v \|_0 \equiv \left( \int_{\Omega} | v |^2 dx \right)^{1/2} < \infty ;$$

the scalar product is simply noted with parentheses :

$$(1.25) \quad (v, w) = \int_{\Omega} v(x) w(x) dx .$$

The Sobolev space  $H^1(\Omega)$  is composed with functions in  $L^2(\Omega)$  whose weak derivatives belong also to space  $L^2(\Omega)$ . The associated norm is denoted by  $\| \bullet \|_1$  and is defined according to

$$(1.26) \quad \| v \|_1 \equiv \left( \| v \|_0^2 + \| \text{grad } v \|_0^2 \right)^{1/2} ,$$

with  $\text{grad } v = \left( \frac{\partial v}{\partial x_1}, \dots, \frac{\partial v}{\partial x_d} \right)^t$  and  $\| \text{grad } v \|_0^2 = \sum_{j=1}^d \left\| \frac{\partial v}{\partial x_j} \right\|_0^2$ . Subspace  $H_0^1(\Omega)$  of space  $H^1(\Omega)$  is composed by functions of  $H^1(\Omega)$  whose trace values on the boundary  $\partial\Omega$  is identically equal to zero. We will denote by  $| \bullet |_1$  the so-called semi-norm associated with space  $H_0^1(\Omega)$  :  $| v |_1^2 \equiv \| \text{grad } v \|_0^2$ . The topological dual space of  $H_0^1(\Omega)$  is denoted by  $H^{-1}(\Omega)$  ; note that this space contains  $L^2(\Omega)$  but contains also distributions that can not be represented by functions.

- We will use also Sobolev space  $H^2(\Omega)$ , composed with functions  $v \in H^1(\Omega)$  whose gradient also belongs to  $H^1(\Omega)$  and the associated norm and semi-norm are defined by the relations

$$\|v\|_2^2 \equiv \|v\|_0^2 + \|\text{grad } v\|_1^2 = \|v\|_0^2 + \|\text{grad } v\|_0^2 + \sum_{1 \leq i, j \leq n} \left\| \frac{\partial^2 v}{\partial x_i \partial x_j} \right\|_0^2,$$

$$\|v\|_2 \equiv \left( \sum_{1 \leq i, j \leq n} \left\| \frac{\partial^2 v}{\partial x_i \partial x_j} \right\|_0^2 \right)^{1/2}.$$

For mathematical foundation about Sobolev spaces, we refer *i.e.* to Adams [Ad75].

- The Sobolev space  $H(\text{div}, \Omega)$  is composed by vector fields  $q = (q_1, \dots, q_d)^t \in (L^2(\Omega))^d$  whose divergence  $\text{div } q \equiv \sum_{j=1}^d \frac{\partial q_j}{\partial x_j}$  is in space  $L^2(\Omega)$ . The norm in space  $H(\text{div}, \Omega)$  is denoted by  $\|\bullet\|_{\text{div}}$  and satisfies the natural relation :

$$(1.27) \quad \|q\|_{\text{div}} \equiv \left( \sum_{j=1}^d \|q_j\|_0^2 + \|\text{div } q\|_0^2 \right)^{1/2}.$$

We will often use the product space  $V \equiv L^2(\Omega) \times H(\text{div}, \Omega)$  composed by pairs  $\eta$  of the form

$$(1.28) \quad \eta = (v, q) \in L^2(\Omega) \times H(\text{div}, \Omega)$$

and its natural associated norm satisfies

$$(1.29) \quad \|\eta\|_V \equiv \left( \|v\|_0^2 + \|q\|_{\text{div}}^2 \right)^{1/2} = \left( \|v\|_0^2 + \|q\|_0^2 + \|\text{div } q\|_0^2 \right)^{1/2}.$$

without more explicitation. In one space dimension, the spaces  $H(\text{div}, ]0, 1[)$  and  $H^1(0, 1)$  are identical and we have in this case

$$(1.30) \quad \|q\|_1 \equiv \left( \|q\|_0^2 + \|\text{div } q\|_0^2 \right)^{1/2}.$$

## 2) Continuous Petrov-Galerkin formulation

- We recall in this section the Petrov-Galerkin formulation of problem (1.1)-(1.2) in the continuous case. Let  $d$  be some integer  $\geq 1$  and  $\Omega \subset \mathbb{R}^d$  be a bounded domain with boundary  $\partial\Omega$ ,  $u$  be the solution for the Dirichlet problem for Poisson equation (2.1)

$$(2.1) \quad \begin{cases} -\sum_{j=1}^d \frac{\partial^2 u}{\partial x_j^2} \equiv -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases}$$

First equation of (2.1) can be splitted into two equations of degree 1 :

$$(2.2) \quad \begin{cases} p = \text{grad } u & \text{in } \Omega \\ \text{div } p + f = 0 & \text{in } \Omega. \end{cases}$$

We multiply the first equation of (2.2) by a test function  $q \in H(\text{div}, \Omega)$  and second equation of (2.2) by a test function  $v \in L^2(\Omega)$ . We integrate by parts the right hand side of the first equation and use the boundary condition in (2.1) to drop out the boundary term. We sum the two results and obtain

$$(2.3) \quad (u, p) \equiv \xi, \quad \xi \in V \equiv L^2(\Omega) \times H(\text{div}, \Omega)$$

$$(2.4) \quad \gamma(\xi, \eta) = \langle \sigma, \eta \rangle, \quad \forall \eta \equiv (v, q) \in V$$

with

$$(2.5) \quad \gamma((u, p), (v, q)) = (p, q) + (u, \operatorname{div} q) + (\operatorname{div} p, v)$$

$$(2.6) \quad \langle \sigma, (v, q) \rangle = -(f, v).$$

We have the following theorem, due to Babuška [Ba71].

**Theorem 1. Continuous mixed formulation.**

Let  $(V, (\bullet, \bullet))$  be a real Hilbert space,  $V'$  its topological dual space,  $\gamma : V \times V \rightarrow \mathbb{R}$  be a continuous bilinear form such that there exists some  $\beta > 0$  satisfying the so-called inf-sup condition :

$$(2.7) \quad \inf_{\|\xi\|_V=1} \sup_{\|\eta\|_V \leq 1} \gamma(\xi, \eta) \geq \beta$$

and a non uniform condition at infinity :

$$(2.8) \quad \forall \eta \in V, (\eta \neq 0 \Rightarrow \sup_{\xi \in V} \gamma(\xi, \eta) = +\infty).$$

Then, for each  $\sigma \in V'$ , the problem of finding  $\xi \in V$  satisfying the relations (2.4) has a unique solution which continuously depends on  $\sigma$  :

$$(2.9) \quad \|\xi\|_V \leq \frac{1}{\beta} \|\sigma\|_{V'}.$$

The proof of this version of Babuška result can be found *e.g.* in our report [Du97].

- We show now that choices (2.3) and (2.5) for the Poisson equation leads to a well-posed problem in the sense of Theorem 1, *i.e.* that inf-sup condition (2.7) and “infinity condition” (2.8) are both satisfied.

**Proposition 1. Continuous inf-sup and infinity conditions.**

Let  $V$  be equal to  $L^2(\Omega) \times H(\operatorname{div}, \Omega)$  and  $\gamma(\bullet, \bullet)$  be the bilinear form defined at relation (2.5). Then  $\gamma(\bullet, \bullet)$  satisfies both inf-sup condition (2.7) and infinity condition (2.8).

**Proof of proposition 1.**

- We first prove inf-sup condition (2.7). Consider  $\xi = (u, p) \in V$  with a unity norm :

$$(2.10) \quad \|\xi\|_V^2 \equiv \|u\|_0^2 + \|p\|_0^2 + \|\operatorname{div} p\|_0^2 = 1.$$

Let  $\varphi \in H_0^1(\Omega)$  be the variational solution of the problem

$$(2.11) \quad \begin{cases} \Delta \varphi = u & \text{in } \Omega, \\ \varphi = 0 & \text{on } \partial\Omega. \end{cases}$$

This function  $\varphi$  continuously depends on function  $u$ , *i.e.* there exists some constant  $C > 0$  independent of  $u$  such that

$$(2.12) \quad \|\varphi\|_1 \leq C \|u\|_0.$$

Consider some  $\beta > 0$  satisfying the inequality

$$(2.13) \quad \sqrt{1 - \beta - (1 + C^2)(\beta + \sqrt{\beta})^2} \geq \beta.$$

We verify in the following that we can construct  $\eta = (v, q) \in V$  with a norm inferior or equal to 1 such that inequality (2.7) holds. We distinguish between three cases, depending on which term among the three in (2.10) is sufficiently large.

- If we have

$$(2.14) \quad \|p\|_0^2 \geq \beta,$$

we set  $\eta \equiv (v, q)$  defined by  $v = -u$  and  $q = p$ . We have clearly, according to (2.5),  $\gamma(\xi, \eta) = \|p\|_0^2$  and inequality (2.7) is a direct consequence of (2.14) in this case.

- If inequality (2.14) is in defect and if moreover we have

$$(2.15) \quad \|u\|_0 \geq \sqrt{1+C^2}(\beta + \sqrt{\beta}),$$

we set  $v = 0$  and

$$q = \frac{1}{\sqrt{1+C^2} \|u\|_0} \text{grad } \varphi$$

with  $\varphi$  introduced in (2.11). Then it follows from relation (2.12) that the norm  $\|\eta\|_V$  of  $\eta = (v, q)$  is not greater than 1 because  $\|q\|_0 \leq \frac{C}{\sqrt{1+C^2}}$ . We have moreover

$$\begin{aligned} \gamma(\xi, \eta) &\geq (p, q) + (u, \text{div } q) + (\text{div } p, v) \\ &\geq (u, \text{div } q) - \|p\|_0 \|q\|_0 \\ &\geq \frac{\|u\|_0}{\sqrt{1+C^2}} - \sqrt{\beta} \end{aligned}$$

and due to (2.15) this last quantity is greater than  $\beta$ ; inequality (2.7) is established in this second case.

- If inequalities (2.14) and (2.15) are both in defect, we set  $v = \frac{\text{div } p}{\|\text{div } p\|_0}$  and  $q = 0$ . Then  $\eta = (v, q)$  is of unity norm and  $\gamma(\xi, \eta) = \|\text{div } p\|_0$ . But from equality (2.10) we have also

$$\begin{aligned} \|\text{div } p\|_0^2 &= 1 - \|u\|_0^2 - \|p\|_0^2 \\ &\geq 1 - (1+C^2)(\beta + \sqrt{\beta})^2 - \beta \geq \beta^2 \end{aligned}$$

due to relation (2.13). Then the inf-sup inequality (2.7) is established.

