# Table des matières

**1 Introduction** **2**

**I Mise en oeuvre des méthodes de Galerkin discontinues pour une équation d'advection-diffusion** **6**

**2 Méthode de Galerkin continue** **7**

**3 Méthodes de Galerkin discontinues** **11**
3.1 Méthode de Nitsche . . . . . 11
3.1.1 Application à l'équation hyperbolique . . . . . 16
3.1.2 Application à l'équation d'advection-diffusion . . . . . 22
3.2 Méthode de Baumann-Oden . . . . . 29

**4 Problème d'évolution** **32**

**II Estimations d'erreur** **37**

**5 Estimations d'erreur a priori** **38**

**6 Estimations d'erreur a posteriori** **43**

**7 Méthodes adjointes et dualité** **48**
7.1 Préambule . . . . . 48
7.2 Mise en oeuvre et analyse des résultats . . . . . 53

**8 Conclusions et perspectives** **58**

# Chapitre 1

# Introduction

La formulation de l'équation d'advection-diffusion qu'on se propose d'étudier est la suivante :

$$\begin{aligned} \nabla \cdot (\beta u) - \epsilon \Delta u &= S \qquad \text{sur } \Omega \subset \mathbf{R}^n \ , \\ u &= f \qquad \text{sur } \Gamma_D \ , \\ (\epsilon \nabla u) \cdot n &= g \qquad \text{sur } \Gamma_N \ , \end{aligned} \tag{1.1}$$

où $u$ (un champ scalaire) représente une concentration (par exemple, un polluant dans un cours d'eau), $\beta \in \mathbf{R}^n$ est un champ de vitesses donné, $\epsilon \geq 0$ un coefficient de diffusion et S est un terme source.

On a également :

$$\Gamma_D \cup \Gamma_N = \partial\Omega, \qquad \text{où } \partial\Omega \text{ est la frontière de } \Omega, \tag{1.2}$$

$$\Gamma_D \cap \Gamma_N = \phi, \tag{1.3}$$

$$\Gamma_D \neq \phi. \tag{1.4}$$

Nous pouvons définir les parties entrantes et sortantes de la frontière :

$$\Gamma_D \supseteq \Gamma_- = \{x \in \partial\Omega : \beta \cdot n(x) < 0\}, \qquad \Gamma_+ = \partial\Omega \setminus \Gamma_-. \tag{1.5}$$

Nous supposerons par la suite que $\beta$ et $\epsilon$ sont constants en $x \in \Omega$.

L'étude d'une méthode numérique visant à résoudre cette équation linéaire peut servir de base à l'établissement de méthodes numériques pouvant résoudre des problèmes non-linéaires tels que, par exemple, les équations de Navier-Stokes.

### Méthode des caractéristiques

Mais intéressons-nous quelques instants au cas $\epsilon = 0$. Si $\epsilon$ est nul, nous avons alors affaire à l'équation purement hyperbolique $\beta \cdot \nabla u = S$. Comme conditions initiales, on demandera donc que

$$u = f \qquad \text{sur } \Gamma_- \ . \tag{1.6}$$

Les lignes de courant $x(s)$ de tangente alignée sur $\beta$ sont données par :

$$\frac{dx_i}{ds} = \beta_i(x) \qquad i = 1, \cdots, d. \tag{1.7}$$

$$x(0) = \tilde{x} \tag{1.8}$$

où $\tilde{x} \in \Gamma_-$.

Or,

$$\frac{d}{ds}(u(x(s))) = \sum_{i=1}^{d} \frac{\partial u}{\partial x_i}\frac{dx_i}{ds} = \sum_{i=1}^{d} \frac{\partial u}{\partial x_i}\beta_i(x) = \beta \cdot \nabla u, \tag{1.9}$$

il faut donc résoudre

$$\frac{d}{ds}(u(x(s))) = S \tag{1.10}$$

avec $u(\tilde{x}) = f$ si $\tilde{x} \in \Gamma_-$.

On peut donc déterminer $u$ par intégration le long de la caractéristique passant par $\tilde{x}$. Il est important de remarquer, à ce stade, que la solution u peut présenter des discontinuités d'une caractéristique à l'autre.

Considérons le problème suivant dans $\mathbf{R}^2$ :

### Problème 1

$$\frac{\partial u}{\partial x} + \frac{\partial u}{\partial y} = 0 \qquad \textit{avec } max(x, y) \leq 1 \textit{ et } min(x, y) \geq 0, \tag{1.11}$$

$$u(0, y) = 1 \textit{ et } u(x, 0) = 0 \qquad \textit{où } 0 < y \leq 1 \textit{ et } 0 \leq x \leq 1. \tag{1.12}$$

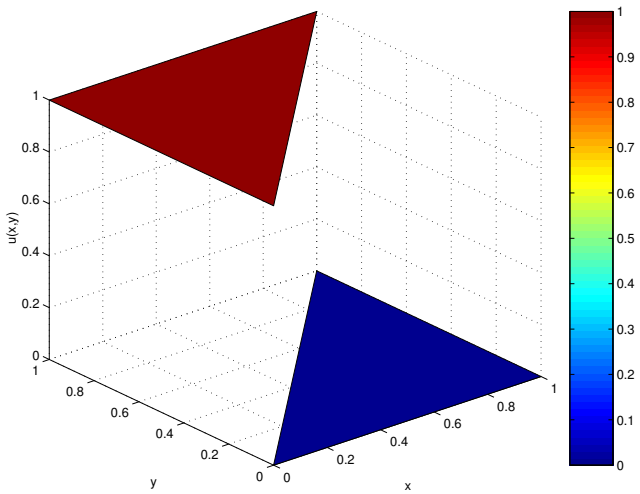



FIG. 1.1 – Discontinuité en y = x

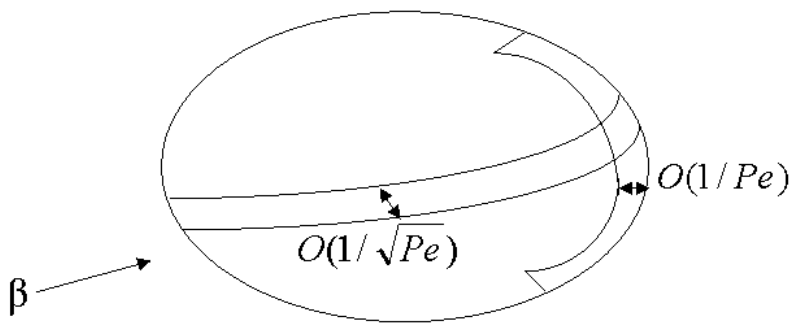
Cela correspond à $\beta_1 = \beta_2 = 1$ et $S = 0$. La solution de ce problème est donnée par $u(x, y) = 1$ si $y > x$ et $u(x, y) = 0$ sinon.

### Couches limites

Revenons à notre problème elliptique initial donné par (1.1) et interrogeons-nous à propos de la régularité de la solution. Comme nous l'avons remarqué plus haut, la solution du problème hyperbolique ($\epsilon = 0$) peut produire des discontinuités. Si $\epsilon$ est strictement positif, une telle discontinuité va *diffuser* tout au long de la caractéristique et ce, sur une région de largeur $\mathrm{O}(\sqrt{\frac{1}{Pe}})$ où Pe est le nombre de Péclet du problème :

$$Pe = \frac{||\beta||L}{\epsilon}. \tag{1.13}$$

De même, si la solution du problème hyperbolique en $\Gamma_+$ ne coincide pas avec les conditions en $\Gamma_+$ du problème elliptique, une *couche limite de sortie* apparaîtra et cette dernière aura une largeur $\mathrm{O}(\frac{1}{Pe})$.




Dans la première partie, nous nous attacherons à formuler un problème faible permettant de résoudre l'équation d'advection-diffusion, nous verrons que la méthode continue ne parvient pas à capturer les couches limites si le maillage est trop grand. Nous formulerons alors les méthodes discontinues qui ont recours à un *crime variationnel*. Nous étudierons les performances de ces méthodes pour l'équation d'advection-diffusion et l'équation hyperbolique. En particulier, nous analyserons l'implication du *paramètre de stabilisation* qui intervient dans ces méthodes. Enfin, nous parlerons du cas transitoire, nous verrons comment combiner les méthodes discontinues avec des méthodes comme, par exemple, la méthode d'*Euler explicite*.

Dans la seconde partie, nous étudierons les propriétés de convergence de la méthode, nous analyserons plus particulièrement la norme *énergétique* de l'erreur, nous verrons qu'elle a une convergence d'ordre 0.5 (ce qui rejoint la théorie) ; nous aborderons ensuite les méthodes résiduelles qui permettent de raffiner le maillage afin de réduire cette norme énergétique de l'erreur. Enfin, nous étudierons les méthodes duales qui permettent de calculer une fonctionnelle linéaire à une tolérance près.

# Première partie

# Mise en oeuvre des méthodes de Galerkin discontinues pour une équation d'advection-diffusion

# Chapitre 2

# Méthode de Galerkin continue

*Nous allons tenter de trouver une formulation faible pertinente au problème d'advection-diffusion tel que présenté en (1.1). Pour cela, nous commencerons par rappeler la méthode de Galerkin*[1] *continue classique ; nous allons nous apercevoir que cette méthode approxime mal les couches limites si le maillage est trop grand.*

Supposons tout d'abord qu'on impose des conditions de Dirichlet nulles sur toute la frontière, c'est-à-dire

$$\Gamma_D = \partial\Omega \text{ et } u|_{\Gamma_D} = 0. \tag{2.1}$$

Nous allons chercher une formulation faible du type

$$a_{AD,1}(u,v) = b_{AD,1}(v) \tag{2.2}$$

où $u, v \in U_0(\Omega)$, $U_0(\Omega)$ étant un espace de fonctions continues (et nulles à la frontière) approprié (en fait, $U_0(\Omega)$ n'est rien d'autre que $H_0^1(\Omega)$). On a, par le théorème de Stokes,

$$\begin{aligned} -\int_\Omega v(\epsilon\Delta u - \beta\cdot\nabla u)dx &= \int_\Omega \epsilon\nabla u\cdot\nabla v + (\beta\cdot\nabla u)v dx \\ &= \int_\Omega S v dx. \end{aligned} \tag{2.3}$$

puisque le second terme résultant de l'intégration par partie,

$$\int_{\partial\Omega} v(\nabla u\cdot n(x))ds$$

[1] Boris Grigorievich Galerkin, 4 mars 1871, Polotsk, Biélorussie; 12 juin 1945, Moscou, URSS

est nul.

Par conséquent, le problème faible sera :
trouver $u \in U_0(\Omega)$ tel que

$$\int_\Omega \epsilon \nabla u \cdot \nabla v + (\beta \cdot \nabla u) v dx = \int_\Omega S v dx \qquad \forall v \in U_0(\Omega), \tag{2.4}$$

c'est-à-dire :

$$a_{AD,1}(u, v) = b_{AD,1}(v) \qquad \forall v \in U_0(\Omega). \tag{2.5}$$

Donnons-nous à présent un espace de dimension finie dans lequel on va chercher une approximation de la solution du problème faible.

**Définition 1** *Soit $\mathcal{T}$ une subdivision de $\Omega$ en éléments ouverts $\kappa$ telle que $\overline{\Omega} = \bigcup_{\kappa \in \mathcal{T}} \overline{\kappa}$ .*

**Définition 2**

$$\tilde{V}_0^h(\Omega, \mathcal{T}) = \{v^h \in \mathcal{C}(\Omega) : v^h|_\kappa \in P_i \text{ et } v^h = 0 \text{ sur } \Gamma_D\}$$

*où $P_i$ est l'espace de dimension finie des polynômes de degré égal ou inférieur à $i$ et définis sur $\kappa$.*

Notons qu'on a que

$$\tilde{V}_0^h(\Omega, \mathcal{T}) \subset U_0(\Omega).$$

La méthode numérique revient donc à

trouver $u_h \in \tilde{V}_0^h(\Omega, \mathcal{T})$ tel que :

$$a_{AD,1}(u_h, v_h) = b_{AD,1}(v_h) \qquad \forall v_h \in \tilde{V}_0^h(\Omega, \mathcal{T}).$$

**Mise en oeuvre**

Nous avons testé cette méthode sur le problème unidimensionnel que voici :

$$\frac{\mathrm{d}u}{\mathrm{d}x} - 0.05 \frac{\mathrm{d}^2 u}{\mathrm{d}x^2} = 0 \qquad u(0) = 1, u(1) = 0.$$

La solution analytique de ce problème est :

$$u(x,y) = 1 - \frac{\exp(20x) - 1}{\exp(20) - 1}$$

Les conditions aux limites ne sont pas toutes les deux nulles, mais cela n'est pas grave, il suffit de considérer des espaces satisfaisant les conditions de Dirichlet.

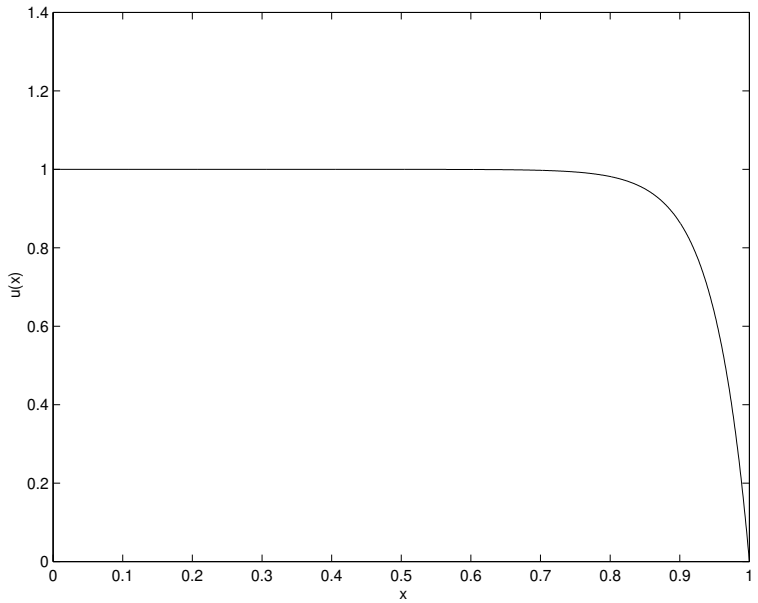



FIG. 2.1 – Solution analytique du problème unidimensionnel

Pour la résolution numérique, nous nous sommes donné deux pas de maillage différents, l'un de longueur 0.2, l'autre de longueur 0.1. On remarque (Figure 2.2) que les solutions ont des comportements radicalement différents : là où le maillage est trop grand, la solution présente des oscillations tandis que lorsqu'on raffine le maillage, la solution numérique a une allure proche de la solution analytique.