- We prove now the infinity condition (2.8). Let  $\eta = (v, q)$  be a non-zero pair of functions in the product space  $L^2(\Omega) \times H(\text{div}, \Omega)$ . We again distinguish between three cases.

(i) If  $\text{div } q \neq 0$ , we set  $u = \lambda \text{div } q$ ,  $p = 0$  and  $\xi = (u, p)$ . Then  $\gamma(\xi, \eta) = \lambda \|\text{div } p\|_0^2$  tends to  $+\infty$  as  $\lambda$  tends to  $+\infty$ .

(ii) If  $\text{div } q = 0$  and  $v \neq 0$ , let  $\varphi \in H_0^1(\Omega)$  be the variational solution of the problem

$$\begin{cases} \Delta \varphi = v & \text{in } \Omega \\ \varphi = 0 & \text{on } \partial\Omega \end{cases}$$

and  $\tilde{p} = \text{grad } \varphi$ . Then  $(\tilde{p}, q) = (\text{grad } \varphi, q) = -(\varphi, \text{div } q) = 0$ . We set  $u = 0$ ,  $p = \lambda \tilde{p}$  and  $\xi = (u, p)$ . We have  $\gamma(\xi, \eta) = \lambda(\text{div } \tilde{p}, v) = \lambda \|v\|_0^2$  which tends to  $+\infty$  as  $\lambda$  tends to  $+\infty$ .

(iii) If  $\operatorname{div} q = 0$  and  $v = 0$ , vector  $q$  is non null by hypothesis. Then  $u = 0$ ,  $p = \lambda q$  and  $\xi = (u, p)$  show that  $\gamma(\xi, \eta) = (p, q) = \lambda \|q\|_0^2$  which tends to  $+\infty$  as  $\lambda$  tends to  $+\infty$ . Inequality (2.8) is established and the proof of Proposition 1 is completed.  $\square$

### 3) Discrete mixed Petrov-Galerkin formulation for finite volumes

• We consider again the unidimensional problem (1.1)-(1.2) on domain  $\Omega = ]0, 1[$ , the mesh  $\mathcal{T}$  introduced in (1.3), a discrete approximation space  $U_{\mathcal{T}}$  of Hilbert space  $L^2(\Omega)$  defined in (1.11) and a discrete finite dimensional approximation space  $P_{\mathcal{T}}$  of Sobolev space  $H(\operatorname{div}, \Omega)$  defined at relation (1.12). We modify in the following the mixed finite element formulation (1.13)-(1.15) of problem (1.1)(1.2) and consider the discrete mixed Petrov-Galerkin formulation :

$$(3.1) \quad u_{\mathcal{T}} \in U_{\mathcal{T}}, \quad p_{\mathcal{T}} \in P_{\mathcal{T}}$$

$$(3.2) \quad (p_{\mathcal{T}}, q) + (u_{\mathcal{T}}, \operatorname{div} q) = 0, \quad \forall q \in Q_{\mathcal{T}}^{\psi}$$

$$(3.3) \quad (\operatorname{div} p_{\mathcal{T}}, v) + (f, v) = 0, \quad \forall v \in U_{\mathcal{T}}.$$

We remark that the only difference with (1.13)-(1.15) consists in the choice of test function  $q$  in relation (3.2) : in the classical mixed formulation,  $q$  belongs to space  $P_{\mathcal{T}}$  (see relation (1.14)) whereas in the present one, we suppose in equation (3.2) that  $q$  belongs to space  $Q_{\mathcal{T}}^{\psi}$ . The trial functions (space  $P_{\mathcal{T}}$ ) and the weighting functions (space  $Q_{\mathcal{T}}^{\psi}$ ) for the discretization of the equation  $p = \operatorname{grad} u$  are now not identical. Therefore we have replaced a classical mixed formulation by a Petrov-Galerkin one, in a way suggested several years ago by Hughes [Hu78] and Johnson-Nävert [JN81] for advection-diffusion problems, more recently in a similar context by Thomas and Trujillo [TT99].

We define the space  $Q_{\mathcal{T}}^{\psi}$  in the way described below.

**Definition 1. Space of weighting functions.**

Let  $\psi : [0, 1] \rightarrow \mathbb{R}$  be a continuous function satisfying the localization condition

$$(3.4) \quad \psi(0) = 0, \quad \psi(1) = 1,$$

let  $\mathcal{T}$  be a mesh given in relation (1.3) and defined by vertices  $x_j$  and finite elements  $K$  of the form  $K_{j+1/2} = ]x_j, x_{j+1}[$ . We define a basis function  $\psi_j$  of space  $Q_{\mathcal{T}}^{\psi}$  by affine transformation of function  $\psi$  :

$$(3.5) \quad \psi_j(x) = \begin{cases} \psi\left(\frac{x - x_{j-1}}{h_{j-1/2}}\right) & \text{if } x_{j-1} \leq x \leq x_j \\ \psi\left(\frac{x_{j+1} - x}{h_{j+1/2}}\right) & \text{if } x_j \leq x \leq x_{j+1} \\ 0 & \text{elsewhere.} \end{cases}$$

The space  $Q_{\mathcal{T}}^{\psi}$  is defined as the set of linear combinations of functions  $\psi_j$  :

$$(3.6) \quad q \in Q_{\mathcal{T}}^{\psi} \quad \text{iff} \quad \exists q_0, \dots, q_n \in \mathbb{R} \quad \text{such that} \quad q = \sum_{j=0}^n q_j \psi_j.$$

- The interest of such weighting functions is to be able to diagonalize the mass matrix  $(\varphi_i, \psi_j)$  ( $0 \leq i, j \leq n$ ) composed with the basis  $(\varphi_i)_{0 \leq i \leq n}$  of space  $P_{\mathcal{T}}$  and the basis  $(\psi_j)_{0 \leq j \leq n}$  of linear space  $Q_{\mathcal{T}}^{\psi}$ . We have the following result :

**Proposition 2. Orthogonality.**

Let  $\psi$  be defined as in definition 1 and satisfying moreover the orthogonality condition

$$(3.7) \quad \int_0^1 (1-x) \psi(x) dx = 0.$$

Then the mass matrix  $(\varphi_i, \psi_j)$  ( $0 \leq i, j \leq n$ ) associated with equation (3.2) is diagonal :

$$(3.8) \quad \exists H_j \in \mathbb{R}, \quad (\varphi_i, \psi_j) = H_j \delta_{i,j}, \quad 0 \leq i, j \leq n.$$

**Proof of proposition 2.**

- The proof of relation (3.8) is elementary. If  $i$  and  $j$  are two different integers, the support of function  $\varphi_i \psi_j$  is reduced to a null Lebesgue measure set except if  $i = j - 1$  or  $i = j + 1$ . In the first case, we have

$$\begin{aligned} \int_0^1 \varphi_{j-1}(x) \psi_j(x) dx &= \int_{x_{j-1}}^{x_j} \varphi_{j-1}(x) \psi_j(x) dx \\ &= h_{j-1/2} \int_0^1 (1-y) \psi(y) dy, \quad j = 1, \dots, n \end{aligned}$$

with the change of variable  $x = x_{j-1} + h_{j-1/2} y$  compatible with relations (3.5). The last expression in the previous computation is null due to (3.7).

- In a similar way, in the second case, we have :

$$\begin{aligned} \int_0^1 \varphi_{j+1}(x) \psi_j(x) dx &= \int_{x_j}^{x_{j+1}} \varphi_{j+1}(x) \psi_j(x) dx \\ &= h_{j+1/2} \int_0^1 (1-y) \psi(y) dy, \quad j = 1, \dots, n \end{aligned}$$

with a new variable  $y$  defined by the relation  $x = x_{j+1} - h_{j+1/2} y$  and thanks to relation (3.5). The resulting integral remains equal to zero due to the orthogonality condition (3.7).

- When  $j = i$ , previous calculations show that

$$\begin{aligned} \int_0^1 \varphi_j(x) \psi_j(x) dx &= \int_{x_{j-1}}^{x_j} \varphi_j(x) \psi_j(x) dx + \int_{x_j}^{x_{j+1}} \varphi_j(x) \psi_j(x) dx \\ &= (h_{j-1/2} + h_{j+1/2}) \int_0^1 y \psi(y) dy, \quad j = 1, \dots, n-1. \end{aligned}$$

If  $h_j$  is the expression defined in (1.8), the value of  $H_j$  is simply expressed by :

$$(3.9) \quad H_j = 2 h_j \int_0^1 x \psi(x) dx, \quad j = 0, \dots, n$$

and Proposition 2 is then proven. □

- We can now specify a choice of shape function  $\psi$  in order to recover finite volumes with mixed Petrov-Galerkin formulation : since relation (3.2) used with test function  $q = \psi_j$  shows (with notations given at relations (1.16) and (1.17)) :

$$(3.10) \quad H_j p_j = u_{j+1/2} - u_{j-1/2}, \quad j = 0, \dots, n,$$

the finite volumes are reconstructed if relation (3.10) is identical to the heuristic definition (1.5), *i.e.* due to (3.9), if we have the following compatibility condition between finite volumes and mixed Petrov-Galerkin formulation :

$$(3.11) \quad \int_0^1 x \psi(x) dx = \frac{1}{2}.$$

The next proposition show that cubic spline function can be chosen as localization  $\psi$  function.