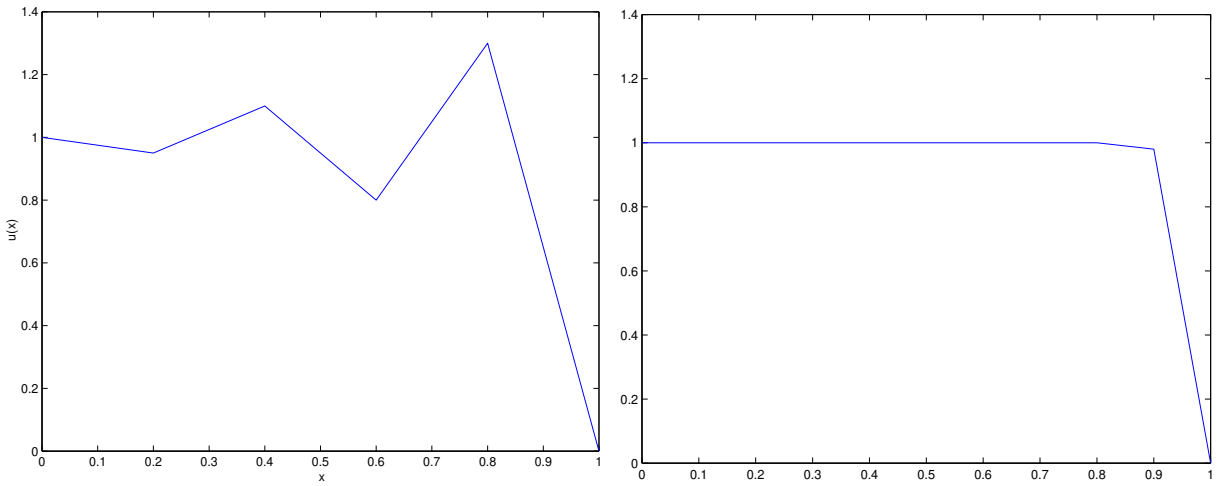



FIG. 2.2 – Solutions numeriques du problème unidimensionnel avec un maillage de 0.2 et de 0.1

# Chapitre 3

# Méthodes de Galerkin discontinues

*Dans ce chapitre, nous étudierons une première méthode discontinue où les conditions aux frontières seront fortement respectées ; de là, nous écrirons une méthode plus générale où les conditions aux limites seront faiblement respectées ; enfin, nous aborderons une variante, la méthode de Baumann et Oden.*

*L'idée est que la solution numérique peut présenter des sauts afin de mieux approximer la solution là où elle présente une forte pente. Par conséquent, on commettra un crime variationnel en cherchant l'approximation numérique en dehors de l'espace dans lequel se trouve la solution du problème faible.*

*Nous appliquerons cette méthode à l'étude de l'équation hyperbolique et de l'équation d'advection-diffusion.*

## 3.1 Méthode de Nitsche

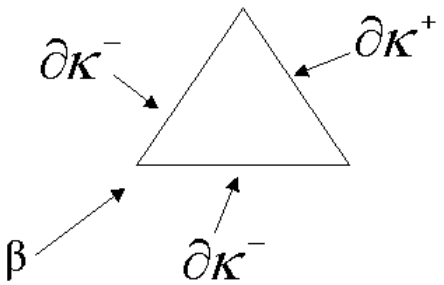
### Conditions de Dirichlet fortement imposées

Définissons tout d'abord les parties entrantes et sortantes de $\partial\kappa$, les traces intérieures et extérieures de $v$ ainsi que le saut $[v]$ qui nous permettront de faire apparaître des discontinuités dans la solution numérique.

**Définition 3** *Soit $\partial\kappa$ la frontière de $\kappa$. Les parties entrantes et sortantes de $\partial\kappa$ sont données respectivement par :*

$$\partial\kappa^- = \{x \in \partial\kappa : \beta \cdot n_\kappa(x) < 0\}, \qquad \partial\kappa^+ = \{x \in \partial\kappa : \beta \cdot n_\kappa(x) \geq 0\}$$

*où $n_\kappa(x)$ est la normale sortante en $x \in \partial\kappa$.*

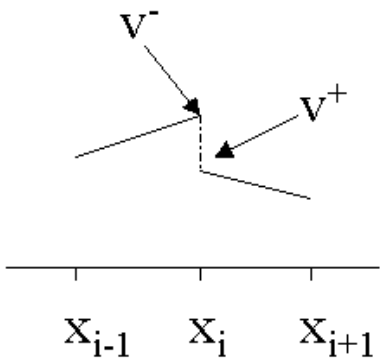



**Définition 4** *Les traces intérieures $v^+(x)$ et $v^-(x)$ où $x \in \partial\kappa$ valent :*

$$v^+(x) = \lim_{s\to 0^+} v(x+s\beta), \qquad v^-(x) = \lim_{s\to 0^-} v(x+s\beta).$$

**Définition 5** *Le saut $[v]$ est défini par :*

$$[v] = v^+ - v^-.$$




Le problème faible s'écrit alors :
trouver $u \in U_0(\Omega)$ tel que

$$a_{AD,2}(u,v) = b_{AD,2}(v) \qquad \forall v \in U_0(\Omega). \tag{3.1}$$

où

$$\begin{aligned} a_{AD,2}(u,v) = \sum_{\kappa\in\mathcal{T}} \Big\{ & \int_\kappa \epsilon\nabla u\cdot\nabla v + (\beta\cdot\nabla u)v dx \\ & -\epsilon\int_{\partial\kappa^-}[v]\langle\nabla u\cdot n_\kappa(x)\rangle ds - \epsilon\int_{\partial\kappa^-}[u]\langle\nabla v\cdot n_\kappa(x))\rangle ds \\ & +\eta_\kappa\int_{\partial\kappa^-}[u][v]ds \\ & -\int_{\partial\kappa^-}(\beta\cdot n_\kappa(x))[u]v^+ds\Big\} \end{aligned} \tag{3.2}$$

et

$$b_{AD}(v) = \sum_{\kappa\in\mathcal{T}}\Big\{\int_\kappa S v dx\Big\} \tag{3.3}$$

où on a défini

$$\langle\nabla u\cdot n_\kappa(x)\rangle = \frac{1}{2}(\nabla u^-\cdot n_\kappa(x) + \nabla u^+\cdot n_\kappa(x)). \tag{3.4}$$

Le premier terme de $a_{AD,2}(u,v)$ provient de la méthode continue ; le deuxième de l'intégration par parties de $v\Delta u$ sur $\kappa$, le troisième a pour but de symétriser le problème lorsque $\beta$ est nul et d'assurer que le saut est proche de zéro.

Le quatrième fait en sorte que la coercivité et la continuité de $a_{AD,2}(u,v)$ sont vérifiées lorsque $\eta$ est assez grand (et pour $\beta$ nul) (on trouvera dans [8] une démonstration), il assure également que le saut est proche de zéro de même que le cinquième terme qui provient de l'application du théorème de la divergence au terme hyperbolique $\int_\kappa(\beta\cdot\nabla u)v dx$.

Supposons $\beta$ nul, $a_{AD,2}(u,v)$ étant symétrique, coercive et continue, le problème a une et une seule solution (par le théorème de Lax-Milgram) ; de plus, on peut trouver une interprétation du terme de pénalité : en effet, le problème faible qui vient d'être énoncé est équivalent au problème de minimisation

$$\min_{u\in U_0(\Omega)} J(u) = \frac{1}{2}a_{AD,2}(u,u) - b_{AD,2}(u). \tag{3.5}$$

Ce qui veut dire que si on remplace $\eta$ par $\eta+\eta_0$, on aura

$$\begin{aligned} J(u) = & \frac{1}{2}a_{AD,2,\eta}(u,u) - b_{AD,2,\eta}(u) \\ & +\eta_0\Big\{\int_{\partial\kappa^-}[u]^2 ds\Big\}. \end{aligned} \tag{3.6}$$

Tout cela pour dire que plus $\eta_0$ est grand, plus l'intégrale du carré du saut de $u$ entre deux éléments sera petite ; plus intuitivement, plus $\eta_0$ est grand, plus $u$ est "continue".

La méthode numérique vient alors immédiatement. Il faut tout d'abord se donner un espace de fonctions discontinues :

**Définition 6**

$$V_0^h(\Omega, \mathcal{T}) = \{v^h \in L^2(\Omega) : v^h|_\kappa \in P_{i_\kappa}(\kappa) \text{ et } v^h = 0 \text{ sur } \Gamma_D\}$$

*où $P_{i_\kappa}(\kappa)$ est l'espace de dimension finie des polynômes de degré égal ou inférieur à $i_\kappa$ et définis sur $\kappa$.*

Il faut attirer l'attention sur le fait que $V_0^h(\Omega, \mathcal{T}) \nsubseteq U_0(\Omega)$ et, par conséquent, on commet un *crime variationnel* en choisissant cet espace de dimension finie n'appartenant pas à l'espace des fonctions dans lequel on cherche la solution.

La méthode numérique s'énonce :
trouver $u_h \in V_0^h(\Omega, \mathcal{T})$ tel que :

$$a_{AD,2}(u_h, v_h) = b_{AD,2}(v_h) \qquad \forall v_h \in V_0^h(\Omega, \mathcal{T}). \tag{3.7}$$

**Conditions aux limites faiblement imposées**

Une fois la méthode de Nitsche aux conditions de Dirichlet fortement imposées, la généralisation au cas général est immédiate. On suppose qu'on a des conditions aux limites répondant aux définitions (1.2), (1.3) et (1.4). L'espace des fonctions continues sera $U(\Omega)$ avec $u$ quelconque sur $\partial\Omega$ et $U_0(\Omega) \subset U(\Omega)$.

Le problème faible s'écrit :
trouver $u \in U(\Omega)$ tel que

$$a_{AD}(u, v) = b_{AD}(v) \qquad \forall v \in U(\Omega) \tag{3.8}$$

où

$$
\begin{aligned}
a_{AD}(u,v) = \sum_{\kappa\in\mathcal{T}} \Big\{ & \int_\kappa \epsilon\nabla u\cdot\nabla v + (\beta\cdot\nabla u)v dx - \epsilon\int_{\partial\kappa\cap\Gamma_D} v(\nabla u\cdot n_\kappa(x))ds \\
& -\epsilon\int_{\partial\kappa^-\setminus\Gamma_D} [v]\langle\nabla u\cdot n_\kappa(x)\rangle ds - \epsilon\int_{\partial\kappa\cap\Gamma_D} u(\nabla v\cdot n_\kappa(x))ds \\
& -\epsilon\int_{\partial\kappa^-\setminus\Gamma_D} [u]\langle\nabla v\cdot n_\kappa(x))\rangle ds + \eta_\kappa\int_{\partial\kappa\cap\Gamma_D} uv ds \\
& +\eta_\kappa\int_{\partial\kappa^-\setminus\Gamma_D} [u][v]ds - \int_{\partial\kappa^-\setminus\Gamma_-} (\beta\cdot n_\kappa(x))[u]v^+ ds \\
& -\int_{\partial\kappa^-\cap\Gamma_-} (\beta\cdot n_\kappa(x))u^+v^+ ds\Big\}
\end{aligned}
\tag{3.9}
$$

et

$$
\begin{aligned}
b_{AD}(v) = \sum_{\kappa\in\mathcal{T}} \Big\{ & \int_\kappa Sv dx - \epsilon\int_{\partial\kappa\cap\Gamma_D} \nabla v\cdot n_\kappa(x) f ds \\
& +\int_{\partial\kappa\cap\Gamma_N} vg ds + \eta_\kappa\int_{\partial\kappa\cap\Gamma_D} vf ds \\
& -\int_{\partial\kappa^-\cap\Gamma_-} (\beta\cdot n_\kappa(x))fv^+ ds\Big\}.
\end{aligned}
\tag{3.10}
$$

On a simplement ajouté le fait que les conditions aux limites devaient être satisfaites faiblement.
La méthode numérique s'écrit dès lors immédiatement :

trouver $u_h \in V^h(\Omega,\mathcal{T})$ tel que :

$$a_{AD}(u_h, v_h) = b_{AD}(v_h) \qquad \forall v_h \in V^h(\Omega,\mathcal{T})$$

où

$$V^h(\Omega,\mathcal{T}) = \{v^h \in L^2(\Omega) : v^h|_\kappa \in P_{i_\kappa}(\kappa)\}.$$

Notons qu'on a $\tilde{V}_0^h \subset V_0^h \subset V^h$ et que $V^h \nsubseteq U(\Omega)$. Ici aussi, on commet un crime variationnel.

### 3.1.1 Application à l'équation hyperbolique

On a donc $\epsilon = 0$. Le problème s'écrit alors :

$$\beta \cdot \nabla u = S \qquad \text{sur } \Omega \ , \tag{3.11}$$

$$u = f \qquad \text{sur } \Gamma_- \ . \tag{3.12}$$

Si on pose, en outre, que $\eta_\kappa$ est nul, la formulation faible présente une structure particulière. En effet, la méthode se réécrit :

trouver $u_h \in V^h(\Omega, \mathcal{T})$ tel que

$$a_H(u_h, v) = b_H(v_h) \qquad \forall v_h \in V^h(\Omega, \mathcal{T}) \tag{3.13}$$

où on a

$$\begin{aligned} a_H(u_h, v_h) \;=\; & \sum_{\kappa \in \mathcal{T}} \int_\kappa (\beta \cdot \nabla u) v dx \\ & - \sum_{\kappa \in \mathcal{T}} \int_{\partial \kappa^- \setminus \Gamma_-} (\beta \cdot n_\kappa(x))[u] v^+ ds \\ & - \sum_{\kappa \in \mathcal{T}} \int_{\partial \kappa^- \cap \Gamma_-} (\beta \cdot n_\kappa(x)) u^+ v^+ ds \end{aligned} \tag{3.14}$$

et

$$b_H(v_h) = \sum_{\kappa \in \mathcal{T}} \int_\kappa S v dx - \sum_{\kappa \in \mathcal{T}} \int_{\partial \kappa^- \cap \Gamma_-} (\beta \cdot n_\kappa(x)) f v^+ ds. \tag{3.15}$$

Soit un élément $\kappa \in \mathcal{T}$ et supposons à présent $u^-$ connu sur $\partial\kappa^-$. Alors on a, sur un seul élément et similairement à (3.13), la méthode suivante : trouver $u_h \in P_{i_\kappa}(\kappa)$ tel que

$$a_\kappa(u_h, v_h) = b_\kappa(v_h) \qquad \forall v_h \in P_{i_\kappa}(\kappa) \tag{3.16}$$

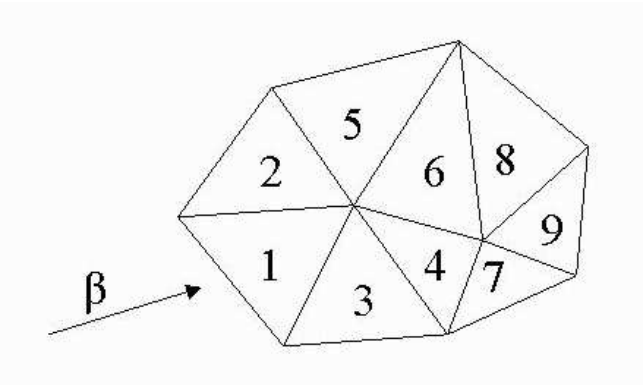
où on définit

$$a_\kappa(u, v) = \int_\kappa (\beta \cdot \nabla u) v dx - \int_{\partial \kappa^-} u^+ v^+ (\beta \cdot n_\kappa(x)) ds \tag{3.17}$$

et

$$b_\kappa(v) = \int_\kappa S v dx - \int_{\partial \kappa^-} u^- v^+ (\beta \cdot n_\kappa(x)) ds. \tag{3.18}$$

Ce n'est rien d'autre que le problème (3.13) mais formulé sur un seul sous-élément de $\Omega$. Donc, si $u^-$ est donné sur $\partial\kappa^-$, on peut déterminer $u$ sur




$\kappa$ (et a fortiori, $u^-$ sur $\partial\kappa^+$). De cette manière, on peut calculer $u$ de proche en proche en commençant par les éléments $\kappa$ tels que $\partial\kappa^- \subset \Gamma_-$ sachant que $u^-$ y vaut $f$.

Dans ce petit exemple, on calculera d'abord $u$ sur le triangle 1 puique les conditions d'entrée $f$ y sont entièrement connues. Connaissant $u$ sur le triangle 1, on peut calculer $u$ sur le triangle 2, les conditions d'entrée étant données par $f$ sur un côté et par $u$ sur le deuxième. On peut progresser de cette manière pour calculer $u$ de triangle en triangle.
Si $\eta_\kappa$ avait été non nul, nous n'aurions pas pu procéder de proche en proche car les équations auraient été couplées entre elles.