**Proposition 3. Spline example.**

Let  $\psi : [0, 1] \rightarrow \mathbb{R}$  be a continuous function satisfying the localization condition (3.4), orthogonality condition (3.7) and the compatibility condition with finite volumes (3.11). Then function  $\psi$  is uniquely defined if we suppose moreover that  $\psi$  is polynomial of degree  $\leq 3$ . We have

$$(3.12) \quad \psi(x) = \frac{1}{2} + 3(2x - 1) - \frac{5}{2}(2x - 1)^3 = -9x + 30x^2 - 20x^3.$$

**Proof of proposition 3.**

- It is an elementary calculus. First, due to (3.4), it is natural to search  $\psi$  of the form  $\psi(x) = x(1 + \alpha(1 - x) + \beta(1 - x)^2)$ . Secondly it comes simply from (3.7) and (3.11) that

$$\int_0^1 \psi(x) dx = \int_0^1 x \psi(x) dx = \frac{1}{2}.$$

Then due to the explicit value of some polynomial integrals

$$\int_0^1 x(1 - x) dx = \frac{1}{6}, \quad \int_0^1 x^2(1 - x) dx = \frac{1}{12}, \quad \int_0^1 x^2(1 - x)^2 dx = \frac{1}{30},$$

we can express  $\int_0^1 \psi(x) dx$  and  $\int_0^1 x \psi(x) dx$  in terms of unknowns  $\alpha$  and  $\beta$  :

$$\int_0^1 \psi(x) dx = \frac{1}{2} + \frac{\alpha}{6} + \frac{\beta}{12}, \quad \int_0^1 x \psi(x) dx = \frac{1}{3} + \frac{\alpha}{12} + \frac{\beta}{30}.$$

We deduce that  $\alpha = 10$ ,  $\beta = -20$  and relation (3.12) holds. □

#### 4) Discrete inf-sup condition

- For unidimensional Poisson equation with homogeneous boundary condition, the finite volume method is now formulated as a discrete approximation (3.1)-(3.3) associated with the bilinear form  $\gamma(\bullet, \bullet)$  defined in relation (2.5) and the following finite dimensional subspaces  $V_1$  and  $V_2$  of continuous space  $V = L^2(\Omega) \times H(\text{div}, \Omega)$  :

$$(4.1) \quad V_1 = U_{\mathcal{T}} \times P_{\mathcal{T}}$$

$$(4.2) \quad V_2 = U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}.$$

With these notations, problem (3.1) (3.2) (3.3) can be formulated as follows :

$$(4.3) \quad \xi_1 = (u_{\mathcal{T}}, p_{\mathcal{T}}) \in V_1$$

$$(4.4) \quad \gamma(\xi_1, \eta) = \langle \sigma, \eta \rangle, \quad \forall \eta \in V_2$$

with linear form  $\sigma$  defined in (2.6). We have the following approximation theorem [Ba71].

**Theorem 2. General approximation result.**

Let  $V$  be a real Hilbert space and  $\gamma$  be a continuous bilinear form like in Theorem 1 with a continuity modulus denoted by  $M$  :

$$(4.5) \quad |\gamma(\xi, \eta)| \leq M \|\xi\|_V \|\eta\|_V, \quad \forall \xi, \eta \in V.$$

Let  $V_1$  and  $V_2$  be two closed subspaces of space  $V$  such that we have the following two properties : on one hand, there exists some constant  $\delta$  associated with the uniform discrete inf-sup condition

$$(4.6) \quad \inf_{\xi \in V_1, \|\xi\|_V=1} \sup_{\eta \in V_2, \|\eta\|_V \leq 1} \gamma(\xi, \eta) \geq \delta$$

and on the other hand, the discrete infinity condition

$$(4.7) \quad \forall \eta \in V_2 \setminus \{0\}, \quad \sup_{\xi \in V_1} \gamma(\xi, \eta) = +\infty$$

is satisfied. Then problem (4.3)(4.4) has a unique solution  $\xi_1 \in V_1$ . If  $\xi$  is the solution of continuous problem (2.3)(2.4) (obtained simply with  $V_1 = V_2 = V$ ), we have the following control of the approximation error by the interpolation error :

$$(4.8) \quad \|\xi - \xi_1\|_V \leq \left(1 + \frac{M}{\delta}\right) \|\xi - \zeta\|_V, \quad \forall \zeta \in V_1.$$

- Theorem 2 plays an analogous role than the so-called Cea lemma [Ce64] in classical analysis of the error for conforming finite elements (Ciarlet-Raviart [CR72]). It states that when constant  $\delta$  in estimate (4.6) is independent of the choice of spaces  $V_1$  and  $V_2$  (uniform inf-sup discrete condition) the error  $\|\xi - \xi_1\|_V$  is dominated by the interpolation error  $\inf_{\zeta \in V_1} \|\xi - \zeta\|_V$ , that establishes convergence with an optimal order when  $V_1$  is growing more and more towards space  $V$ . The two next propositions compare discrete  $L^2$  norms when interpolation function  $\psi$ , satisfying the two conditions (3.4) and (3.7), is moreover submitted to the following **compatibility interpolation condition**

$$(4.9) \quad \psi(\theta) + \psi(1 - \theta) \equiv 1, \quad \forall \theta \in [0, 1]$$

does not satisfy it. Note that for the spline example (3.12), compatibility interpolation condition was satisfied. We suppose also that the mesh  $\mathcal{T}$  can be chosen in the class  $\mathcal{U}_{\alpha, \beta}$  of uniformly regular meshes.

**Definition 2. Uniformly regular meshes.**

Let  $\alpha, \beta$  be two real numbers such that

$$(4.10) \quad 0 < \alpha < 1 < \beta.$$

The class  $\mathcal{U}_{\alpha, \beta}$  of uniformly regular meshes is composed by all the meshes  $\mathcal{T}$  associated with  $n_{\mathcal{T}}$  ( $n_{\mathcal{T}} \in \mathbb{N}$ ) vertices  $x_j^{\mathcal{T}}$  satisfying

$$(4.11) \quad 0 = x_0^\mathcal{T} < x_1^\mathcal{T} < \cdots < x_{n_\mathcal{T}-1}^\mathcal{T} < x_{n_\mathcal{T}}^\mathcal{T} = 1$$

and such that the corresponding measures  $h_{j+1/2}^\mathcal{T}$  of elements  $K_{j+1/2}^\mathcal{T}$

$$h_{j+1/2}^\mathcal{T} = x_{j+1}^\mathcal{T} - x_j^\mathcal{T}, \quad j = 0, 1, \dots, n_\mathcal{T} - 1$$

satisfy the condition

$$(4.12) \quad \frac{\alpha}{n_\mathcal{T}} \leq h_{j+1/2}^\mathcal{T} \leq \frac{\beta}{n_\mathcal{T}}, \quad \forall j = 0, 1, \dots, n_\mathcal{T} - 1, \quad \forall \mathcal{T} \in \mathcal{U}_{\alpha, \beta}.$$

• We remark that the ratio  $h_{j+1/2}^\mathcal{T} / h_{j-1/2}^\mathcal{T}$  of successive cells has not to be close to 1 but remains bounded from below by  $\alpha/\beta$  and from above by  $\beta/\alpha$ . We will denote by  $h_\mathcal{T}$  the maximal stepsize of mesh  $\mathcal{T}$  :

$$(4.13) \quad h_\mathcal{T} = \max_{j=0,1,\dots,n_\mathcal{T}-1} h_{j+1/2}^\mathcal{T}.$$

**Proposition 4. Stability when changing the interpolant function.**

Let  $\psi$  be a continuous function  $[0, 1] \rightarrow \mathbb{R}$  satisfying the conditions (3.4), (3.7) and the compatibility interpolation condition (4.9). Let  $\mathcal{T}$  be some mesh of the interval  $[0, 1]$  composed with  $n_\mathcal{T} = n$  elements,  $P_\mathcal{T}$  be the space of continuous  $P_1$  functions associated with mesh  $\mathcal{T}$  and defined in (1.12) and  $Q_\mathcal{T}^\psi$  be the analogous space, but associated with the use of  $\psi$  for interpolation and defined in (3.5)(3.6). Consider  $(q_0, q_1, \dots, q_n) \in \mathbb{R}^{n+1}$ ,

$$q = \sum_{j=0}^n q_j \psi_j \in Q_\mathcal{T}^\psi \quad \text{and} \quad \tilde{q} = \sum_{j=0}^n q_j \varphi_j \in P_\mathcal{T}.$$

We have the estimations

$$(4.14) \quad \frac{3}{2} \delta \|\tilde{q}\|_0^2 \leq \|q\|_0^2 \leq 12 \tilde{\delta} \|\tilde{q}\|_0^2$$

with strictly positive constants  $\delta$  and  $\tilde{\delta}$  defined by

$$(4.15) \quad \delta = \int_0^1 (\psi(\theta))^2 d\theta - \left| \int_0^1 \psi(\theta) \psi(1-\theta) d\theta \right|$$

$$(4.16) \quad \tilde{\delta} = \int_0^1 (\psi(\theta))^2 d\theta.$$

**Proof of proposition 4.**

• It is not immediate that  $\delta$  is strictly positive. From Cauchy-Schwarz inequality we have

$$(4.17) \quad \left| \int_0^1 \psi(\theta) \psi(1-\theta) d\theta \right| \leq \int_0^1 (\psi(\theta))^2 d\theta$$

which proves that  $\delta \geq 0$ . If there is exact equality in inequality (4.17), the case of equality in Cauchy-Schwarz inequality show that the two functions in the scalar product at the left hand side of (4.17) are proportional :

$$\exists \lambda, \mu \in \mathbb{R}, (\lambda, \mu) \neq (0, 0), \quad \forall \theta \in [0, 1], \quad \lambda \psi(1-\theta) + \mu \psi(\theta) = 0.$$

Taking  $\theta = 0$  in previous inequality, localization condition (3.4) shows that  $\lambda = 0$ . In a similar manner, the choice of the particular value  $\theta = 1$  implies  $\mu = 0$ , which is finally

not possible because  $(\lambda, \mu) \neq (0, 0)$ . Therefore the equality case in (4.17) is excluded and  $\delta > 0$ .

- We evaluate now the  $L^2$  norm of  $q = \sum_{j=0}^n q_j \psi_j$ . We get

$$\begin{aligned}
 \|q\|_0^2 &= \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} \left( q_j \psi\left(\frac{x_{j+1}-x}{h_{j+1/2}}\right) + q_{j+1} \psi\left(\frac{x-x_j}{h_{j+1/2}}\right) \right)^2 dx \\
 &= \sum_{j=0}^{n-1} h_{j+1/2} \int_0^1 [q_j \psi(1-\theta) + q_{j+1} \psi(\theta)]^2 d\theta, \quad x = x_j + \theta h_{j+1/2} \\
 &= \sum_{j=0}^{n-1} h_{j+1/2} \left[ (q_j^2 + q_{j+1}^2) \int_0^1 (\psi(\theta))^2 d\theta + 2 q_j q_{j+1} \int_0^1 \psi(\theta) \psi(1-\theta) d\theta \right] \\
 &\geq \sum_{j=0}^{n-1} h_{j+1/2} (q_j^2 + q_{j+1}^2) \left[ \int_0^1 (\psi(\theta))^2 d\theta - \left| \int_0^1 \psi(\theta) \psi(1-\theta) d\theta \right| \right] \\
 &\quad + \sum_{j=0}^{n-1} h_{j+1/2} (|q_j| - |q_{j+1}|)^2 \left| \int_0^1 \psi(\theta) \psi(1-\theta) d\theta \right| \\
 (4.18) \quad \|q\|_0^2 &\geq \delta \sum_{j=0}^{n-1} h_{j+1/2} (q_j^2 + q_{j+1}^2).
 \end{aligned}$$

We have an analogous inequality concerning  $\tilde{q} = \sum_{j=0}^n q_j \varphi_j$ , by replacing the number  $\delta$  by its precise value when  $\psi(\bullet)$  is replaced by an affine interpolation between data, ie function  $\mathbb{R} \ni \theta \mapsto \theta \in \mathbb{R}$ . We deduce from (4.18) in this particular case :

$$(4.19) \quad \|\tilde{q}\|_0^2 \geq \frac{1}{6} \sum_{j=0}^{n-1} h_{j+1/2} (q_j^2 + q_{j+1}^2).$$

In an analogous way, we have

$$\|q\|_0^2 \leq \sum_{j=0}^{n-1} h_{j+1/2} (|q_j| + |q_{j+1}|)^2 \int_0^1 (\psi(\theta))^2 d\theta$$

*i.e.*

$$(4.20) \quad \|q\|_0^2 \leq 2 \tilde{\delta} \sum_{j=0}^{n-1} h_{j+1/2} (|q_j|^2 + |q_{j+1}|^2).$$

We have the same inequality when the interpolant function  $q$  is replaced by  $\tilde{q}$ , and  $\delta$  replaced by its value when  $\psi(\bullet)$  is replaced by affine interpolation  $\theta \mapsto \theta$  :

$$(4.21) \quad \|\tilde{q}\|_0^2 \leq \frac{2}{3} \sum_{j=0}^{n-1} h_{j+1/2} (|q_j|^2 + |q_{j+1}|^2).$$

- From (4.20) and (4.19) we deduce

$$\|q\|_0^2 \leq 2 \tilde{\delta} \sum_{j=0}^{n-1} h_{j+1/2} (|q_j|^2 + |q_{j+1}|^2) \leq 12 \tilde{\delta} \|\tilde{q}\|_0^2$$

that establishes the second inequality of (4.14). Using estimates (4.18) and (4.21) we have

$$\|q\|_0^2 \geq \delta \sum_{j=0}^{n-1} h_{j+1/2} (|q_j|^2 + |q_{j+1}|^2) \geq \frac{3}{2} \delta \|\tilde{q}\|_0^2$$

and the proof of inequality (4.14) is completed.  $\square$

• We show now that if condition (4.7) of compatibility interpolation condition is not satisfied, the uniform inf-sup condition (4.6) cannot be satisfied for any family of uniformly regular meshes. In other words, trial functions in space  $Q_{\mathcal{T}}^{\psi}$  oscillate too much and stability is in defect.

**Theorem 3. Lack of inf-sup condition.**

Let  $\psi : [0, 1] \rightarrow \mathbb{R}$  be a continuous function satisfying conditions (3.4), (3.7) and the negation of compatibility interpolation condition, *i.e.*

$$(4.22) \quad \exists \theta \in ]0, 1[, \quad \psi(\theta) + \psi(1-\theta) \neq 1.$$

Then for any family  $\mathcal{U}_{\alpha, \beta}$  of uniformly regular meshes ( $0 < \alpha < 1 < \beta$ ), the inf-sup condition (4.6) is not satisfied for spaces  $V_1 = U_{\mathcal{T}} \times P_{\mathcal{T}}$  and  $V_2 = U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}$  and meshes  $\mathcal{T}$  of  $\mathcal{U}_{\alpha, \beta}$ :

$$(4.23) \quad \begin{cases} \forall (\alpha, \beta), 0 < \alpha < 1 < \beta, \quad \forall D > 0, \\ \exists \mathcal{T} \in \mathcal{U}_{\alpha, \beta}, \quad \exists \xi \in U_{\mathcal{T}} \times P_{\mathcal{T}} \text{ such that } \|\xi\| = 1 \text{ and} \\ \forall \eta \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}, \quad \|\eta\| \leq 1 \quad \Rightarrow \quad \gamma(\xi, \eta) \leq D. \end{cases}$$

**Proof of theorem 3.**

• The first point what we have to show is that if relation (4.22) is satisfied, then we have

$$(4.24) \quad \left| \int_0^1 \frac{d\psi}{d\theta}(\theta) \frac{d\psi}{d\theta}(1-\theta) d\theta \right| < \int_0^1 \left( \frac{d\psi}{d\theta} \right)^2 d\theta.$$

The large inequality between the two sides of (4.24) just express Cauchy-Schwarz inequality. If the equality is realized, functions  $\frac{d\psi}{d\theta}(\bullet)$  and  $\frac{d\psi}{d\theta}(1-\bullet)$  are linearly dependent:

$$(4.25) \quad \exists (\lambda, \mu) \in \mathbb{R}, (\lambda, \mu) \neq (0, 0), \quad \forall \theta \in [0, 1], \quad \lambda \frac{d\psi}{d\theta}(\theta) - \mu \frac{d\psi}{d\theta}(1-\theta) = 0.$$

Then function  $\theta \mapsto \lambda \psi(\theta) + \mu \psi(1-\theta)$  is equal to some constant whose value is equal to  $\mu$  (take  $\theta = 0$  and apply (3.4)). Moreover, taking  $\theta = 1$ , we get  $\lambda = \mu$  and we obtain in this way

$$(4.26) \quad \mu (\psi(\theta) + \psi(1-\theta) - 1) = 0, \quad \forall \theta \in [0, 1].$$

Joined with relation (4.22),  $\mu$  is necessarily equal to zero and finally  $\lambda = \mu = 0$  which express the contradiction.

• We set

$$(4.27) \quad \epsilon = \int_0^1 \left( \frac{d\psi}{d\theta} \right)^2 d\theta - \int_0^1 \frac{d\psi}{d\theta}(\theta) \frac{d\psi}{d\theta}(1-\theta) d\theta$$

and  $\epsilon > 0$  due to (4.24). We evaluate now the  $L^2$  norm of  $\operatorname{div} q = \frac{dq}{dx} = \frac{d}{dx} \left( \sum_{j=0}^{n-1} q_j \psi_j \right)$ :

$$\begin{aligned}
 \|\operatorname{div} q\|_0^2 &= \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} \left( \frac{d}{dx} \left[ q_j \psi \left( \frac{x_{j+1} - x}{h_{j+1/2}} \right) + q_{j+1} \psi \left( \frac{x - x_j}{h_{j+1/2}} \right) \right] \right)^2 dx \\
 &= \sum_{j=0}^{n-1} \frac{1}{h_{j+1/2}} \int_0^1 \left( -q_j \frac{d\psi}{d\theta}(1-\theta) + q_{j+1} \frac{d\psi}{d\theta}(\theta) \right)^2 d\theta \\
 &= \sum_{j=0}^{n-1} \frac{1}{h_{j+1/2}} \left[ (q_j^2 + q_{j+1}^2) \int_0^1 \left( \frac{d\psi}{d\theta}(\theta) \right)^2 d\theta \right] \\
 &\quad - 2 \sum_{j=0}^{n-1} \frac{1}{h_{j+1/2}} \left[ q_j q_{j+1} \int_0^1 \frac{d\psi}{d\theta}(\theta) \frac{d\psi}{d\theta}(1-\theta) d\theta \right] \\
 &\geq \epsilon \sum_{j=0}^{n-1} \frac{(q_j^2 + q_{j+1}^2)}{h_{j+1/2}} + \sum_{j=0}^{n-1} \frac{(|q_j| - |q_{j+1}|)^2}{h_{j+1/2}} \left| \int_0^1 \frac{d\psi}{d\theta}(\theta) \frac{d\psi}{d\theta}(1-\theta) d\theta \right|.
 \end{aligned}$$

Then

$$(4.28) \quad \|\operatorname{div} q\|_0^2 \geq \epsilon \sum_{j=0}^{n-1} \frac{(q_j^2 + q_{j+1}^2)}{h_{j+1/2}}.$$

• We establish now (4.23) which express the negation of uniform inf-sup condition. Consider a mesh  $\mathcal{T}$  composed with  $n$  elements uniformly distributed :

$$0 = x_0 < x_1 = \frac{1}{n} < \dots < x_k = \frac{k}{n} < \dots < x_{n-1} < x_n = 1$$

with integer  $n$  chosen such that

$$(4.29) \quad \frac{2}{\sqrt{n\epsilon}} \leq D.$$

It is clear that for each pair  $(\alpha, \beta)$  satisfying relation (4.10), mesh  $\mathcal{T}$  defined previously belongs to  $\mathcal{U}_{\alpha, \beta}$  ( $h_{j+1/2}^{\mathcal{T}}$  is exactly equal to  $\frac{1}{n_{\mathcal{T}}}$  with notations proposed at Definition 1).

Introduce  $u(x) \equiv 1$ ,  $p(x) \equiv 0$  and  $\xi \equiv (u, p) = (1, 0)$  which is clearly of norm equal to unity in space  $V = L^2(0, 1) \times H^1(0, 1)$ . For each  $\eta = (v, q)$  in subspace  $U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}$ , we have

$$\gamma(\xi, \eta) = (1, \operatorname{div} q) = q(x_n) - q(x_0).$$

From inequality (4.28) we have :

$$|q_j|^2 \leq \frac{1}{\epsilon} h_{j+1/2} \|\operatorname{div} q\|_0^2 \leq \frac{1}{n\epsilon}, \quad \forall j = 0, \dots, n$$

when  $\mathcal{T}$  is chosen as above and  $\eta$  with a norm less or equal to 1 in space  $L^2(0, 1) \times H^1(0, 1)$  (see (1.29)). Then we have

$$(4.30) \quad |\gamma(\xi, \eta)| \leq \frac{2}{\sqrt{n\epsilon}} \leq D$$

if relation (4.29) is realized. Relation (4.23) is proven and uniform inf-sup condition is in defect.  $\square$

## 5) Convergence of finite volumes in the one dimensional case

• We have proven in section 4 (Theorem 3) that if the compatibility interpolation condition

$$(5.1) \quad \psi(\theta) + \psi(1 - \theta) \equiv 1, \quad \forall \theta \in [0, 1]$$

is not realized, there is no hope to obtain convergence in usual Hilbert spaces for the finite volume method (1.5)-(1.9) formulated as a mixed Petrov-Galerkin finite element method (3.1)-(3.3) associated with a family  $\mathcal{U}_{\alpha,\beta}$  of uniformly regular meshes  $\mathcal{T}$ , shape functions  $\xi_{\mathcal{T}} = (u_{\mathcal{T}}, p_{\mathcal{T}}) \in U_{\mathcal{T}} \times P_{\mathcal{T}}$ , weighting functions  $\eta = (v, q) \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}$  and bilinear form

$$(5.2) \quad \gamma(\xi_{\mathcal{T}}, \eta) = (p_{\mathcal{T}}, q) + (u_{\mathcal{T}}, \operatorname{div} q) + (\operatorname{div} p_{\mathcal{T}}, v).$$

On the contrary, if compatibility interpolation condition (5.1) is realized, we have convergence and the following result holds.