### Mise en oeuvre

Nous avons subdivisé notre domaine $\Omega$ en triangles $\kappa$ quelconques. Nous nous sommes également choisi un espace $V^h(\Omega, \mathcal{T})$ où $i_\kappa = 1$ pour tout $\kappa$ (ce n'est rien d'autre que l'espace des fonctions linéaires par morceau). En d'autres mots,

$$u_h = \sum_{\kappa \in \mathcal{T}} \sum_{i=1}^{3} U_{\kappa,i} \phi_{\kappa,i} \tag{3.19}$$

où $\phi_{\kappa,i}$ est nul sur $\Omega \setminus \kappa$ et est linéaire sur $\kappa$ avec, si on considère le prolongement de $\phi_{\kappa,i}$ sur $\overline{\kappa}$,

$$\phi_{\kappa,i} = \begin{cases} 1 & \text{sur le sommet } i \text{ du triangle } \overline{\kappa}, \\ 0 & \text{sur les deux autres sommets.} \end{cases}$$

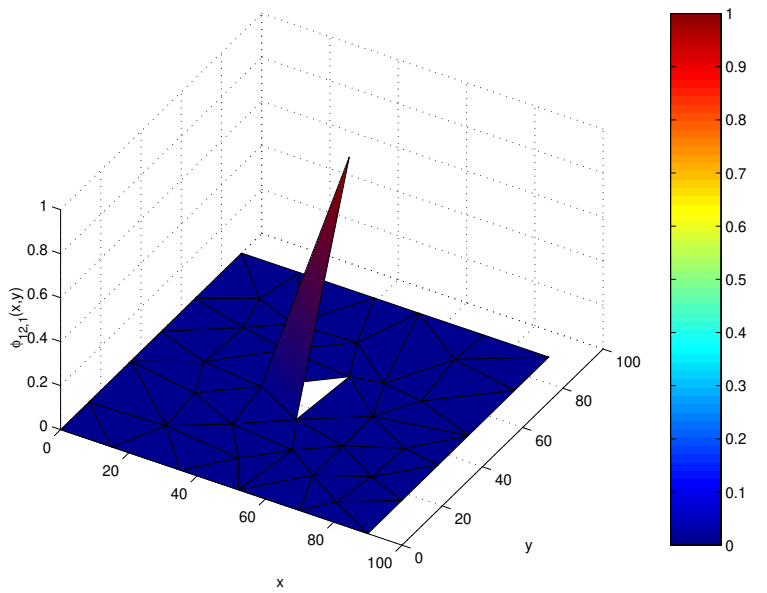



FIG. 3.1 – Fonction de forme

Nous devons donc résoudre le système matriciel donné par

$$\sum_{j=1}^{N} U_j . a_H(\phi_j, \phi_k) = b_H(\phi_k). \tag{3.20}$$

où N est le nombre de triangles multiplié par 3.

Reprenons l'exemple présenté précédemment en (1.11) et (1.12). Il est parfaitement résolu (aux erreurs d'arrondi près) si la ligne de discontinuité (c'est-à-dire la droite $y = x$) coincide avec la triangulation (Figure 3.2). Par contre, dans la Figure 3.3, on a modifié le paramètre $\beta$ : $\beta = (1, 2)$. La ligne de discontinuité représentée par la droite $y = 2x$ ne coincide pas avce la triangulation et par conséquent, le problème n'est pas parfaitement résolu.

Dans les figures (3.4) et (3.5) (qui représentent la solution du même problème, mais sous différents angles de vue), on a repris le $\beta$ d'origine mais on a travaillé sur un maillage légèrement modifié. Là non plus, il n'y a pas de coincidence entre la maillage et la ligne de discontinuité et donc, la solution fournie n'est pas la solution exacte contrairement au maillage de la Figure 3.2. Notons cependant qu'on trouve la solution exacte en amont de la modification apportée au maillage, c'est dû au fait que la solution en aval

n'influence pas la solution en amont.

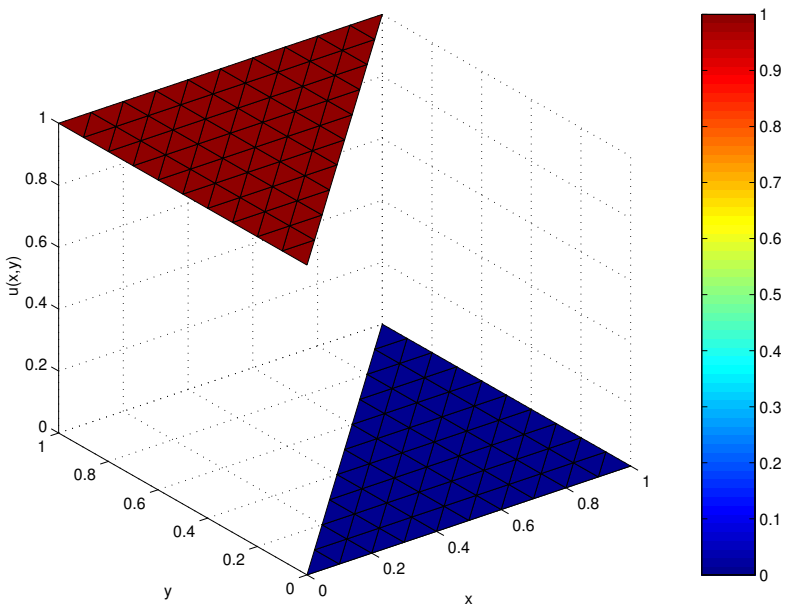



FIG. 3.2 – Solution du Problème 1, parfaite coincidence entre le maillage et la ligne de discontinuité $y = x$

Signalons que la matrice $a_H(\phi_j, \phi_k)$ présente une structure particulière : une numérotation intelligente des triangles peut la rendre triangulaire inférieure par blocs (3x3). Cela vient du fait qu'on peut calculer la solution $u$ de triangle en triangle, ce qui se traduit matriciellement par le fait que pour calculer $U_j$, $U_{j+1}$ et $U_{j+2}$, on n'a besoin que de $U_1, \cdots, U_{j-1}$.
Cette triangulation n'aurait pas été rendue possible si on avait pris un paramètre de stabilisation $\eta$ non nul.

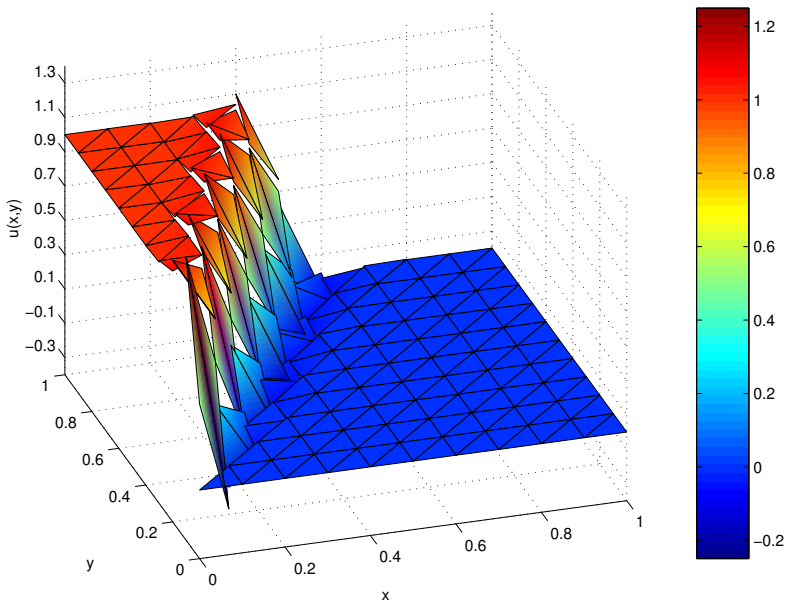



FIG. 3.3 – Solution du Problème 1 avec un $\beta$ modifié

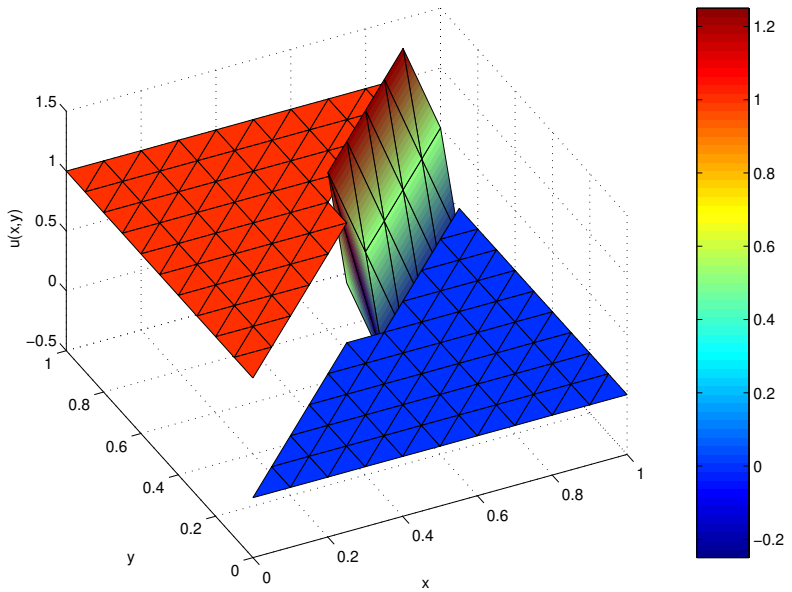



FIG. 3.4 – Solution du Problème 1 avec un maillage légèrement modifié

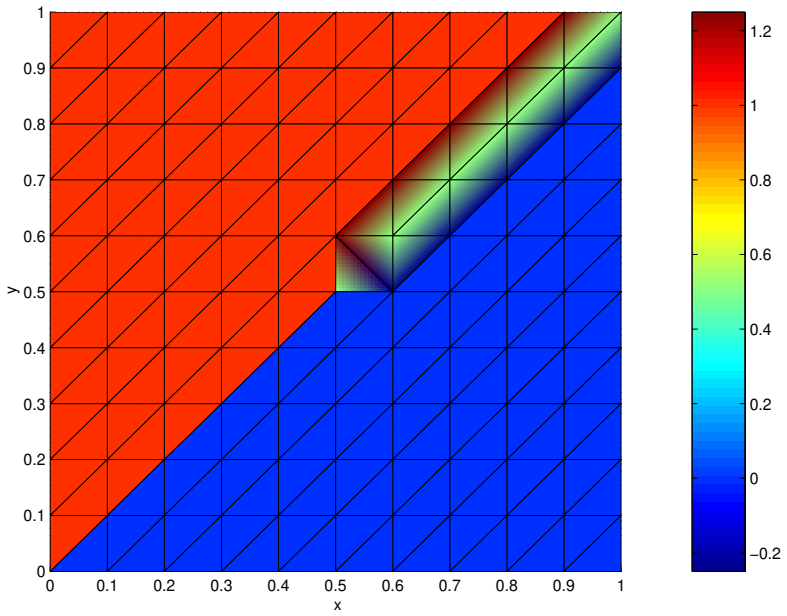



FIG. 3.5 – Solution du Problème 1 avec un maillage légèrement modifié, angle de vue différent

### 3.1.2 Application à l'équation d'advection-diffusion

Nous avons implémenté la méthode de Nitsche en reprenant le même espace discret $V^h(\Omega, \mathcal{T})$ que pour l'équation hyperbolique. Nous avons redéfini

$$\eta_\kappa = \frac{\gamma}{diam(\kappa)} \tag{3.21}$$

où $diam(\kappa)$ est le *diamètre* du triangle. On prendra comme valeur du diamètre le diamètre convexe de l'élément, ici, le plus grand côté du triangle. Le paramètre $\gamma$ est une constante qu'on va se charger de déterminer au mieux.

Signalons que la méthode de résolution du système linéaire décrite pour le problème d'advection pure ne peut plus fonctionner ici. En effet, comme il s'agit d'une équation elliptique, chaque élément est influencé par ses voisins. Matriciellement, on n'aura donc plus une structure triangulaire. On a alors choisi la méthode de la triangulation de Gauss-Jordan (il existe des méthodes moins coûteuses mais là n'est pas l'objectif de ce travail).

On s'est également donné un problème présentant une couche limite et une couche limite de sortie :

**Problème 2**

$$-\epsilon\Delta u + \beta \cdot \nabla u = 0 \qquad avec\ max(x,y) \leq 1\ et\ min(x,y) \geq 0, \tag{3.22}$$

*où, par défaut,* $\epsilon = 0.001$ *et* $\beta = (\cos(10°), \sin(10°))'$.
*De plus,*

$$u(0,y) = \begin{cases} 1 & si\ y > \frac{1}{2} \\ 0 & sinon; \end{cases}$$

$$u(x,0) = u(1,y) = 0$$

*et*

$$\frac{\partial u}{\partial y}(x,1) = 0.$$

Nous avons donc calculé la solution numérique pour différentes valeurs de $\gamma$, le coefficient de pénalité :
$\gamma = 0$, $\gamma = 0.035$, $\gamma = 1$ et $\gamma = 10$.