### Theorem 4. Convergence of 1D finite volumes in Hilbert spaces.

Let  $\psi : [0, 1] \rightarrow \mathbb{R}$  be a continuous function, satisfying  $\psi(0) = 0$ , the compatibility interpolation condition (5.1) and orthogonality condition

$$(5.3) \quad \int_0^1 (1-x) \psi(x) dx = 0.$$

Let  $\mathcal{U}_{\alpha,\beta}$  ( $0 < \alpha < 1 < \beta$ ) be a family of regular meshes  $\mathcal{T}$  in the sense given in definition 2,  $U_{\mathcal{T}}$  and  $P_{\mathcal{T}}$  be interpolation spaces of piecewise constant functions in each element and continuous piecewise linear functions,  $Q_{\mathcal{T}}^{\psi}$  be the space of weighting functions proposed at Definition 1 : function  $\psi_j$  is defined in (3.5) and function  $q \in Q_{\mathcal{T}}^{\psi}$  satisfies

$$(5.4) \quad q = \sum_{j=0}^n q_j \psi_j.$$

Then for each  $f \in L^2$ , the solution  $\xi_{\mathcal{T}} = (u_{\mathcal{T}}, p_{\mathcal{T}}) \in U_{\mathcal{T}} \times P_{\mathcal{T}}$  of the finite volume method for the approximation of the solution  $\xi \equiv (u, p = \operatorname{grad} u)$  of Dirichlet problem for one-dimensional Poisson equation

$$(5.5) \quad -\Delta u = f \text{ in } ]0, 1[, \quad u(0) = u(1) = 0$$

is given by solving problem (3.1)-(3.3) :

$$(5.6) \quad \gamma(\xi_{\mathcal{T}}, \eta) = (f, v), \quad \forall \eta = (v, q) \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}$$

where bilinear form  $\gamma(\bullet, \bullet)$  is defined in (5.2).

Moreover when  $f$  belongs to space  $H^1(0, 1)$ , there exists some constant  $C > 0$  depending only on  $\alpha$  and  $\beta$  such that

$$(5.7) \quad \|u - u_{\mathcal{T}}\|_0 + \|p - p_{\mathcal{T}}\|_1 \leq C h_{\mathcal{T}} \|f\|_1, \quad \forall \mathcal{T} \in \mathcal{U}_{\alpha,\beta}$$

where  $h_{\mathcal{T}}$  is the maximal step size of mesh  $\mathcal{T}$  precisely defined in (4.13).

**Remark 1.** A simple but fundamental remark is that the finite volume method (1.5)-(1.9) corresponds exactly to the mixed Petrov-Galerkin finite element formulation, independently of the choice of interpolation function  $\psi$  satisfying (5.1). This is due to the fact that the heuristic relation (1.5) holds if the following relation

$$(5.8) \quad \int_0^1 \theta \psi(\theta) d\theta = \frac{1}{2}$$

(see also (3.9) and (3.11)) is satisfied. But relation (5.8) derives clearly from relation (5.3) by integration of identity (5.1) after multiplication by  $\theta$ . Compatibility interpolation condition (5.1) gives an acute link between consistency (relation (5.8)) and convergence (inf-sup condition (4.6)). We have proven that the heuristic relation (1.5) is the only possible finite volume scheme associated with a stable mixed Petrov Galerkin formulation.

- Some propositions are useful to be established, before proving completely Theorem 4, first established with other techniques by Baranger *et al* [BMO96] and also studied with finite difference techniques by Eymard, Gallouët and Herbin [EGH2k].

**Proposition 5.  $H^1$  continuity of  $P_1$  interpolation.**

Let  $\Pi_{\mathcal{T}}$  be the classical  $P_1$  interpolation operator in space  $P_{\mathcal{T}}$ , defined by

$$(5.9) \quad (\Pi_{\mathcal{T}}\mu)(x_j) = \mu(x_j), \quad \forall \mu \in H^1(0, 1), \quad \forall x_j \text{ vertex of mesh } \mathcal{T}.$$

When mesh  $\mathcal{T}$  describes a family  $\mathcal{U}_{\alpha,\beta}$  of uniformly regular meshes, we have the following property :

$$(5.10) \quad \exists C_1 > 0, \quad \forall \mu \in H^1(0, 1), \quad \|\Pi_{\mathcal{T}}\mu\|_1 \leq C_1 \|\mu\|_1.$$

**Proposition 6. Discrete stability.**

Let  $\alpha$  and  $\beta$  be such that  $0 < \alpha < 1 < \beta$  and  $\mathcal{U}_{\alpha,\beta}$  be a family of uniformly regular meshes. When  $\psi$  is chosen satisfying hypotheses of Theorem 4, there exists some constant  $C > 0$  such that

$$(5.11) \quad \left\{ \begin{array}{l} \forall \mathcal{T} \in \mathcal{U}_{\alpha,\beta} \quad \forall u \in U_{\mathcal{T}} \quad \exists q \in Q_{\mathcal{T}}^{\psi}, \\ (u, \operatorname{div} q) = \|u\|_0^2 \quad \text{and} \quad \|q\|_1 \leq C \|u\|_0. \end{array} \right.$$

**Proof of proposition 6.** • Let  $u$  be given in  $U_{\mathcal{T}}$  and  $\varphi \in H_0^1(0, 1)$  be the variational solution of the problem

$$(5.12) \quad \Delta \chi = u \text{ on } ]0, 1[, \quad \chi(0) = \chi(1) = 0.$$

Then (see *e.g.* [Ad75]),  $\chi$  belongs to space  $H^2$  and there exists some constant  $C_2$  independent on  $u$  such that

$$\|\chi\|_2 \leq C_2 \|u\|_0.$$

Let  $\tilde{q} = \Pi_{\mathcal{T}}(\operatorname{grad} \chi)$  be the usual  $P_1$  interpolate of  $\operatorname{grad} \chi$ . From Proposition 5, we have

$$(5.13) \quad \|\tilde{q}\|_1 \leq C_1 \|\operatorname{grad} \chi\|_1 \leq C_1 \|\chi\|_2 \leq C_1 C_2 \|u\|_0 = C_3 \|u\|_0.$$

Writing  $\tilde{q} = \sum_{j=0}^n q_j \varphi_j \in P_{\mathcal{T}}$ , we introduce the second interpolant function  $q = \sum_{j=0}^n q_j \psi_j \in Q_{\mathcal{T}}^{\psi}$  and we have, for any  $v \in U_{\mathcal{T}}$

$$\begin{aligned}
 (\operatorname{div} q, v) &= \sum_{j=0}^{n-1} v_{j+1/2} \int_{x_j}^{x_{j+1}} \operatorname{div} q \, dx \\
 &= \sum_{j=0}^{n-1} v_{j+1/2} (q_{j+1} - q_j) \\
 &= \sum_{j=0}^{n-1} v_{j+1/2} \left( \frac{d\chi}{dx}(x_{j+1}) - \frac{d\chi}{dx}(x_j) \right) \\
 &= \sum_{j=0}^{n-1} v_{j+1/2} \int_{x_j}^{x_{j+1}} \Delta \chi \, dx \\
 &= (u, v), \quad \forall v \in U_{\mathcal{T}}.
 \end{aligned}$$

In particular (choose  $v = u$ ), the equality  $(u, \operatorname{div} q) = \|u\|_0^2$  of relation (5.11) is established.

• We show now the stability inequality of relation (5.11), between  $\|q\|_1$  and  $\|u\|_0$ . We have, from relation (4.14) of Proposition 4 and estimations (5.13)

$$\begin{aligned}
 \|q\|_0^2 &\leq 12 \tilde{\delta} \|\tilde{q}\|_0^2 \leq 12 \tilde{\delta} C_3^2 \|u\|_0^2 \\
 \|q\|_1^2 &= \sum_{j=0}^{n-1} \frac{1}{h_{j+1/2}} (q_{j+1} - q_j)^2 \int_0^1 \left( \frac{d\psi}{d\theta} \right)^2 d\theta \\
 &= \int_0^1 \left( \frac{d\psi}{d\theta} \right)^2 d\theta \quad \|\tilde{q}\|_1^2 \\
 &\leq \int_0^1 \left( \frac{d\psi}{d\theta} \right)^2 d\theta \quad C_3^2 \|u\|_0^2.
 \end{aligned}$$

and since (5.1) holds,

From these inequalities, we deduce inequality  $\|q\|_1 \leq C \|u\|_0$ , with

$$C = \left( 12 \tilde{\delta} + \int_0^1 \left( \frac{d\psi}{d\theta} \right)^2 d\theta \right)^{1/2} C_3$$

and Proposition 6 is established.  $\square$

**Proposition 7. Uniform discrete inf-sup condition.**

Let  $\psi : [0, 1] \rightarrow \mathbb{R}$  be a continuous function satisfying  $\psi(0) = 0$ , orthogonality condition (5.3) and compatibility interpolation condition (5.1). Let  $\tilde{\delta}$  be defined according to relation (4.16) and

$$(5.14) \quad K = \frac{4}{3} \left( 1 + \sqrt{12 \tilde{\delta}} \right).$$

Let  $\alpha$  and  $\beta$  be real numbers such that  $0 < \alpha < 1 < \beta$ ,  $\mathcal{U}_{\alpha, \beta}$  be a family of uniformly regular meshes,  $\gamma(\bullet, \bullet)$  be the bilinear form defined in (5.2),  $C$  be the constant associated with inequality (5.11) in Proposition 6 and  $\rho > 0$  be chosen such that

$$(5.15) \quad \rho + \sqrt{K\rho} \leq \frac{1}{C} \sqrt{1 - \rho^2 - K\rho}.$$

Then we have the following uniform discrete inf-sup condition :

$$(5.16) \quad \begin{cases} \forall \mathcal{T} \in \mathcal{U}_{\alpha, \beta}, \quad \forall \xi = (u, p) \in U_{\mathcal{T}} \times P_{\mathcal{T}}, \|\xi\| = 1, \\ \exists \eta = (v, q) \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}, \quad \|\eta\| \leq 1 \quad \text{and} \quad \gamma(\xi, \eta) \geq \rho. \end{cases}$$

**Proof of proposition 7.**

- As in Proposition 1, we distinguish between three cases. If we have the condition

$$(5.17) \quad \|\operatorname{div} p\|_0 \geq \rho,$$

let  $\eta \equiv (v, q)$  be defined by  $v = \frac{\operatorname{div} p}{\|\operatorname{div} p\|_0}$  and  $q = 0$ . Then, due to relation (5.2), we have  $\gamma(\xi, \eta) = (\operatorname{div} p, v) = \|\operatorname{div} p\|_0 \geq \rho$  and inequality (5.16) is proven in this simple case.

- When (5.17) is in defect, we suppose also that  $p$  is sufficiently large :

$$(5.18) \quad \|\operatorname{div} p\|_0 \leq \rho \quad \text{and} \quad \|p\|_0^2 \geq K\rho.$$

We set  $p = \sum_{j=0}^n p_j \varphi_j$  and introduce  $q \in Q_{\mathcal{T}}^{\psi}$  according to the relation

$$(5.19) \quad q = \frac{1}{\sqrt{12\delta}} \sum_{j=0}^n p_j \psi_j.$$

From inequality (4.14) and the hypothesis done on  $\xi = (u, p)$ , we have

$$\|q\|_0 \leq \|p\|_0 \leq 1$$

and moreover :

$$\begin{aligned} (p, q) &= \frac{1}{\sqrt{12\delta}} \sum_{j=0}^n p_j^2 (\varphi_j, \psi_j) \\ &= \frac{1}{\sqrt{12\delta}} \sum_{j=0}^{n-1} h_{j+1/2} (p_j^2 + p_{j+1}^2) \int_0^1 \theta \psi(\theta) d\theta \quad \text{due to (3.9) and (1.8)} \\ &= \frac{1}{4\sqrt{3\delta}} \sum_{j=0}^{n-1} h_{j+1/2} (p_j^2 + p_{j+1}^2) \\ &\geq \frac{1}{8} \sqrt{\frac{3}{\delta}} \|p\|_0^2 \quad \text{due to (4.21)}. \end{aligned}$$

We introduce  $\eta = (0, q)$ . Then we have shown that  $\|\eta\| \leq 1$  and we have also

$$\begin{aligned} \gamma(\xi, \eta) &= (p, q) + (u, \operatorname{div} q) \\ &= (p, q) + \frac{1}{\sqrt{12\delta}} (u, \operatorname{div} p) \\ &\geq \frac{1}{8} \sqrt{\frac{3}{\delta}} K\rho - \frac{1}{\sqrt{12\delta}} \rho = \rho \end{aligned}$$

due to (5.14). Then (5.16) holds in this second case.

- In the third case, we suppose

$$(5.20) \quad \|\operatorname{div} p\|_0 \leq \rho, \quad \|p\|_0^2 \leq K\rho.$$

Then because the norm of  $\xi$  is exactly equal to 1, we have

$$\|u\|_0^2 = 1 - \|p\|_0^2 - \|\operatorname{div} p\|_0^2 \geq 1 - K\rho - \rho^2$$

which is strictly positive because the right hand side of inequality (5.15) is strictly positive ( $\rho > 0$ ). Let  $q$  be associated with  $u$  according to relation (5.11) of proposition 6 :

$$(5.21) \quad q \in Q_{\mathcal{T}}^{\psi}, \quad (u, \operatorname{div} q) = \|u\|_0^2, \quad \|q\|_1 \leq C \|u\|_0.$$

Then  $\eta \equiv (0, \frac{1}{C \|u\|_0} q)$  has a norm not greater than 1 and due to relation (5.2), we have

$$\begin{aligned} \gamma(\xi, \eta) &= \left( p, \frac{q}{C \|u\|_0} \right) + \left( u, \frac{\operatorname{div} q}{C \|u\|_0} \right) \\ &\geq -\|p\|_0 + \frac{1}{C} \|u\|_0 && \text{due to (5.21)} \\ &\geq -\sqrt{K\rho} + \frac{1}{C} \sqrt{1 - \rho^2 - K\rho} && \text{due to (5.20)} \\ &\geq \rho && \text{due to (5.15)} \end{aligned}$$

that ends the establishment of uniform inf-sup condition (5.16).  $\square$

- We need also interpolation results, that are classical (see, *e.g.* [CR72]). We detail them for completeness.

**Proposition 8. Interpolation errors.**

Let  $v \in L^2(0, 1)$  and  $q \in H^1(0, 1)$  be two given functions,  $M_{\mathcal{T}}$  and  $\Pi_{\mathcal{T}}$  the piecewise constant ( $P_0$ ) and continuous piecewise linear ( $P_1$ ) interpolation operators on mesh  $\mathcal{T}$  defined in finite dimensional spaces  $U_{\mathcal{T}}$  and  $P_{\mathcal{T}}$  respectively by the following relations

$$(5.22) \quad (M_{\mathcal{T}}v)(x) = \frac{1}{h_{j+1/2}} \int_{x_j}^{x_{j+1}} v(y) dy, \quad x_j < x < x_{j+1}$$

$$(5.23) \quad (\Pi_{\mathcal{T}}q)(x) = q(x_j) \frac{x_{j+1} - x}{h_{j+1/2}} + q(x_{j+1}) \frac{x - x_j}{h_{j+1/2}}, \quad x_j \leq x \leq x_{j+1}.$$

Then if  $v \in H^1(0, 1)$  and  $q \in H^2(0, 1)$ , we have the interpolation error estimates :

$$(5.24) \quad \|v - M_{\mathcal{T}}v\|_0 \leq C h_{\mathcal{T}} \left\| \frac{dv}{dx} \right\|_0$$

$$(5.25) \quad \|q - \Pi_{\mathcal{T}}q\|_1 \leq C h_{\mathcal{T}} \left\| \frac{d^2q}{dx^2} \right\|_0$$

where  $h_{\mathcal{T}}$ , defined in (4.13), is the maximal step size in mesh  $\mathcal{T}$  and  $C$  is some constant independant of  $\mathcal{T}$ ,  $v$  and  $q$ .

**Proof of Theorem 4.**

- First the Poisson equation (5.5) is formulated under the Petrov-Galerkin form (2.3)-(2.4) in linear space  $V = L^2(0, 1) \times H^1(0, 1)$ . Then Proposition 1 about continuous inf-sup condition and infinity condition and Theorem 1 show that the first hypothesis of Theorem 2 is satisfied.

- Secondly let  $\mathcal{U}_{\alpha, \beta}$  be a family of uniformly regular meshes  $\mathcal{T}$ . The discrete inf-sup condition is satisfied with a constant  $\delta$  in the right hand side of (4.6) which does not depend on  $\mathcal{T}$ , due to Proposition 7 and in particular inequality (5.16).

• We prove now the infinity condition (4.7) between  $V_1 = U_{\mathcal{T}} \times P_{\mathcal{T}}$  and  $V_2 = U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}$ . Let  $\eta = (v, q)$  be a non-zero pair in  $V_2$ .

★ If  $\operatorname{div} q \neq 0$ , let  $u = \lambda \operatorname{div} q$  and  $p = 0$ . We set  $\xi = (u, p) \in U_{\mathcal{T}} \times P_{\mathcal{T}}$  and we have  $\gamma(\xi, \eta) = \lambda \|\operatorname{div} q\|^2$  which tends to  $+\infty$  when  $\lambda$  tends to infinity.

★ If  $\operatorname{div} q = 0$ , and  $v \neq 0$ , we construct  $p$  as the linear interpolate of  $\operatorname{grad} \varphi$ , where  $\varphi \in H_0^1(0, 1)$  is the variational solution of Poisson problem  $\Delta \varphi = v$ . Then  $(p, q) = \left( \int_0^1 p(x) dx \right) q$  because  $\operatorname{div} q = 0$  implies that  $q$  is equal to some constant. But

$$\begin{aligned} \int_0^1 p(x) dx &= \sum_{j=1}^{n-1} \int_{x_j}^{x_{j+1}} \left( \frac{d\varphi}{dx}(x_j) \frac{x_{j+1} - x}{h_{j+1/2}} + \frac{d\varphi}{dx}(x_{j+1}) \frac{x - x_j}{h_{j+1/2}} \right) dx \\ &= \sum_{j=1}^{n-1} \int_{x_j}^{x_{j+1}} \frac{d\varphi}{dx}(x) dx \end{aligned}$$

because  $\operatorname{grad} \varphi$  is affine in each element  $]x_j, x_{j+1}[$  since  $\Delta \varphi = v$  is a constant in each such interval. We deduce that  $\int_0^1 p(x) dx = 0$  due to the homogeneous Dirichlet boundary conditions for function  $\varphi$ . We take  $\xi = (0, \lambda p)$ . Then

$$\gamma(\xi, \eta) = (\lambda p, q) + (0, \operatorname{div} q) + (\lambda \operatorname{div} p, v) = \lambda \|v\|_0^2$$

and this expression tends towards  $+\infty$  as  $\lambda$  tends to  $+\infty$ .

★ If  $\operatorname{div} q$  and  $v$  are both equal to zero,  $q$  is a constant function which is not null because  $\eta \neq 0$ . If we take  $u = 0$  and  $p = \lambda q$  (this last choice is possible because, due to (5.1),  $P_{\mathcal{T}}$  and  $Q_{\mathcal{T}}^{\psi}$  contain the constant functions), we get  $\gamma(\xi, \eta) = \lambda \|q\|_0^2$  and this expression tends to  $+\infty$  as  $\lambda$  tends to  $+\infty$ . Therefore the discrete infinity condition (4.7) is satisfied.

• The conclusion of Theorem 2 ensures the majoration of the error in  $L^2(0, 1) \times H^1(0, 1)$  norm (left hand side of relations (4.8) and (5.7)) by the interpolation error (right hand side of relation (4.8)). From Proposition 8, the interpolation error is of order one and we have

$$(5.26) \quad \|u - u_{\mathcal{T}}\|_0 + \|p - p_{\mathcal{T}}\|_1 \leq C h_{\mathcal{T}} \left( \left\| \frac{du}{dx} \right\|_0 + \left\| \frac{d^2p}{dx^2} \right\|_0 \right)$$

when  $\mathcal{T}$  belongs to family  $\mathcal{U}_{\alpha, \beta}$  of uniformly regular meshes. The final estimale (5.7) is a consequence of regularity of the solution  $u$  of the homogeneous Dirichlet Poisson problem (5.5) when  $f$  belongs to  $H^1(0, 1)$  :

$$(5.27) \quad u \in H^3(0, 1) \quad \text{and} \quad \|u\|_3 \leq \tilde{C} \|f\|_1$$

Joined with (5.26), this inequality ends the proof of Theorem 4.  $\square$

## 6) First order for least squares

• We have established with Theorem 4 that convergence of the finite volume method but the result suffers from the fact that a too important regularity is necessary for the datum of homogeneous Dirichlet problem of Poisson equation

$$(6.1) \quad -\Delta u = f \quad \text{in } ]0, 1[, \quad u(0) = u(1) = 0.$$

The dream would be to use the interpolation result

$$(6.2) \quad \| u - M_{\mathcal{T}} u \|_0 \leq C h_{\mathcal{T}} \left\| \frac{du}{dx} \right\|_0$$

but if  $u$  belongs only in  $H_0^1(0, 1)$ , its gradient  $p = \frac{du}{dx}$  belongs only in  $L^2(0, 1)$  and there is no hope to define the interpolate  $\Pi_{\mathcal{T}} p$  for a so poor regular function and consequently to define fluxes at interfaces between two finite elements or the  $L^2(0, 1)$  scalar product  $(f, v)$ .