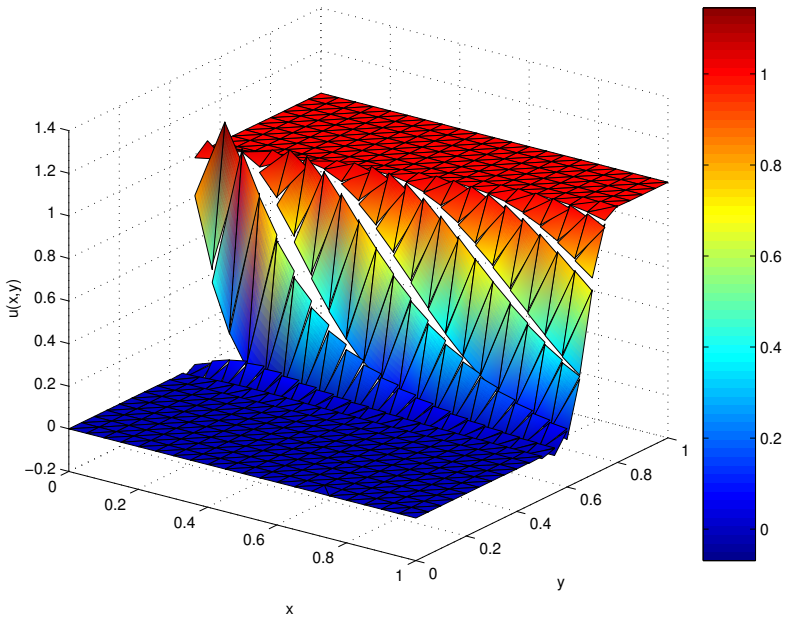



FIG. 3.6 – Problème 2, $\gamma$ nul

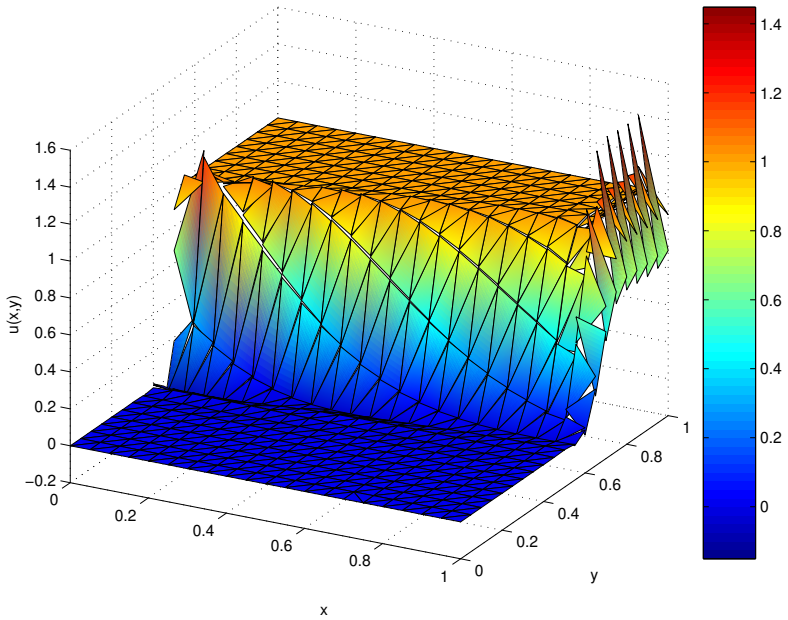



FIG. 3.7 – Problème 2, $\gamma = 0.035$

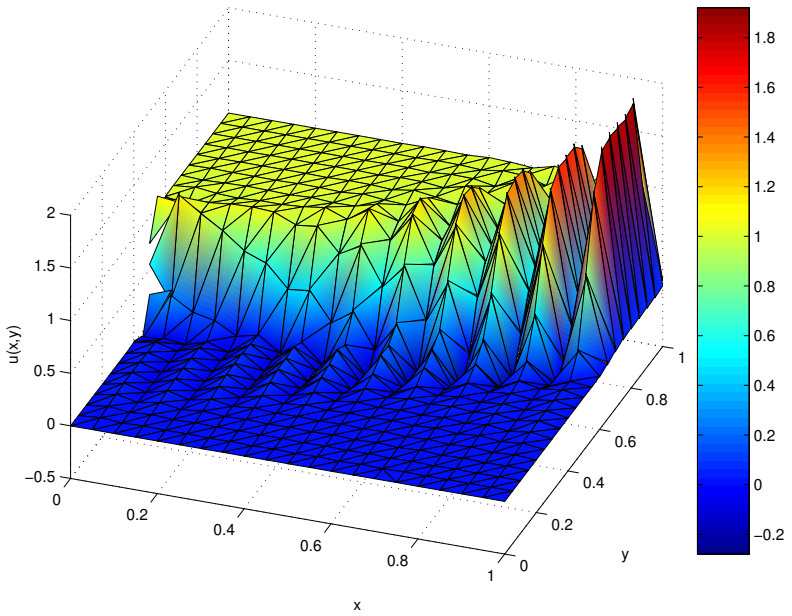



FIG. 3.8 – Problème 2, $\gamma = 1.0$

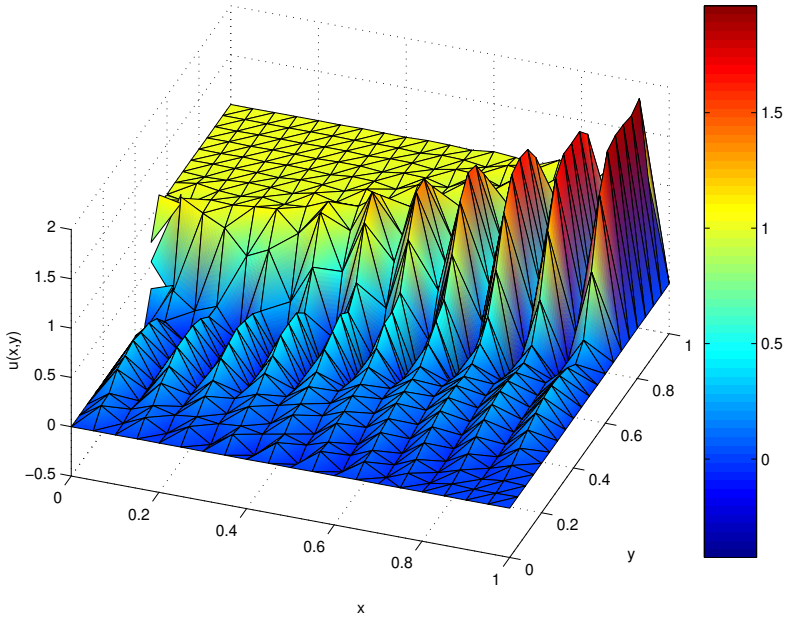



FIG. 3.9 – Problème 2, $\gamma = 10.0$

Ce qu'on avait annoncé concernant l'influence du paramètre de pénalité $\gamma$ se vérifie : quand $\gamma$ est nul, les conditions de Dirichlet à la sortie ne sont pas du tout approximées tandis que quand $\gamma$ grandit, elles sont de mieux en mieux approximées mais au prix d'oscillations qui apparaissent.

Il y a donc un $\gamma$ optimal qu'il faut déterminer. On va, pour cela, étudier l'erreur d'approximation sur un problème dont on connaît la solution.

**Problème 3**

$$-\epsilon\Delta u + \beta \cdot \nabla u = 0 \qquad avec\ max(x,y) \leq 1\ et\ min(x,y) \geq 0, \tag{3.23}$$

*et où* $\epsilon = 0.01$ *et* $\beta = (1,0)'$.
*En outre,*

$$u(0,y) = 1$$
$$u(1,y) = 0$$

*et*

$$\frac{\partial u}{\partial y}(x,1) = \frac{\partial u}{\partial y}(x,0) = 0.$$

Comme le problème précédent, ce problème présente une couche limite de sortie en $x = 1$ (couche limite de sortie qui était la cause de nos problèmes). La solution exacte est donnée par

$$u(x,y) = 1 - \frac{\exp(\frac{x}{\epsilon}) - 1}{\exp(\frac{1}{\epsilon}) - 1} \tag{3.24}$$

On calcule alors l'erreur, c'est-à-dire l'approximation de laquelle on retire la solution exacte. On choisira comme norme la norme suggérée dans [4] :

$$||e||_h^2 = ||e||_{L^2}^2 + \sum_{\kappa \in \mathcal{T}} h_\kappa^2 ||\nabla e \cdot \nabla e||_{L_2}^2. \tag{3.25}$$

Ici, $h_\kappa$ représente la taille du maillage. Comme on s'est choisi un maillage régulier pour traîter cet exemple, $h_\kappa$ est la longueur du côté adjacent à l'angle droit de nos triangles. On calcule la norme de l'erreur pour différentes valeurs de $\gamma$ et de $h_\kappa$.

On se rend compte dans la Figure 3.10 qu'il y a une valeur de $\gamma$ qui minimise la norme de l'erreur, cette valeur tourne aux alentours de 0.2. Notons qu'on obtient sensiblement le même graphe pour une norme $H^1$ plus classique.
La Figure 3.11 représente la solution du Problème 2 avec le $\gamma$ optimal qui vient d'être calculé.

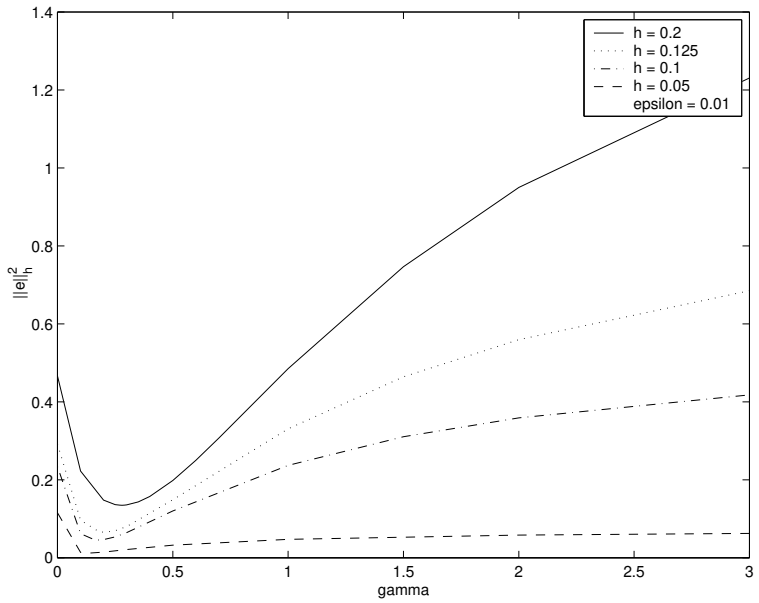



FIG. 3.10 – Carré de la norme de l'erreur d'approximation du Problème 3 en fonction de $\gamma$ et pour différentes tailles du maillage

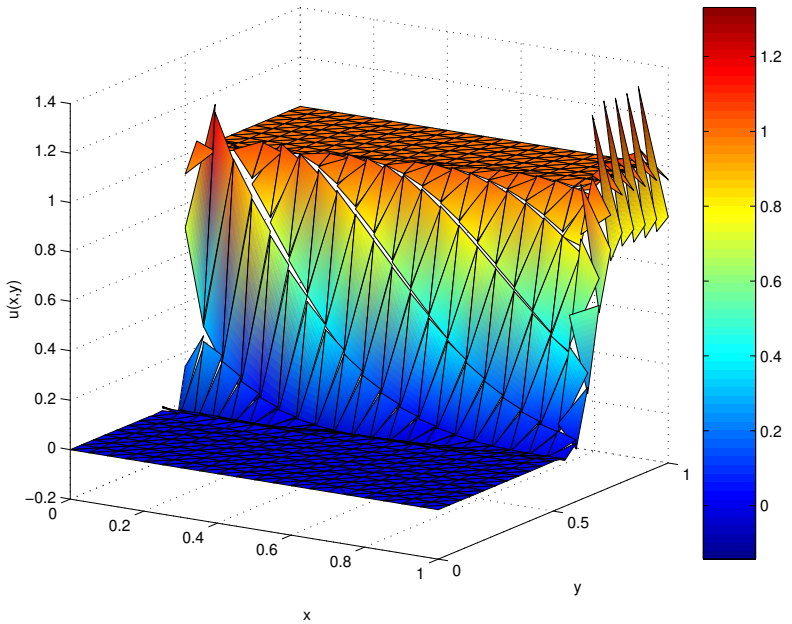



FIG. 3.11 – Problème 2, $\gamma$ optimal

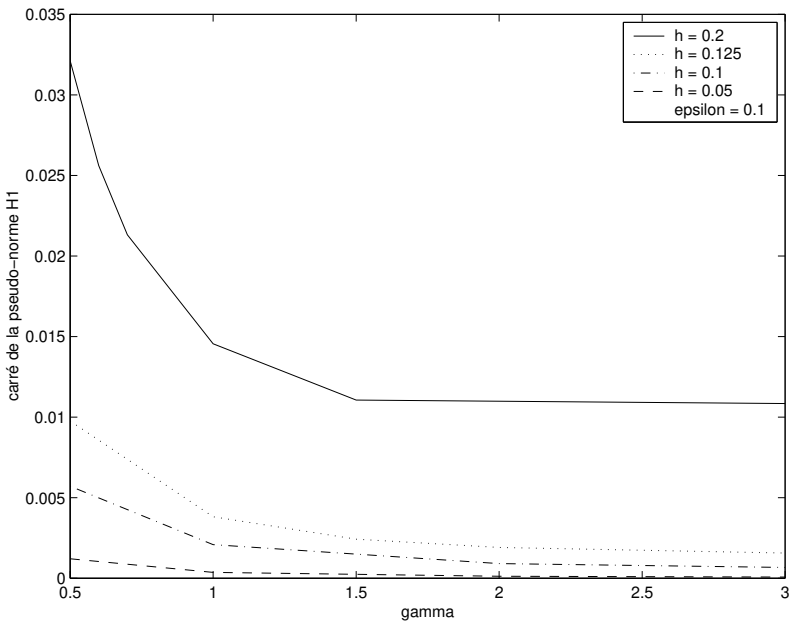



FIG. 3.12 – Carré de la norme de l'erreur d'approximation du Problème 3 en fonction de $\gamma$ et pour différentes tailles du maillage, $\epsilon = 0.1$

On peut refaire le même graphe mais pour un Problème 3 légèrement modifié : on va poser à présent le coefficient de diffusion $\epsilon$ égal à 0.1 (contre 0.01 auparavant).

Nous pouvons voir dans la Figure 3.12 que le minimum n'est jamais atteint, en d'autres mots, que la fonction est strictement monotone. Cela veut dire que le $\gamma$ optimal est infini ou, en d'autres mots, une approximation continue est meilleure. Ceci est parfaitement logique car la largeur de la couche limite vaut $O(\epsilon) = 0.1$, or, tous les maillages sont plus petits ou du même ordre que cette largeur, la couche limite peut alors être bien approchée par une approximation continue. On en déduit qu'il vaut mieux, pour ce cas, utiliser une méthode continue plutôt qu'une méthode discontinue et plus coûteuse.

Pour le cas $\epsilon = 0.01$, la taille des maillages était bien plus grande que la couche limite de sortie, c'est pour cela que la méthode discontinue donnait une meilleure approximation.

A la lumière de ce qui vient d'être dit, on est en droit de se demander à partir de quelle valeur du coefficient de diffusion $\epsilon$, pour un maillage fixé, on a un infimum de $||e||_h$ en l'infini ; autrement dit, pour quel $\epsilon$ vaut-il mieux

choisir une méthode continue plutôt qu'une méthode discontinue ? Le graphe suivant (Figure 3.13) répond à cette question.

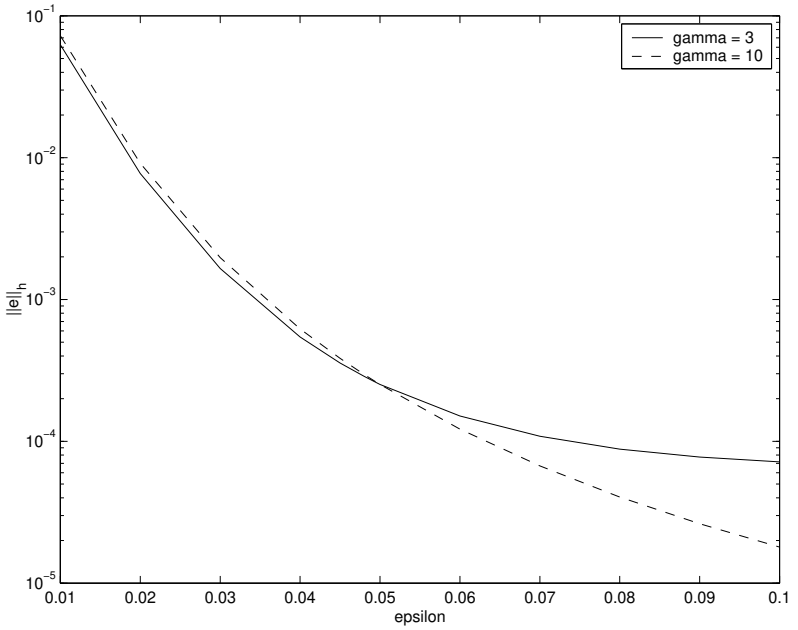



FIG. 3.13 – Norme de l'erreur en fonction de $\epsilon$ avec $\gamma = 3$ et $\gamma = 10$

Ce graphique donne la norme de l'erreur du Problème 3 en fonction du coefficient de diffusion $\epsilon$ et ce, pour deux valeurs de $\gamma$ différentes (3 et 10). Le maillage a une taille de 0.05.

Nous pouvons voir que pour des valeurs de $\epsilon$ inférieures à 0.05, la norme de l'erreur en $\gamma = 3$ est inférieure à la norme de l'erreur en $\gamma = 10$. On en déduit que le minimum de la norme de l'erreur est atteint en des $\gamma$ bornés. Par contre, si $\epsilon > 0.05$, l'infimum de la norme de l'erreur sera atteint en un $\gamma$ infini (puisque $\gamma = 10$ est meilleur que $\gamma = 3$ et que $\gamma = 10$ est assimilé à l'infini). On retrouve le résultat annoncé, si la taille du maillage est inférieure à la taille de la couche limite, une méthode continue est à conseiller (car moins coûteuse) ; dans le cas contraire, on préfèrera une méthode discontinue comme celle qui vient d'être exposée.

## 3.2 Méthode de Baumann-Oden

Un des objectifs de cette méthode est de rendre la matrice de rigidité mieux conditionnée et ce, en changeant le signe du troisième terme de (3.2). Si on passe directement au cas général, la méthode s'énonce :
trouver $u \in U(\Omega)$ tel que

$$a_{OB}(u,v) = b_{OB}(v) \qquad \forall v \in U(\Omega) \tag{3.26}$$

avec

$$\begin{aligned} a_{OB}(u,v) = \sum_{\kappa\in\mathcal{T}} \Big\{ & \int_\kappa \epsilon\nabla u\cdot\nabla v + (\beta\cdot\nabla u)v dx - \epsilon\int_{\partial\kappa\cap\Gamma_D} v(\nabla u\cdot n_\kappa(x))ds \\ & -\epsilon\int_{\partial\kappa^-\setminus\Gamma_D} [v]\langle\nabla u\cdot n_\kappa(x)\rangle ds + \epsilon\int_{\partial\kappa\cap\Gamma_D} u(\nabla v\cdot n_\kappa(x))ds \\ & +\epsilon\int_{\partial\kappa^-\setminus\Gamma_D} [u]\langle\nabla v\cdot n_\kappa(x))\rangle ds + \eta_\kappa\int_{\partial\kappa\cap\Gamma_D} uv ds \\ & +\eta_\kappa\int_{\partial\kappa^-\setminus\Gamma_D} [u][v]ds + \int_{\partial\kappa^-\setminus\Gamma_-} (\beta\cdot n_\kappa(x))[u]v^+ ds \\ & -\int_{\partial\kappa^-\cap\Gamma_-} (\beta\cdot n_\kappa(x))u^+v^+ ds\Big\} \end{aligned} \tag{3.27}$$

et

$$\begin{aligned} b_{OB}(v) = \sum_{\kappa\in\mathcal{T}} \Big\{ & \int_\kappa Sv dx + \epsilon\int_{\partial\kappa\cap\Gamma_D} \nabla v\cdot n_\kappa(x) f ds \\ & + \int_{\partial\kappa\cap\Gamma_N} vg ds + \eta_\kappa\int_{\partial\kappa\cap\Gamma_D} vf ds \\ & + \int_{\partial\kappa^-\cap\Gamma_-} (\beta\cdot n_\kappa(x))fv^+ ds\Big\}. \end{aligned} \tag{3.28}$$

Cette méthode présente deux avantages, elle ne requiert pas de terme de pénalité si on utilise des polynômes d'ordre supérieur à deux et, pour des polynômes linéaires, elle est plus stable : en effet, si on étudie le nombre de condition des matrices obtenues par la méthode de Nitsche ou de Baumann-Oden, la seconde méthode est plus favorable(nombre de condition plus petit). On peut s'en rendre compte en regardant la Figure 3.14 : on y a porté le nombre de condition de la méthode de Baumann-Oden divisé par le nombre de condition de la méthode de Nitsche en fonction de $\gamma$(Problème 3, h = 0.125).

Remarquons que pour $\gamma$ tendant vers l'infini, ce rapport tend vers 1. C'est on ne peut plus logique puisque si $\gamma$ tend vers l'infini, le terme de pénalité

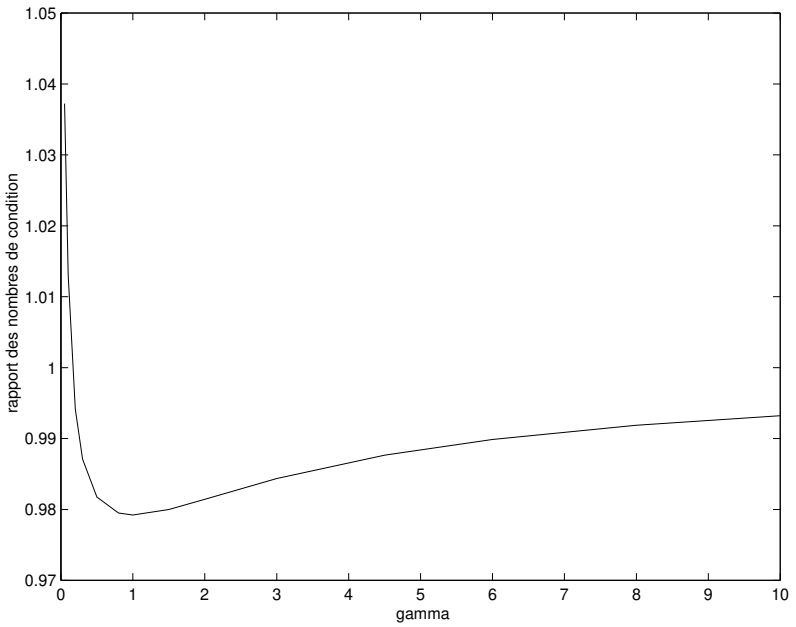



FIG. 3.14 – Nombre de condition de la matrice de rigidité obtenue par la méthode de Baumann-Oden divisé par celui obtenu par la méthode de Nitsche en fonction de $\gamma$

sera dominant dans la matrice de rigidité or, ce terme de pénalité est identique dans les deux méthodes. Donc, lorsque $\gamma$ tend vers l'infini, les matrices de rigidité sont identiques et les nombres de condition égaux.

On peut également se demander si les deux méthodes donnent des résultats significativement différents. Nous avons dès lors calculé la solution du Problème 2 avec le paramètre de stabilité optimal et nous avons comparé le résultat avec la solution précédemment calculée(Figure 3.15).

On s'aperçoit que la différence maximale est de l'ordre de 0.05, cette différence étant la plus marquée dans les couches limites. On peut donc conclure que les méthodes donnent des résultats proches (pour des fonctions de forme linéaires).

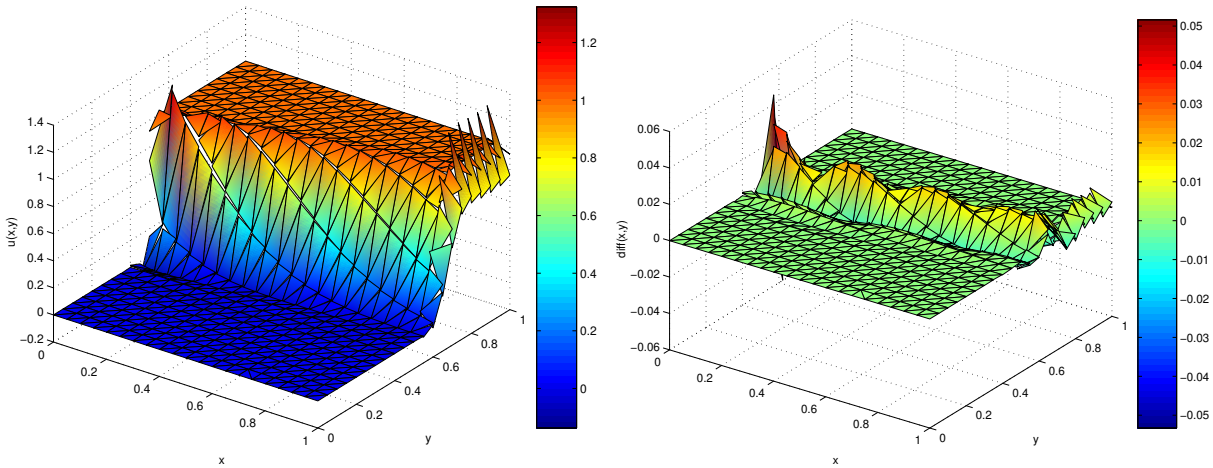



FIG. 3.15 – Solution du Problème 2 calculée par la méthode de Baumann et Oden, différence des solutions obtenues par les deux méthodes différentes

# Chapitre 4

# Problème d'évolution

*Dans ce chapitre, on va brièvement s'intéresser cas transitoire de l'équation d'advection-diffusion. On sait que la stabilité d'un problème d'évolution est liée aux valeurs propres du système discrétisé. Nous allons donc étudier le comportement des valeurs propres des matrices de rigidité du problème stationnaire.*

L'équation différentielle transitoire s'écrit :

$$\frac{\partial u}{\partial t} + \nabla \cdot (\beta u) - \epsilon \Delta u = S \qquad \text{sur } \Omega \text{ x } [0, T] \ . \tag{4.1}$$

Les conditions de Neumann et de Dirichlet sont connues pour tout $t \in$ [0,T]. De plus, on connaît $u(., 0)$, c'est-à-dire qu'on connaît $u$ en $t = 0$. Enfin, $S$ dépend de $x$ et de $t$.

Le schéma numérique de la méthode d'Euler s'écrit :

$$\frac{1}{\Delta t}(\mathbf{u}(t + \Delta t) - \mathbf{u}(t)) = -\mathbf{A}_{AD,h}\mathbf{u}(t) + \mathbf{B}_{AD,h} \tag{4.2}$$

où $\mathbf{A}_{AD,h}$ est la matrice de rigidité du problème et où $\mathbf{u}(t)$ est le vecteur solution $[U_1(t),\ U_2(t), \cdots,\ U_N(t)]$.

Nous avons testé cette méthode pour le problème transitoire associé au Problème 1 (problème hyperbolique) présenté auparavant.

Néanmoins, la théorie nous dit que, pour que la solution n'explose pas, il faut que les valeurs propres de $\mathbf{A}_{AD,h}$ se trouvent dans le disque de rayon $\frac{1}{\Delta t}$ et de centre $(\frac{1}{\Delta t}, 0)$. Vérifions que c'est bien le cas pour la matrice de rigidité du Problème 1 (Figure 4.2).

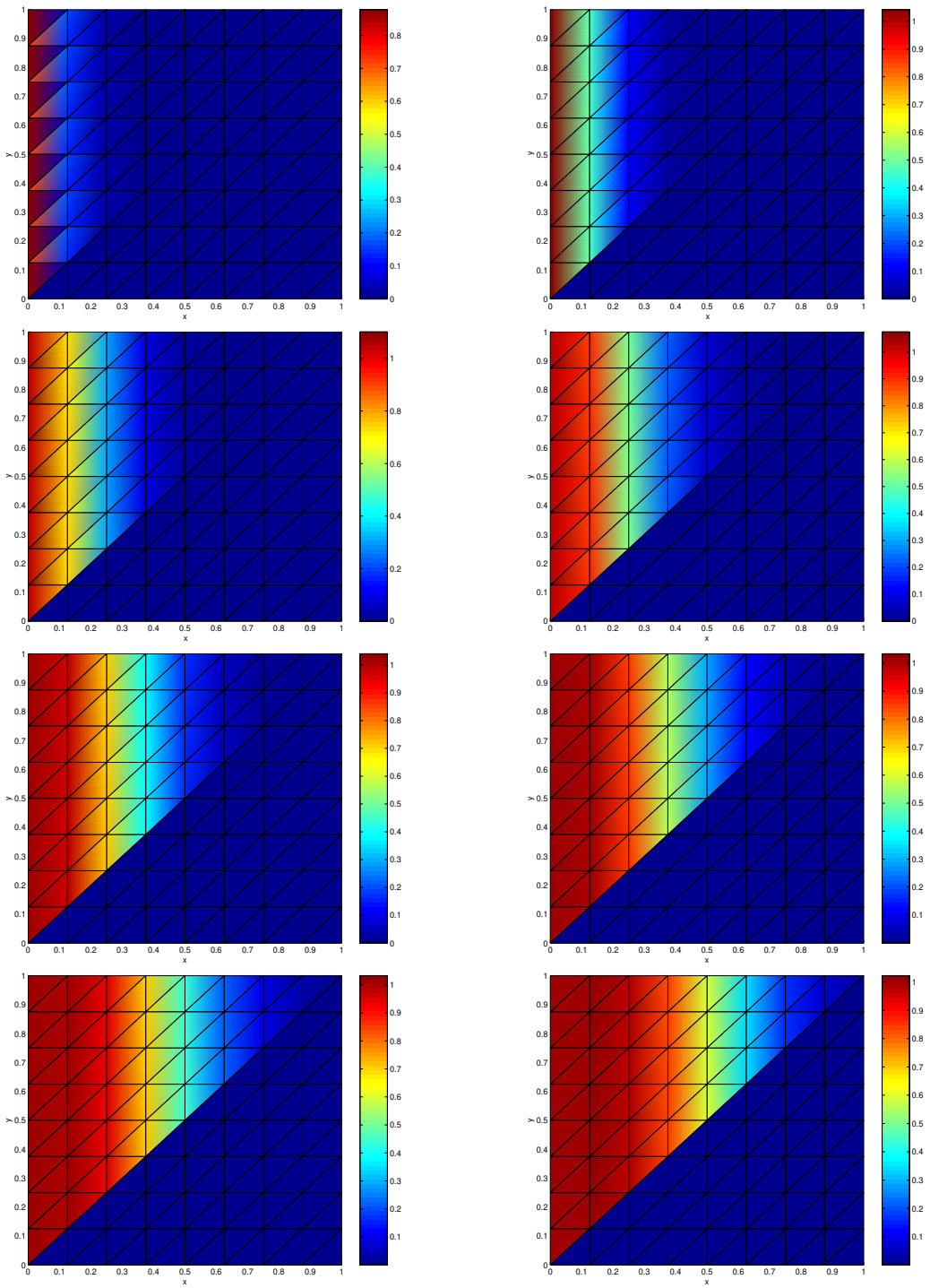


FIG. 4.1 – Problème transitoire associé au Problème 1, $\Delta t = 0.1$

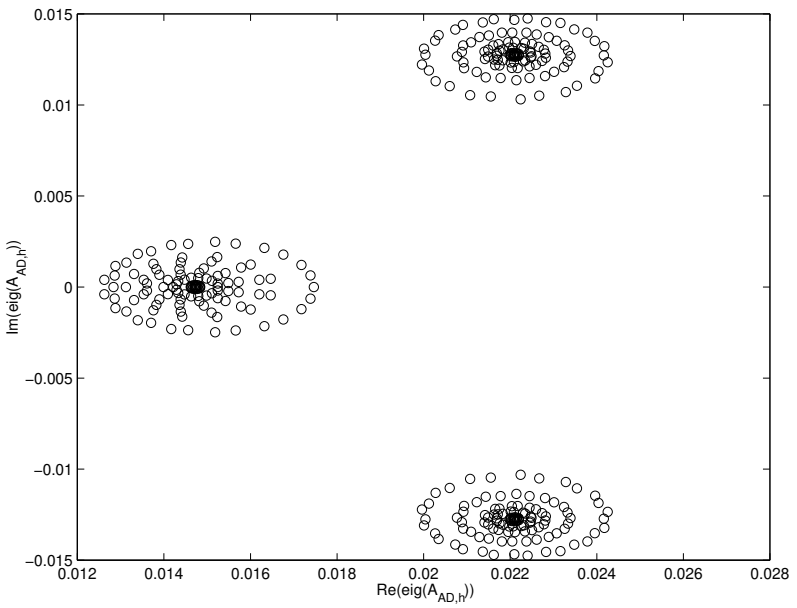



FIG. 4.2 – Valeurs propres de la matrice de rigidité $\mathbf{A}_{AD,h}$ associée au Problème 1 (problème hyperbolique)

Un petit calcul nous indique que le $\Delta t$ maximal qu'on peut choisir vaut 62.03.

Etudions alors les valeurs propres de la matrice de rigidité $A_{AD,h}$ associée à un problème d'advection-diffusion et ce, pour différentes valeurs du paramètre de stabilisation $\gamma$. Nous avons porté dans la Figure 4.3 les valeurs propres de la matrice de rigidité du problème stationnaire 3 (où la taille $h$ du maillage vaut 0.125) et ce, pour $\gamma = 0.2, 1.0$ et 3.0.