- Secondly, the finite element method with linear finite elements show both estimates [CR72] :

$$(6.3) \quad \| u - u_{\mathcal{T}} \|_1 \leq C h_{\mathcal{T}} \left\| \frac{d^2 u}{dx^2} \right\|_0$$

$$(6.4) \quad \| u - u_{\mathcal{T}} \|_0 \leq C h_{\mathcal{T}}^2 \left\| \frac{d^2 u}{dx^2} \right\|_0.$$

Inequality (6.3) is not accessible for present finite volumes because the discrete unknown field  $u_{\mathcal{T}}$  belongs only in  $L^2(0, 1)$  and estimate (6.4) show second order accuracy in the  $L^2$  norm, which is much more precise than the interpolation estimate (6.2) can do. We will show in next theorem that the intermediate result

$$\| u - u_{\mathcal{T}} \|_0 \leq C h_{\mathcal{T}} \| f \|_0$$

holds when  $f$  belongs in  $L^2(0, 1)$ . This result is optimal in the sense that on one hand the  $H^2$  semi-norm in the right hand side of (6.3) and (6.4) demands a minimum of regularity for datum  $f$  and condition  $f \in L^2(0, 1)$  is a good regularity constraint for a distribution which *a priori* belongs to space  $H^{-1}(0, 1)$ . On the other hand, the  $L^2$  error  $\| u - u_{\mathcal{T}} \|_0$  should have the same order that the interpolation error  $\| u - M_{\mathcal{T}} u \|_0$  (see left hand side of (6.2)).

- Nevertheless, note that some kind of superconvergence between the interpolated value  $M_{\mathcal{T}} u$  and the discrete solution  $u_{\mathcal{T}}$ , i.e. estimation of the type

$$\| M_{\mathcal{T}} u - u_{\mathcal{T}} \|_0 \leq C h_{\mathcal{T}}^2$$

have been obtained by Arbogast, Wheeler and Yotov [AWY97] in the case of quasi-uniform grids and sufficiently regular solution  $u$ .

**Theorem 5. A second result of convergence.**

We make the same hypotheses than in Theorem 4 for the interpolation function  $\psi$ , for the family  $\mathcal{U}_{\alpha, \beta}$  ( $0 < \alpha < 1 < \beta$ ) of uniformly regular meshes  $\mathcal{T}$  and we suppose that datum  $f \in L^2(0, 1)$  is given. Then the solution  $u \in H^2(0, 1)$  of problem (6.1) can be approximated by the finite volume method

$$(6.5) \quad \begin{cases} \xi_{\mathcal{T}} = (u_{\mathcal{T}}, p_{\mathcal{T}}) \in U_{\mathcal{T}} \times P_{\mathcal{T}} \\ \gamma(\xi_{\mathcal{T}}, \eta) = (f, v), \quad \forall \eta \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi} \end{cases}$$

with  $U_{\mathcal{T}}, P_{\mathcal{T}}, Q_{\mathcal{T}}^{\psi}$  and  $\gamma(\bullet, \bullet)$  defined in (1.11), (1.12), (3.6) and (2.5) respectively. Moreover there exists some constant  $C$  depending only on  $\alpha$  and  $\beta$  such that

$$(6.6) \quad \|u - u_{\mathcal{T}}\|_0 + \|p - p_{\mathcal{T}}\|_0 \leq C h_{\mathcal{T}} \|f\|_0,$$

with  $h_{\mathcal{T}}$  equal to the maximal size of mesh  $\mathcal{T}$ .

**Proposition 9. Complementary interpolation estimate.**

Let  $q$  be a given function in  $H^1(0, 1)$  and  $\Pi_{\mathcal{T}}q$  be its linear interpolate in space  $P_{\mathcal{T}}$  associated with the mesh  $\mathcal{T}$  and defined in (5.23). Then we have

$$(6.7) \quad \|q - \Pi_{\mathcal{T}}q\|_0 \leq C h_{\mathcal{T}} \left\| \frac{dq}{dx} \right\|_0$$

where  $h_{\mathcal{T}}$  is the maximal step size of mesh  $\mathcal{T}$  and  $C$  some constant independent of  $\mathcal{T}$  and  $q$ .

**Proof of proposition 9.**

• The proof of this proposition is conducted as in Proposition 8. We first establish inequality (6.7) when  $\mathcal{T} = \{0 = x_0 < x_1 = 1\}$  is the trivial mesh of interval  $]0, 1[$ . In this particular case, function  $q - \Pi_{\mathcal{T}}q$  belongs to  $H_0^1(0, 1)$  and the Poincaré estimate show that we have

$$(6.8) \quad \|q - \Pi_{\mathcal{T}}q\|_0 \leq C_1 \left\| \frac{d}{dx}(q - \Pi_{\mathcal{T}}q) \right\|_0.$$

Then we can establish the simple estimation

$$(6.9) \quad \left\| \frac{d}{dx}(\Pi_{\mathcal{T}}q) \right\|_0 \leq \left\| \frac{dq}{dx} \right\|_0$$

because

$$\begin{aligned} \left\| \frac{d}{dx}(\Pi_{\mathcal{T}}q) \right\|_0^2 &= \int_0^1 (q(1) - q(0))^2 dx = \left( \int_0^1 \frac{dq}{dy} dy \right)^2 \\ &\leq \int_0^1 \left( \frac{dq}{dy} \right)^2 dy = \left\| \frac{dq}{dx} \right\|_0^2. \end{aligned}$$

The proof of estimate (6.7) in this particular case follows from triangular inequality based on (6.8) and (6.9) with  $C = 2C_1$ .

• A general mesh  $\mathcal{T} = \{0 = x_0 < x_1 < \dots < x_n = 1\}$  is composed with  $n$  trivial meshes  $\mathcal{T}_{j+1/2} = \{x_j < x_{j+1}\}$  of the interval  $]x_j, x_{j+1}[$ . We adopt the notation (5.34) introduced inside the proof of Proposition 8 and we have :

$$\begin{aligned} \|q - \Pi_{\mathcal{T}}q\|_{0, ]0, 1[}^2 &= \sum_{j=0}^{n-1} \|q - \Pi_{\mathcal{T}}q\|_{0, ]x_j, x_{j+1}[}^2 \\ &= \sum_{j=0}^{n-1} h_{j+1/2} \|\hat{q}_{j+1/2} - \hat{\Pi}\hat{q}_{j+1/2}\|_{0, ]0, 1[}^2 && \text{from (5.36)} \\ &\leq (C_1)^2 \sum_{j=0}^{n-1} h_{j+1/2} \left\| \frac{d}{d\theta} \left( \hat{q}_{j+1/2} - \hat{\Pi}\hat{q}_{j+1/2} \right) \right\|_{0, ]0, 1[}^2 && \text{from (6.8)} \end{aligned}$$

$$\|q - \Pi_{\mathcal{T}} q\|_{0,]0,1[}^2 \leq (C_1)^2 \sum_{j=0}^{n-1} h_{j+1/2}^2 \left\| \frac{d}{dx} (q - \Pi_{\mathcal{T}} q) \right\|_{0,]x_j, x_{j+1}[}^2 \quad \text{from (5.37).}$$

Then

$$(6.10) \quad \|q - \Pi_{\mathcal{T}} q\|_0 \leq C_1 h_{\mathcal{T}} \left\| \frac{d}{dx} (q - \Pi_{\mathcal{T}} q) \right\|_0.$$

In an analogous way than the one that conducted to estimation (6.9), we have :

$$\begin{aligned} \left\| \frac{d}{dx} (\Pi_{\mathcal{T}} q) \right\|_0^2 &= \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} \left( \frac{q(x_{j+1}) - q(x_j)}{h_{j+1/2}} \right)^2 dx \\ &= \sum_{j=0}^{n-1} \frac{1}{h_{j+1/2}} \left( \int_{x_j}^{x_{j+1}} \left( \frac{dq}{dx} \right) dx \right)^2 \\ &\leq \sum_{j=0}^{n-1} \frac{1}{h_{j+1/2}} \left( \int_{x_j}^{x_{j+1}} \left( \frac{dq}{dx} \right)^2 dx \right) \left( \int_{x_j}^{x_{j+1}} dx \right) \\ &= \left\| \frac{dq}{dx} \right\|_0^2 \end{aligned} \quad \text{by Cauchy-Schwarz}$$

$$(6.11) \quad \left\| \frac{d}{dx} \Pi_{\mathcal{T}} q \right\|_0 \leq \left\| \frac{dq}{dx} \right\|_0.$$

Then inequality (6.10) joined with (6.11) and the triangular inequality show (6.7) with  $C = 2C_1$ .  $\square$

### Proof of Theorem 5.

• We divide it into three steps. First we establish that if a pair  $(s_{\mathcal{T}}, m_{\mathcal{T}}) \in U_{\mathcal{T}} \times P_{\mathcal{T}}$  is solution of the discrete finite volume problem in Petrov-Galerkin formulation, with data  $\delta$  and  $\varphi$  in  $L^2(0, 1)$

$$(6.12) \quad (m_{\mathcal{T}}, q) + (s_{\mathcal{T}}, \operatorname{div} q) = (\delta, q) + (\varphi, \operatorname{div} q), \quad \forall q \in Q_{\mathcal{T}}^{\psi}$$

$$(6.13) \quad (\operatorname{div} m_{\mathcal{T}}, v) = 0, \quad \forall v \in U_{\mathcal{T}}$$

then we have a stability estimate

$$(6.14) \quad \|s_{\mathcal{T}}\|_0 + \|m_{\mathcal{T}}\|_1 \leq C \left( \|\delta\|_0 + \|\varphi\|_0 \right)$$

where  $C$  is a constant dependent only on parameters  $\alpha, \beta$  of the class  $\mathcal{U}_{\alpha, \beta}$  of uniform meshes. Since  $\psi$  interpolant function satisfies the interpolation compatibility condition, Proposition 7 establishes that the discrete inf-sup condition is uniformly satisfied :

$$(6.15) \quad \begin{cases} \exists \rho > 0, \quad \forall \mathcal{T} \in \mathcal{U}_{\alpha, \beta}, \quad \forall \xi = (u, p) \in U_{\mathcal{T}} \times P_{\mathcal{T}}, \xi \neq 0, \\ \exists \eta = (v, q) \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}, \quad \|\eta\| \leq 1 \quad \text{and} \quad \gamma(\xi, \eta) \geq \rho \|\xi\|. \end{cases}$$

We use this stability inequality with  $\xi = (s_{\mathcal{T}}, m_{\mathcal{T}})$  solution of problem (6.12)-(6.13). Then there exists  $\eta = (v, q) \in U_{\mathcal{T}} \times Q_{\mathcal{T}}^{\psi}$  such that  $\|\eta\| \leq 1$  and

$$\begin{aligned} \frac{1}{\sqrt{2}} (\|s_{\mathcal{T}}\|_0 + \|m_{\mathcal{T}}\|_1) &\leq \|\xi\| \leq \frac{1}{\rho} \gamma(\xi, \eta) \\ &= \frac{1}{\rho} \left( (\delta, q) + (\varphi, \operatorname{div} q) \right) \\ &\leq \frac{1}{\rho} (\|\delta\|_0 \|q\|_0 + \|\varphi\|_0 \|\operatorname{div} q\|_0) \end{aligned}$$

$$\frac{1}{\sqrt{2}} (\|s_{\mathcal{T}}\|_0 + \|m_{\mathcal{T}}\|_1) \leq \frac{1}{\rho} (\|\delta\|_0 + \|\varphi\|_0) \|\eta\|$$

and inequality (6.14) is a direct consequence of the fact that  $\|\eta\| \leq 1$ .