On s'aperçoit que, plus $\gamma$ augmente, plus le maximum de la partie réelle des valeurs propres augmente, ce qui a un effet désastreux sur le $\Delta t$ maximal qu'on peut se donner.

La Figure 4.4 nous donne l'évolution du $\Delta t$ maximal en fonction du paramètre de stabilisation $\gamma$. On en déduit que, si on utilise la méthode d'Euler explicite, le $\Delta t$ diminue au fur et à mesure que $\gamma$ augmente. On préfèrera dès lors les méthodes d'Euler implicite ou des trapèzes pour lesquelles les régions de stabilité sont le demi-plan des nombres complexes à partie réelle positive (en effet, quel que soit le $\gamma$ qu'on se donne, les valeurs propres sont à partie réelle positive).

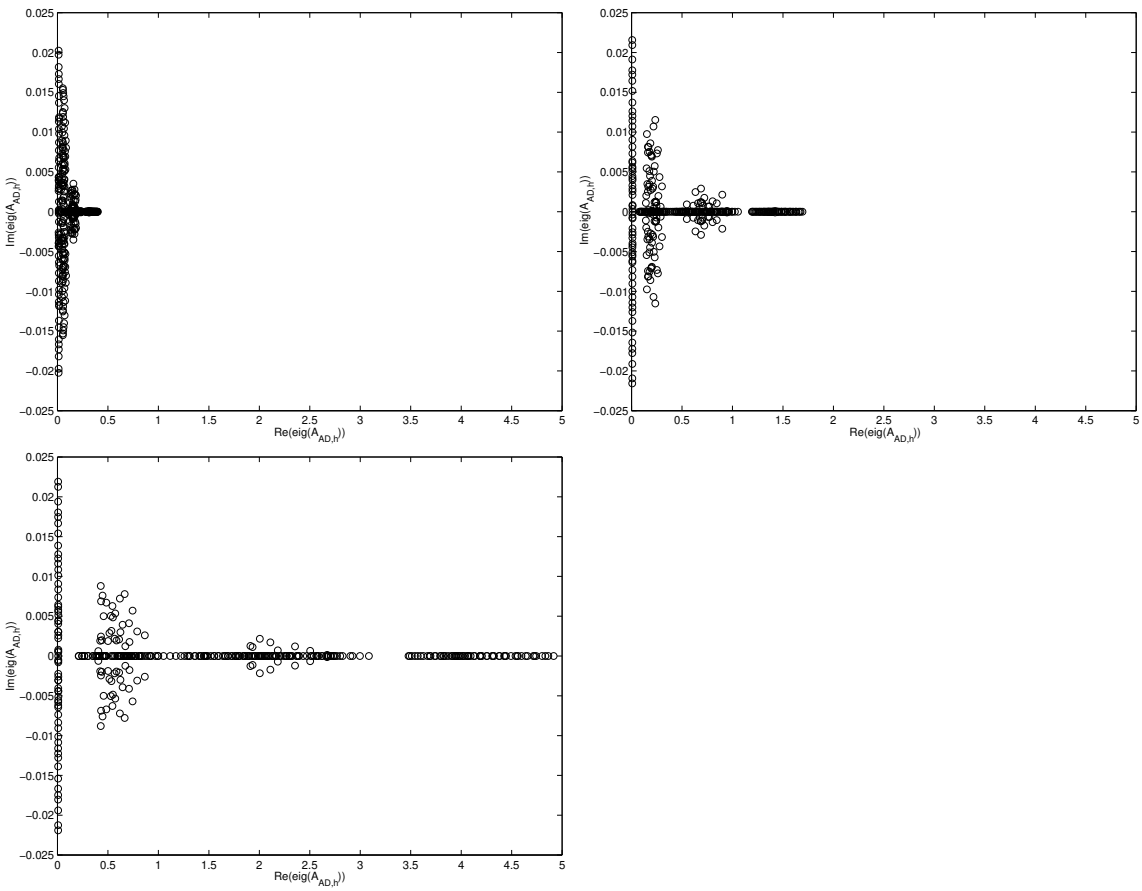



FIG. 4.3 – Valeurs propres de la matrice de rigidité $A_{AD,h}$ pour $\gamma$ valant 0.2, 1 et 3 (problème elliptique)

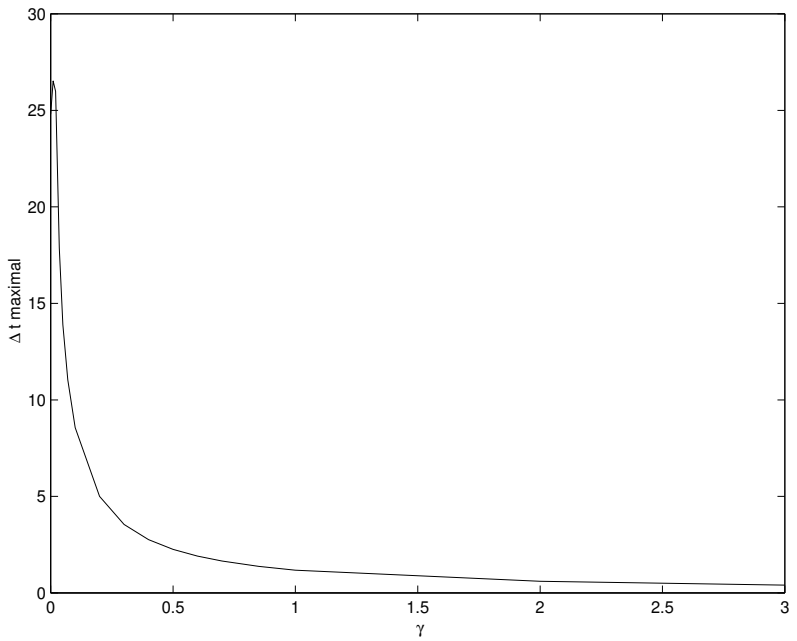



FIG. 4.4 – Evolution du $\Delta t$ maximal en fonction de $\gamma$

# Deuxième partie

# Estimations d'erreur

# Chapitre 5

# Estimations d'erreur a priori

*Nous allons nous attacher à étudier dans ce chapitre la convergence de la norme énergétique $a_{AD}(e,e)$ et d'autres normes usuelles.*

Etudions tout d'abord la norme énergétique. On trouvera dans [6] une preuve que l'ordre de convergence de la méthode de Galerkin discontinue est, borné par $\mathcal{O}(h^{\frac{1}{2}})$, la borne étant atteinte dans le cas purement hyperbolique.

Nous avons porté dans le graphe suivant (Figure 5.1) la valeur de $\ln(\sqrt{a_{AD}(e,e)})$ en fonction de $\ln(\frac{1}{h})$ où $e$ est l'erreur du Problème 3 (calculé avec un paramètre de stabilisation $\gamma$ égal à 0.2). L'ordre de convergence est alors donné par la pente de la courbe.

La valeur absolue de la pente sur le dernier intervalle est de 0.4602 ce qui est inférieur à ce qui a été annoncé (0.5). Néanmoins, un calcul ou deux de plus sur un maillage plus fin nous aurait donné une idée plus précise quant à cet ordre de convergence mais cela nous était impossible car les temps de résolution du système linéaire seraient devenus énormes. Cela dit, on se doute que l'ordre de convergence est proche de 0.5 (et supérieur) car le caractère hyperbolique du problème l'emporte sur le caractère elliptique (coefficient de diffusion de 0.01).

Nous avons également étudié de manière expérimentale les propriétés de convergence d'autres normes usuelles.
Tout d'abord pour la norme $L^2$,

$$||e||_{L^2}^2 = \int_\Omega e^2 dx, \tag{5.1}$$

nous obtenons un taux de convergence de 1.35.

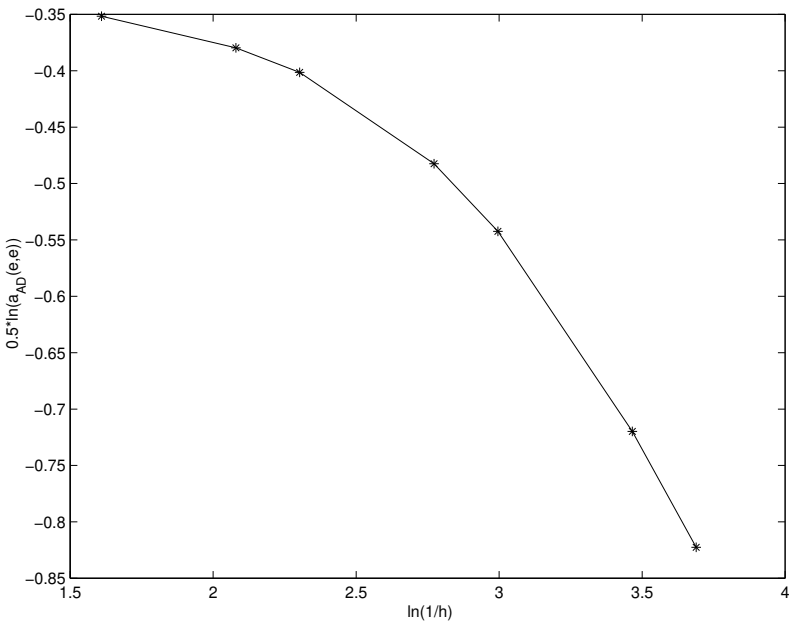



FIG. 5.1 – Convergence de la norme énergétique

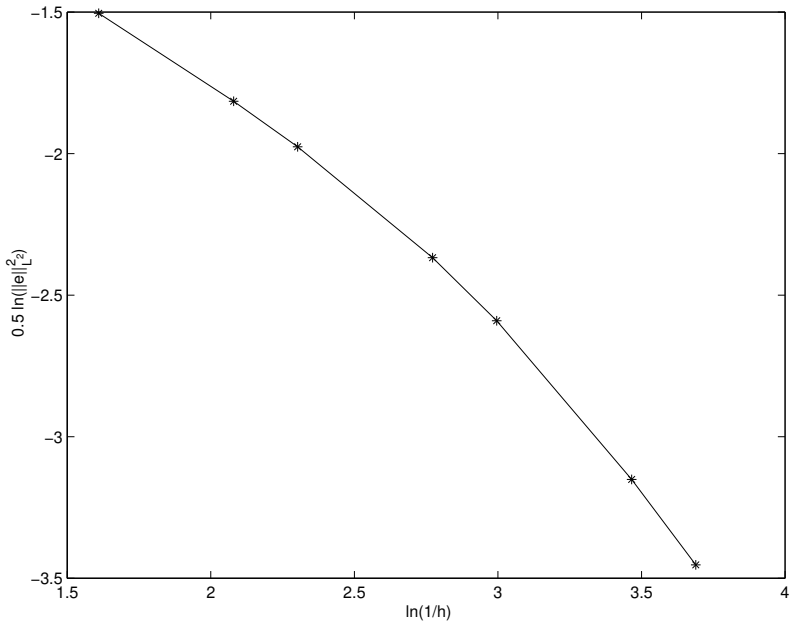



FIG. 5.2 – Convergence de la norme $L^2$

Pour la norme $||.||_h$ présentée en (3.25), le taux de convergence est de 1.66.

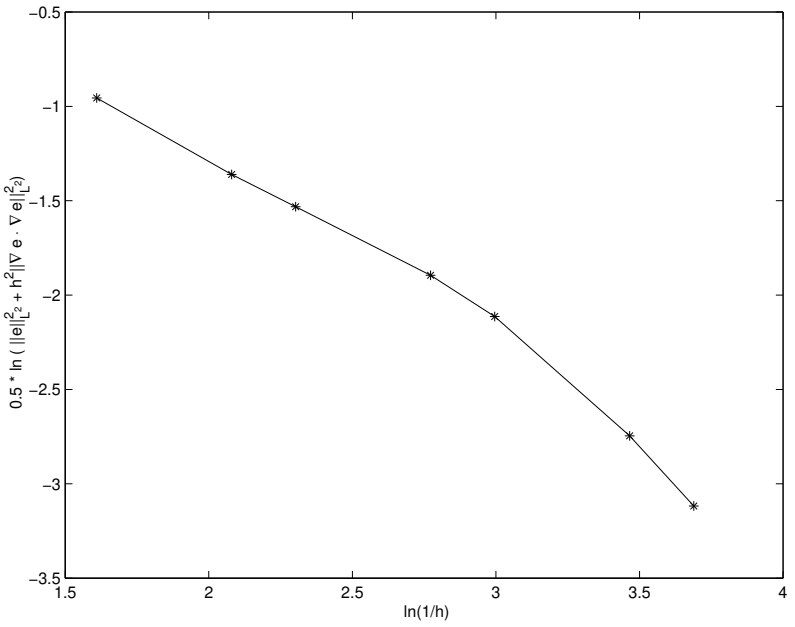



FIG. 5.3 – Convergence de la norme $||.||_h$

Enfin, pour la norme $H^1(\Omega, \mathcal{T})$,

$$||e||^2_{H^1(\Omega,\mathcal{T})} = \sum_{\kappa\in\mathcal{T}} (\int_\kappa e^2 + \nabla e \cdot \nabla e dx), \tag{5.2}$$

on a un taux de convergence de 0.96. On s'aperçoit, dans la Figure 5.4, que la norme augmente, atteint son maximum en $h = 0.05$ puis diminue. Un examen de la Figure 5.5 peut nous aider à comprendre ce phénomène. On y voit que, sur le maillage de taille 0.05, la solution présente des gradients plus importants que sur le maillage de taille 0.2. Par contre, les discontinuités sont plus faibles sur le maillage de taile 0.05 que sur le maillage de taille 0.2. Or, la norme $H^1(\Omega, \mathcal{T})$ ne prend pas en compte les discontinuités mais bien les gradients, ce qui explique que la norme $H^1(\Omega, \mathcal{T})$ croît avant de decroître, tandis que la norme énergétique, qui tient compte des discontinuités, est décroissante partout.

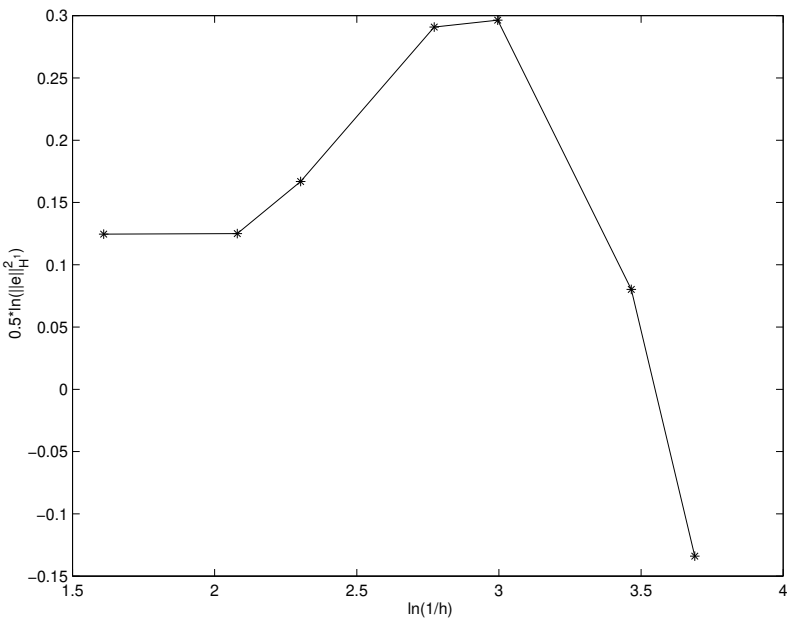



FIG. 5.4 – Convergence de la norme $H^1(\Omega, \mathcal{T})$

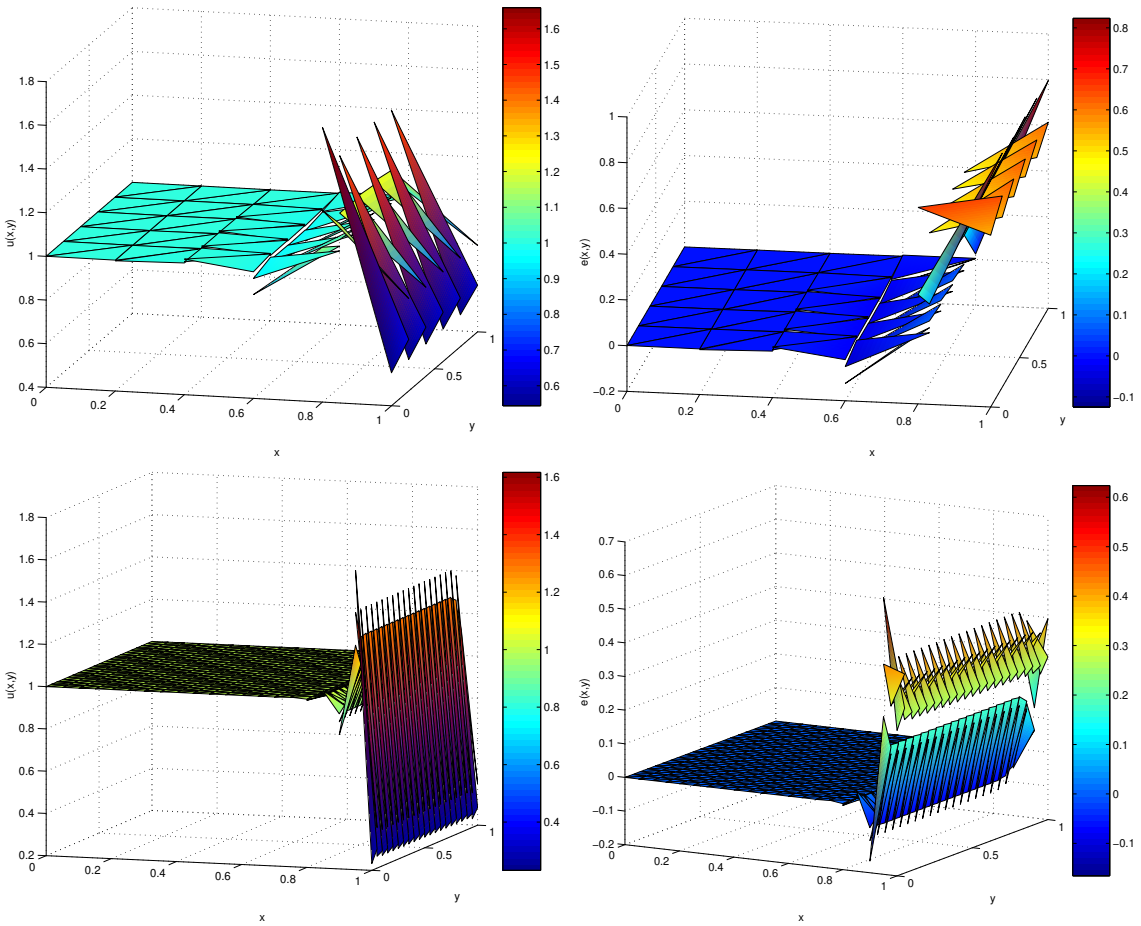



FIG. 5.5 – Solutions du Problème 3 (gauche) et erreurs correspondantes (droite) sur des maillage de taille 0.2 et 0.05

# Chapitre 6

# Estimations d'erreur a posteriori

*L'objectif de cette section est d'établir une méthode permettant de raffiner le maillage là où l'erreur commise est importante et ce, dans le but de réduire l'erreur au sens de la norme énergétique globale. On va, pour cela, évaluer l'erreur résiduelle dans un espace discret contenant strictement l'espace discret dans lequel on a calculé la solution.*

On a

$$a_{AD}(u, v) = a_{AD}(u - u_h + u_h, v) = b_{AD}(v) \qquad \forall v \in U(\Omega) \tag{6.1}$$

et donc, par linéarité,

$$a_{AD}(u - u_h, v) = -a_{AD}(u_h, v) + b_{AD}(v) \qquad \forall v \in U(\Omega) \tag{6.2}$$

où $u$ est la solution exacte de l'équation d'advection-diffusion et $u_h$ la solution calculée.
En d'autres mots,

$$a_{AD}(e, v) = -a_{AD}(u_h, v) + b_{AD}(v) \qquad \forall v \in U(\Omega) \tag{6.3}$$

où $e$ est l'erreur commise. Cette équation nous servira pour estimer l'erreur, néanmoins, on ne peut calculer $b_{AD}(v) - a_{AD}(u_h, v)$ pour toutes les fonctions appartenant à $U(\Omega)$. En outre, calculer $b_{AD}(v) - a_{AD}(u_h, v)$ avec $v \in V^h(\Omega, \mathcal{T})$ ne servirait à rien puisque $b_{AD}(v) - a_{AD}(u_h, v)$ serait alors nul (par construction).

Il faut donc se donner un espace contenant $V^h(\Omega, \mathcal{T})$. On va appeler cet espace $V^h_+(\Omega, \mathcal{T})$ avec

$$V_+^h(\Omega,\mathcal{T}) = V^h(\Omega,\mathcal{T}) \cup B^h(\Omega,\mathcal{T}) \qquad (6.4)$$

où $B^h(\Omega,\mathcal{T})$ est l'espace vectoriel ayant pour base $\phi_\kappa^+$ (fonction bulle telle que définie dans [11]).
$\phi_\kappa^+$ est défini comme suit : $\phi_\kappa^+$ est linéaire par morceau et

$$\phi_\kappa^+ = \begin{cases} 0 & \text{sur } \Omega \setminus \kappa \\ 1 & \text{à l'intersection des médianes de } \kappa. \end{cases}$$

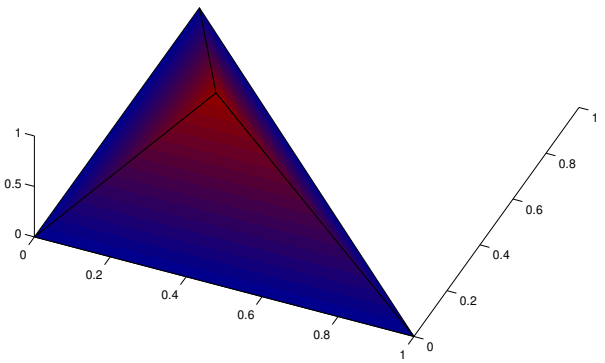



FIG. 6.1 – Fonction bulle

Sachant que $u_h \in V^h(\Omega,\mathcal{T})$, on va calculer l'erreur $a_{AD}(e,e)$ dans $V_+^h(\Omega,\mathcal{T})$ : on peut écrire

$$a_{AD}(e_h^+, v_h^+) = -a_{AD}(u_h, v_h^+) + b_{AD}(v_h^+) \qquad \forall v_h^+ \in V_+^h(\Omega,\mathcal{T}). \qquad (6.5)$$

Or, $V_+^h(\Omega,\mathcal{T}) = V^h(\Omega,\mathcal{T}) \cup B^h(\Omega,\mathcal{T})$ et comme $b_{AD}(v_h^+) - a_{AD}(u_h, v_h^+)$ est nul pour la composante de $v_h^+$ appartenant à $V^h(\Omega,\mathcal{T})$, on a

$$a_{AD}(e_h^+, v_h^+) = -a_{AD}(u_h, v_h^+) + b_{AD}(v_h^+) \qquad \forall v_h^+ \in B^h(\Omega,\mathcal{T}). \qquad (6.6)$$

Ceci étant dit, nous pouvons alors étudier l'erreur commise sur un seul élément. Soit $E_\kappa \in \mathbf{R}$, on obtient

$$a_{AD}(E_\kappa \phi_\kappa^+, \phi_\kappa^+) = -a_{AD}(u_h, \phi_\kappa^+) + b_{AD}(\phi_\kappa^+). \tag{6.7}$$

Travaillons à présent sur $a_{AD}(e,e)$ restreint à $\kappa$ :

$$\begin{aligned}
\sqrt{a_{AD}(e_h^+, e_h^+)} &= \sqrt{a_{AD}(E_\kappa \phi_\kappa^+, E_\kappa \phi_\kappa^+)} \\
&= E_\kappa \sqrt{a_{AD}(\phi_\kappa^+, \phi_\kappa^+)} \\
&= \frac{E_\kappa \cdot a_{AD}(\phi_\kappa^+, \phi_\kappa^+)}{\sqrt{a_{AD}(\phi_\kappa^+, \phi_\kappa^+)}} \\
&= \frac{b_{AD}(\phi_\kappa^+) - a_{AD}(u_h, \phi_\kappa^+)}{\sqrt{a_{AD}(\phi_\kappa^+, \phi_\kappa^+)}}
\end{aligned} \tag{6.8}$$

La méthode est dès lors très simple : il suffit de calculer pour chaque triangle $\kappa$

$$\theta_\kappa = \frac{b_{AD}(\phi_\kappa^+) - a_{AD}(u_h, \phi_\kappa^+)}{\sqrt{a_{AD}(\phi_\kappa^+, \phi_\kappa^+)}}. \tag{6.9}$$

Si

$$\theta_\kappa^2 > \mathrm{TOL}^2 \frac{aire(\kappa)}{aire(\Omega)}, \tag{6.10}$$

où TOL est la tolérance totale requise, on subdivise le triangle $\kappa$ en 4 parties semblables.
On répartit ainsi de manière équivalente l'erreur tolérée sur tous les triangles.

Nous avons testé cette méthode sur le Problème 2 avec une tolérance de 0.1 et un paramètre de stabilisation $\gamma = 0.2$. Nous avons, par conséquent, calculé une première solution sur le maillage de base, raffiné ce maillage par la procédure qui vient d'être décrite, recalculé une nouvelle solution sur le nouveau maillage, raffiné...et ce, par trois fois.

Comme on pouvait s'y attendre, les raffinements des triangles se font là où la solution présente des gradients importants. De plus, on raffine sur le *domaine d'influence* de cette zone.

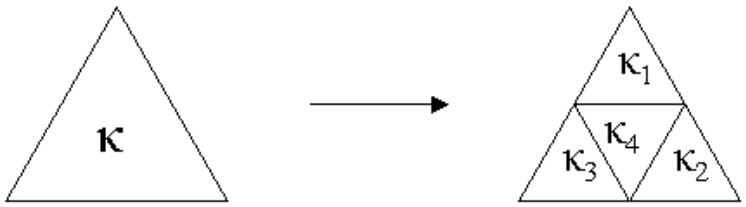



FIG. 6.2 – Subdivision du triangle $\kappa$

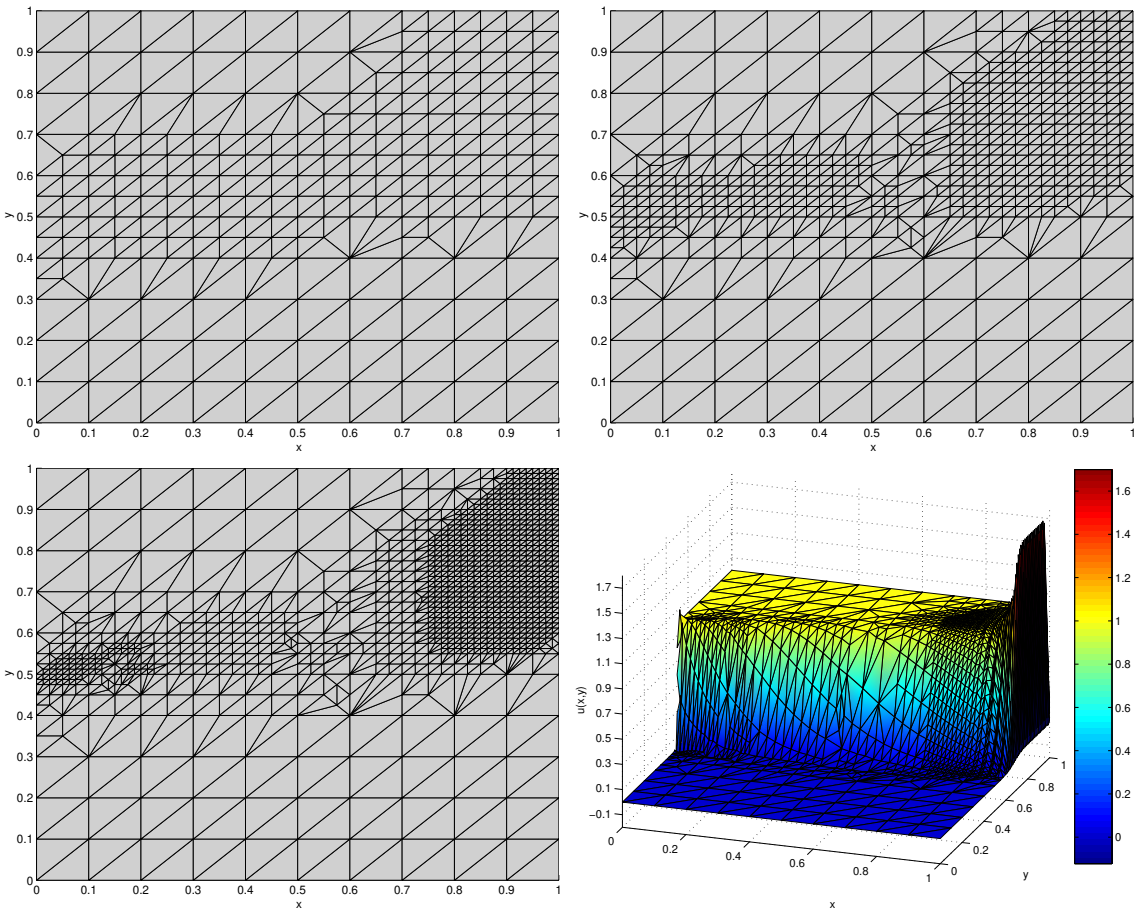



FIG. 6.3 – Raffinements successifs nécessaires à la résolution du Problème 2 et solution obtenue sur le dernier maillage

On remarque que la solution numérique du dernier calcul présente encore une oscillation mais cela est logique puisque la taille minimale du maillage est de 0.0125 soit plus de 10 fois la grandeur de la couche limite.

Etudions à présent les propriétés de convergence de la méthode. Nous avons étudié l'erreur $a_{AD}(e,e)$ commise à chaque itération pour le Problème 3 dont la solution analytique est connue.