• Secondly let  $w_{\mathcal{T}}$  and  $\mu_{\mathcal{T}}$  be two arbitrary functions in spaces  $U_{\mathcal{T}}$  and  $P_{\mathcal{T}}$  respectively. From the continuous mixed formulation

$$(6.16) \quad (p, q) + (u, \operatorname{div} q) = 0 \quad \forall q \in H^1(0, 1)$$

$$(6.17) \quad (\operatorname{div} p, v) + (f, v) = 0 \quad \forall v \in L^2(0, 1)$$

and the discrete Petrov-Galerkin approximation

$$(6.18) \quad (p_{\mathcal{T}}, q) + (u_{\mathcal{T}}, \operatorname{div} q) = 0 \quad \forall q \in Q_{\mathcal{T}}^{\psi}$$

$$(6.19) \quad (\operatorname{div} p_{\mathcal{T}}, v) + (f, v) = 0 \quad \forall v \in U_{\mathcal{T}}.$$

We deduce by difference

$$(6.20) \quad (p_{\mathcal{T}} - \mu_{\mathcal{T}}, q) + (u_{\mathcal{T}} - w_{\mathcal{T}}, \operatorname{div} q) = (p - \mu_{\mathcal{T}}, q) + (u - w_{\mathcal{T}}, \operatorname{div} q) \quad \forall q \in Q_{\mathcal{T}}^{\psi}$$

$$(6.21) \quad (\operatorname{div} (p_{\mathcal{T}} - \mu_{\mathcal{T}}), v) = (\operatorname{div} (p - \mu_{\mathcal{T}}), v) \quad \forall v \in U_{\mathcal{T}}.$$

If we select for  $\mu_{\mathcal{T}}$  the  $P_1$  interpolate of  $p$  in space  $P_{\mathcal{T}}$ , *i.e.*  $\mu_{\mathcal{T}} = \Pi_{\mathcal{T}} p$ , we have  $p(x_j) = \Pi_{\mathcal{T}} p(x_j)$  for each vertex  $x_j$  of mesh  $\mathcal{T}$ , then

$$\int_{x_j}^{x_{j+1}} \operatorname{div} (p - \Pi_{\mathcal{T}} p) \, dx = 0$$

and the same property is true for the right hand side of (6.21). Considering now the particular case of  $w_{\mathcal{T}} = M_{\mathcal{T}} u$ , we deduce from (6.20)(6.21) and previous estimate (6.14) the inequality

$$(6.22) \quad \|p_{\mathcal{T}} - \Pi_{\mathcal{T}} p\|_1 + \|u_{\mathcal{T}} - M_{\mathcal{T}} u\|_0 \leq C \left( \|p - \Pi_{\mathcal{T}} p\|_0 + \|u - M_{\mathcal{T}} u\|_0 \right).$$

Joined with the triangular inequality and majoration of  $L^2$  norm by the  $H^1$  norm, we obtain

$$(6.23) \quad \|p - p_{\mathcal{T}}\|_0 + \|u - u_{\mathcal{T}}\|_0 \leq (1 + C) \left( \|p - \Pi_{\mathcal{T}} p\|_0 + \|u - M_{\mathcal{T}} u\|_0 \right).$$

• The end of the proof is a direct consequence of Propositions 8 and 9 and in particular estimations (5.24) and (6.7) :

$$(6.24) \quad \|p - p_{\mathcal{T}}\|_0 + \|u - u_{\mathcal{T}}\|_0 \leq C h_{\mathcal{T}} \left( \left\| \frac{dp}{dx} \right\|_0 + \|p\|_0 \right)$$

joined with the classical estimate that comes from the variational formulation of problem (6.1) :

$$(6.25) \quad \|p\|_1 \leq C \|f\|_0.$$

The sequence of inequalities (6.24) and (6.25) establishes completely the inequality (6.6) modulo classical conventions in numerical analysis concerning the so-called *constant C*.  $\square$

## 7) Conclusion and acknowledgments

- In two space dimensions, mass lumping of mass matrix of mixed finite elements has defined a particular finite volume method analysed by Baranger et al [BMO96]. Note also that first results for the Laplace equation approached by a simple finite volume method on Delaunay-Voronoi triangular meshes have been obtained by Herbin [He95]. We think also that our one-dimensional result for finite volumes via mixed Petrov-Galerkin finite elements can be generalized in dimension 2 and 3 for regular triangular or tetrahedral meshes with a numerical scheme like the diamond scheme suggested by Noh many years ago [No64], first analysed by Coudière, Vila and Villedieu [CVV99], or our “wedding scheme” proposed in an other context [Du92].
- We have been introduced to techniques of Petrov-Galerkin formulations thanks to a pedagogical initiative of Bernard Larrourou at Ecole Polytechnique. The breakthrough of this research was done during a spring school at Les Houches in may 1996 ; we thank all the participants and in particular Olga Cueto for good working sollicitation. This article has been typed with  $\text{\TeX}$  by the author and we are redevable to the competences of Jean Louis Loday. Second version of this report is due to particular encouragements of Jean-Pierre Croisille. The author thanks also the referee for helpfull suggestions.

## 8) References

- [Ad75] R.A. Adams. *Sobolev spaces*. Academic Press, New York, 1975.
- [AWY97] T. Arbogast, M.F. Wheeler, I. Yotov. Mixed finite elements for elliptic problems with tensor coefficients as cell-centered finite differences, *SIAM J. Numer. Anal.*, vol 34, p. 828-852, 1997.
- [Ba71] I. Babuška. Error-Bounds for Finite Element Method, *Numer. Math.*, vol 16, p. 322-333, 1971.
- [BMO96] J. Baranger, J. F. Maître, F. Oudin. Connection between finite volumes and mixed finite element methods, *Math. Mod. and Numer. Anal.*, vol 30, p. 445-465, 1996.
- [Ce64] J. Cea. Approximation variationnelle des problèmes aux limites, *Ann. Inst. Fourier* (Grenoble), vol 14, p. 345-444, 1964.
- [CR72] P.G. Ciarlet, P.A. Raviart. General Lagrange and Hermite Interpolation in  $\mathbb{R}^n$  with Applications to Finite Element Methods, *Arch. Rational Mech. Anal.*, vol 46, p. 177-199, 1972.
- [CVV99] Y. Coudière, J.P. Vila, P. Villedieu. Convergence rate of a finite volume scheme for a two-dimensional convection-diffusion problem, *Modélisation Mathématique et Analyse Numérique*, vol 33, p. 494-516, 1999.

- [Du92] F. Dubois. Interpolation de Lagrange et volumes finis. Une technique nouvelle pour calculer le gradient d'une fonction sur les faces d'un maillage non structuré, *Aérospatiale Espace & Defense, Internal report ST/S 104109*, february 1992. See also *Lemmes finis pour la dynamique des gaz*, chapter 8, hal-00733937.
- [Du97] F. Dubois. Finite volumes and Petrov-Galerkin finite elements. The unidimensional problem, *Institut Aérotechnique de Saint Cyr*, Report 295, Conservatoire National des Arts et Métiers, october 1997.
- [EGH2k] R. Eymard, T. Gallouët, R. Herbin. Finite Volume Methods, *Handbook of Numerical Analysis* (Ciarlet-Lions Eds), North Holland, Amsterdam, vol 7, p. 715-1022, 2000.
- [Ga92] T. Gallouët. An introduction to Finite Volume Methods, in *Méthodes de Volumes Finis*, cours CEA-EDF-INRIA, Clamart, p. 1-85, 1992.
- [He95] R. Herbin. An error estimate for a finite volume scheme for a diffusion-convection problem in a triangular mesh, *Numer. Meth. for Part. Diff. Equations*, vol 11, p. 165-173, 1995.
- [Hu78] T.J.R. Hughes. A Simple Scheme for Developing "Upwind" Finite Elements, *Int. J. of Numer. Meth. in Eng.*, vol 12, p. 1359-1365, 1978.
- [JN81] C. Johnson, U. Nävert. An analysis of Some Finite Element Methods for Advection-Diffusion Problems, in *Analytical and Numerical Approaches to Asymptotic Problems in Analysis* (Axelsson, Frank, Van des Sluis Eds), North Holland, Amsterdam, p. 99-116, 1981.
- [No64] W.F. Noh. CEL : A Time Dependent Two Space Dimensional, Coupled Euler Lagrange Code, in *Methods in Computational Physics*, vol 3, Academic Press, p. 117-179, 1964.
- [Pa80] S.V. Patankar. *Numerical Heat Transfer and Fluid Flow*, Hemisphere publishing, 1980.
- [RT77] P.A. Raviart, J.M. Thomas. A Mixed Finite Element Method for 2nd Order Elliptic Problems, in *Lectures in Mathematics*, vol 606 (Dold-Eckmann Eds), Springer-Verlag, Berlin, p. 292-315, 1977.
- [TT99] J.M. Thomas, D. Trujillo. Analysis of Finite Volumes Methods, *Université de Pau et des Pays de l'Adour, Applied Mathematics Laboratory*, Internal Report 95-19, 1995, Mixed Finite Volume Methods, *International J. for Numerical Methods in Engineering*, vol 46, p. 1351-1366, 1999.