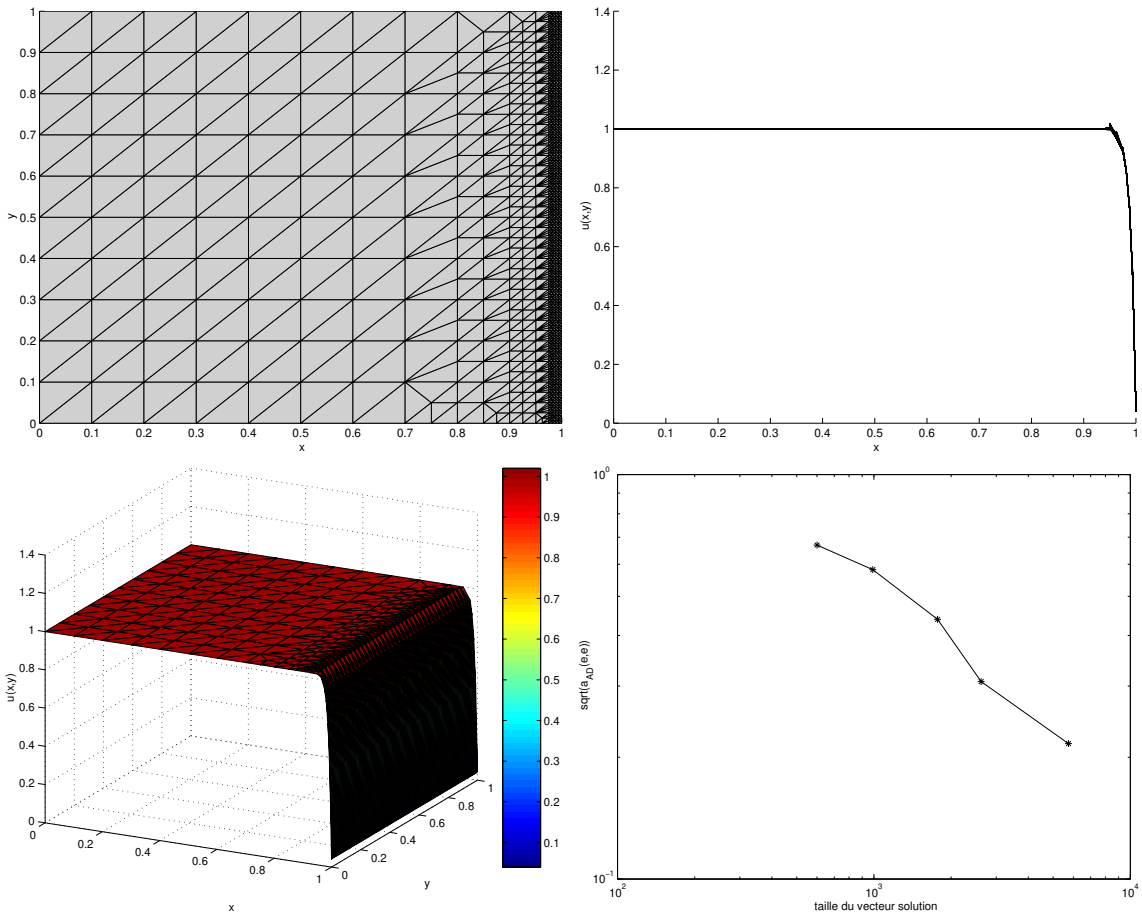



FIG. 6.4 – Solution du Problème 3 sur le dernier maillage calculé, convergence de la norme énergétique en fonction des raffinements successifs (en bas à droite)

Ici, la couche limite de sortie est parfaitement approximée car la taille du maillage (0.00625) est plus petite que la taille de la couche limite de sortie (0.1).

La pente atteint, dans la Figure 6.4 où l'on a porté l'erreur $a_{AD}(e,e)$ commise à chaque itération en fonction du nombre de variables, une valeur absolue maximale de 0.9049.

# Chapitre 7

# Méthodes adjointes et dualité

*L'objectif est de calculer une quantité $J(u)$ où $u$ est la solution de notre problème d'advection-diffusion.*

*C'est fréquemment le cas en ingénierie, par exemple, si $u$ représente une concentration, on peut être intéressé à calculer la moyenne de $u$ sur un sous-ensemble du domaine.*

*Nous allons, par conséquent, établir une méthode afin de déterminer $J(u)$ à une tolérance donnée près.*

## 7.1 Préambule

Jusqu'à présent, nous nous sommes attachés à résoudre le problème variationnel *primal* suivant :
trouver $u \in U(\Omega)$ tel que

$$a_{AD}(u, v) = b_{AD}(v) \qquad \forall v \in U(\Omega) \qquad (P) \tag{7.1}$$

où $a_{AD}(u, v)$ et $b_{AD}(v)$ ont été définis auparavant.

Supposons à présent que nous devons calculer $J(u)$ (que nous supposerons être une fonctionnelle linéaire) pour plusieurs conditions initiales différentes (ce cas se présente également à de nombreuses reprises), c'est-à-dire :
trouver $u_i \in U(\Omega)$ tel que

$$a_{AD}(u_i, v) = b_{AD,i}(v) \qquad \forall v \in U(\Omega) \qquad (P_i), \tag{7.2}$$

$i = 1, \cdots, N$.

Une fois les $u_i$ calculés par $(P_i)$, on calcule les différents $J(u_i)$. Il s'agit là de la *méthode directe* pour laquelle il faut résoudre $N$ systèmes linéaires. Une méthode alternative se présente alors à nous, la *méthode adjointe*, basée sur le problème *dual* :
trouver $z \in U(\Omega)$ tel que

$$a_{AD}(w, z) = J(w) \qquad \forall w \in U(\Omega) \qquad (D). \tag{7.3}$$

On voit directement que

$$a_{AD}(u, z) = J(u) = b_{AD}(z). \tag{7.4}$$

Appliquant alors $(P_i)$ à $(D)$, on a que

$$a_{AD}(u_i, z) = J(u_i) = b_{AD,i}(z). \tag{7.5}$$

Cette dernière équation implique qu'au lieu de résoudre $N$ copies de problème primal $(P_i)$ avec les conditions initiales données par $b_{AD,i}$ puis de calculer $J(u_i)$, il est préférable de résoudre le problème dual $(D)$ et de calculer ensuite $b_{AD,i}(z)$ pour $i = 1, \cdots, N$. Pour $N \gg 1$, la méthode adjointe est bien plus intéressante que la méthode directe.

Evidemment, $(P)$ et $(D)$ ne peuvent être, en général, résolus analytiquement et doivent être résolus de manière numérique. On désirera donc que $J(u)$ soit bien approximé par $J(u_h)$ où $u_h \in V^h(\Omega, \mathcal{T})$. En d'autres mots,

$$|J(u) - J(u_h)| \leq \text{TOL}. \tag{7.6}$$

Or, nous avons

$$\begin{aligned} J(u) - J(u_h) &= a_{AD}(u, z) - a_{AD}(u_h, z) && (7.7)\\ &= b_{AD}(z) - a_{AD}(u_h, z) && (7.8)\\ &= b_{AD}(z - z_h) - a_{AD}(u_h, z - z_h) && (7.9) \end{aligned}$$

puisque

$$a_{AD}(u_h, z_h) = b_{AD}(z_h) \qquad \forall z_h \in V^h(\Omega, \mathcal{T}). \tag{7.10}$$

Bien sûr, la résolution de (7.9) implique le fait qu'on doit connaître $z$, la solution du problème dual, ce qui n'est pas le cas. Il faut alors se donner

un espace dual $\tilde{V}^h(\Omega, \mathcal{T})$. Cet espace aura un maillage $\mathcal{T}$ dual ou un degré polynomial dual.

Repartons de (7.9) : on peut écrire, par linéarité :

$$\begin{aligned} J(u) - J(u_h) &= b_{AD}(z - \tilde{z}_h) - a_{AD}(u_h, z - \tilde{z}_h) \\ &\quad + b_{AD}(\tilde{z}_h - z_h) - a_{AD}(u_h, \tilde{z}_h - z_h), \end{aligned} \tag{7.11}$$

et donc,

$$\begin{aligned} |J(u) - J(u_h)| &\leq |b_{AD}(z - \tilde{z}_h) - a_{AD}(u_h, z - \tilde{z}_h)| \\ &\quad + |b_{AD}(\tilde{z}_h - z_h) - a_{AD}(u_h, \tilde{z}_h - z_h)|. \end{aligned} \tag{7.12}$$

On a donc

$$V^h(\Omega, \mathcal{T}) \subset \tilde{V}^h(\Omega, \mathcal{T}) \subset U(\Omega)$$

et

$$\begin{aligned} u \text{ et } z &\in U(\Omega), \\ u_h \text{ et } z_h &\in V^h(\Omega, \mathcal{T}), \\ \tilde{z}_h &\in \tilde{V}^h(\Omega, \mathcal{T}) \end{aligned}$$

Or, il a été montré au travers d'expérimentations numériques que le premier terme ($|b_{AD}(z - \tilde{z}_h) - a_{AD}(u_h, z - \tilde{z}_h)|$) était un ordre de grandeur plus petit que le second ($|b_{AD}(\tilde{z}_h - z_h) - a_{AD}(u_h, \tilde{z}_h - z_h)|$).
On peut dès lors, sans compromettre la fiabilité de l'algorithme adaptatif, remplacer le critère (7.6) par

$$|b_{AD}(\tilde{z}_h - z_h) - a_{AD}(u_h, \tilde{z}_h - z_h)| < \text{TOL} \tag{7.13}$$

On procède alors de la manière suivante :

- calculer $\tilde{z}_h \in \tilde{V}^h(\Omega, \mathcal{T})$ tel que

$$a_{AD}(\tilde{u}_h, \tilde{z}_h) = J(\tilde{u}_h) \qquad \forall \tilde{u}_h \in \tilde{V}^h(\Omega, \mathcal{T}) \tag{7.14}$$

- calculer $u_h \in V^h(\Omega, \mathcal{T})$ tel que

$$a_{AD}(u_h, v_h) = b_{AD}(v_h) \qquad \forall v_h \in V^h(\Omega, \mathcal{T}) \tag{7.15}$$

- comparer $|b_{AD}(\tilde{z}_h - z_h) - a_{AD}(u_h, \tilde{z}_h - z_h)|$ à TOL où $z_h$ est la composante de $\tilde{z}_h$ dans $V^h(\Omega, \mathcal{T})$.

En particulier, nous nous sommes donné comme espace $\tilde{V}^h(\Omega, \mathcal{T})$ l'espace $V_+^h(\Omega, \tilde{\mathcal{T}})$ défini dans le chapitre précédent.

On a alors

$$(\tilde{z}_h - z_h) = \sum_{\kappa \in \mathcal{T}} Z_\kappa \phi_\kappa^+. \tag{7.16}$$

On peut alors remplacer (7.13) par

$$|b_{AD}(Z_\kappa \phi_\kappa^+) - a_{AD}(u_h, Z_\kappa \phi_\kappa^+)| < \frac{aire(\kappa)}{aire(\Omega)} \mathrm{TOL} \qquad \forall \kappa \in \mathcal{T}. \tag{7.17}$$

Si ce critère n'est pas vérifié pour un triangle, on raffine ce dernier par la procédure présentée dans le chapitre précédent.

Comme fonctionnelle $J(u)$, nous avons choisi la moyenne de $u$ sur un disque de rayon 0.001 inclus dans $\Omega$, c'est-à-dire $\frac{1}{\pi r^2} \int_{\text{disque}} u dx$.

### Signification physique du problème dual

Une question légitime est alors : quelle est la signification physique du problème dual ? Ecrivons le problème $a_{AD}(v, u) = J(v)$ et comparons-le avec le problème $a_{AD}(u, v) = b_{AD}(v)$.

Tout d'abord, commençons par remarquer que $J(u) = \frac{1}{\pi r^2} \int_{\text{disque}} u dx$ correspond à une source de valeur 1 concentrée sur le disque. Il n'y a pas d'autres termes, on en déduit que les conditions de Dirichlet sont nulles ainsi que les conditions de Neumann. En effet, on a

$$\begin{aligned} b_{AD}(v) = \sum_{\kappa \in \mathcal{T}} \Big\{ & \int_\kappa S v dx - \epsilon \int_{\partial\kappa \cap \Gamma_D} \nabla v \cdot n_\kappa(x) f ds \\ & + \int_{\partial\kappa \cap \Gamma_N} v g ds + \eta_\kappa \int_{\partial\kappa \cap \Gamma_D} v f ds \\ & + \int_{\partial\kappa^- \cap \Gamma_-} (\beta \cdot n_\kappa(x)) f v^+ ds \Big\}. \end{aligned} \tag{7.18}$$

qui est bien égal à $J$ si les conditions de Dirichlet sont nulles de même que les conditions de Neumann et si $S$ vaut 1.

Travaillons à présent sur $a_{AD}(v,u)$ et supposons que le domaine est constitué d'un seul élément.

$$\begin{aligned} a_{AD,dual}(v,u) = \int_\Omega \epsilon \nabla u \cdot \nabla v + (\beta_{dual} \cdot \nabla v) u dx - \epsilon \int_{\Gamma_D} u(\nabla v \cdot n(x)) ds \\ -\epsilon \int_{\Gamma_D} v(\nabla u \cdot n(x)) ds + \eta \int_{\Gamma_D} uv ds \\ - \int_{\Gamma_-} (\beta_{dual} \cdot n(x)) u^+ v^+ ds \Big\} \end{aligned} \tag{7.19}$$

Les termes contenant $\epsilon$ ou $\eta_\kappa$ se retrouvent dans $a_{AD}(u,v)$ puisqu'ils sont symétriques. On en déduit que $\Gamma_D$ reste inchangé par le passage au problème dual, la frontière de Dirichlet du problème primal est la frontière de Dirichlet du problème dual.

En outre,

$$\begin{aligned} \int_\Omega (\beta_{dual} \cdot \nabla v) u dx - \int_{\Gamma_-} (\beta_{dual} \cdot n(x)) u^+ v^+ ds = \\ - \int_\Omega (\beta_{dual} \cdot \nabla u) v dx + \int_{\Gamma_+} (\beta_{dual} \cdot n(x)) u^+ v^+ ds = \\ \int_\Omega ((-\beta_{dual}) \cdot \nabla u) v dx - \int_{\Gamma_+} ((-\beta_{dual}) \cdot n(x)) u^+ v^+ . \end{aligned} \tag{7.20}$$

On en déduit que $\beta_{dual} = -\beta$ et donc, que la partie entrante dans le problème primal ($\Gamma_{-,primal}$) devient la partie sortante ($\Gamma_{+,dual}$) dans le problème dual et inversement. La généralisation au cas où nous avons plusieurs éléments est immédiate.

Comme les conditions de Dirichlet du problème $(P)$ sont définies au moins sur $\Gamma_{-,primal} = \Gamma_{+,dual}$ et que, sur $\Gamma_{-,dual}$, la solution est nulle, on en déduit que la frontière de Dirichlet de $(D)$ est donnée par $\Gamma_{+,dual} \cup \Gamma_{-,dual} = \partial\Omega$ et que les conditions y sont nulles.

En bref, la vitesse *duale* est l'opposé de la vitesse *primale* et les conditions de Dirichlet sont nulles sur toute la frontière. Le problème dual détermine ainsi la zone d'influence de la fonctionnelle $J(u)$ ; en d'autres termes, la solution du problème dual fait apparaître les triangles qui influencent le plus la solution là où on a placé notre disque d'intérêt, de plus, l'importance relative de ces triangles dans la solution en ce disque apparaît également.

## 7.2 Mise en oeuvre et analyse des résultats

Nous avons implémenté la procédure qui vient d'être décrite et l'avons testée sur un problème type, le Problème 3. Le centre du disque sur lequel on calcule de la moyenne se situe en (0.9667,0.8333), c'est-à-dire dans la couche limite de sortie.

De plus, comme dans la section consacrée à l'étude a posteriori de l'erreur, nous réappliquons le critère sur la solution qui vient d'être calculée à l'aide de ce dernier.

Enfin, la tolérance TOL sera de 0.01.

On remarque (Figure 7.1) que le système raffine aux alentours du point critique (qui se situe dans la couche limite) mais que toute la couche limite n'est pas raffinée.

En effet, la solution en (0.9667,0.8333) est peu influencée par la solution en (0.9667,0.2), par exemple, et donc, on ne raffine pas aux alentours de (0.9667,0.2) bien que la solution y soit mal approximée.

On peut alors s'intéresser aux propriétés de convergence de l'erreur. C'est ce qui est fait dans la Figure 7.2. On y a porté les valeurs de $J(e) = J(u) - J(u_h)$ en fonction de la taille du vecteur solution de $(P)$ (graphe logarithmique). On peut y voir que l'erreur augmente avant de rediminuer, c'est dû à une oscillation qui disparaît au fur et à mesure des itérations. Remarquons également que la dernière itération du cas où la tolérance est de 0.5 n'apporte plus rien.

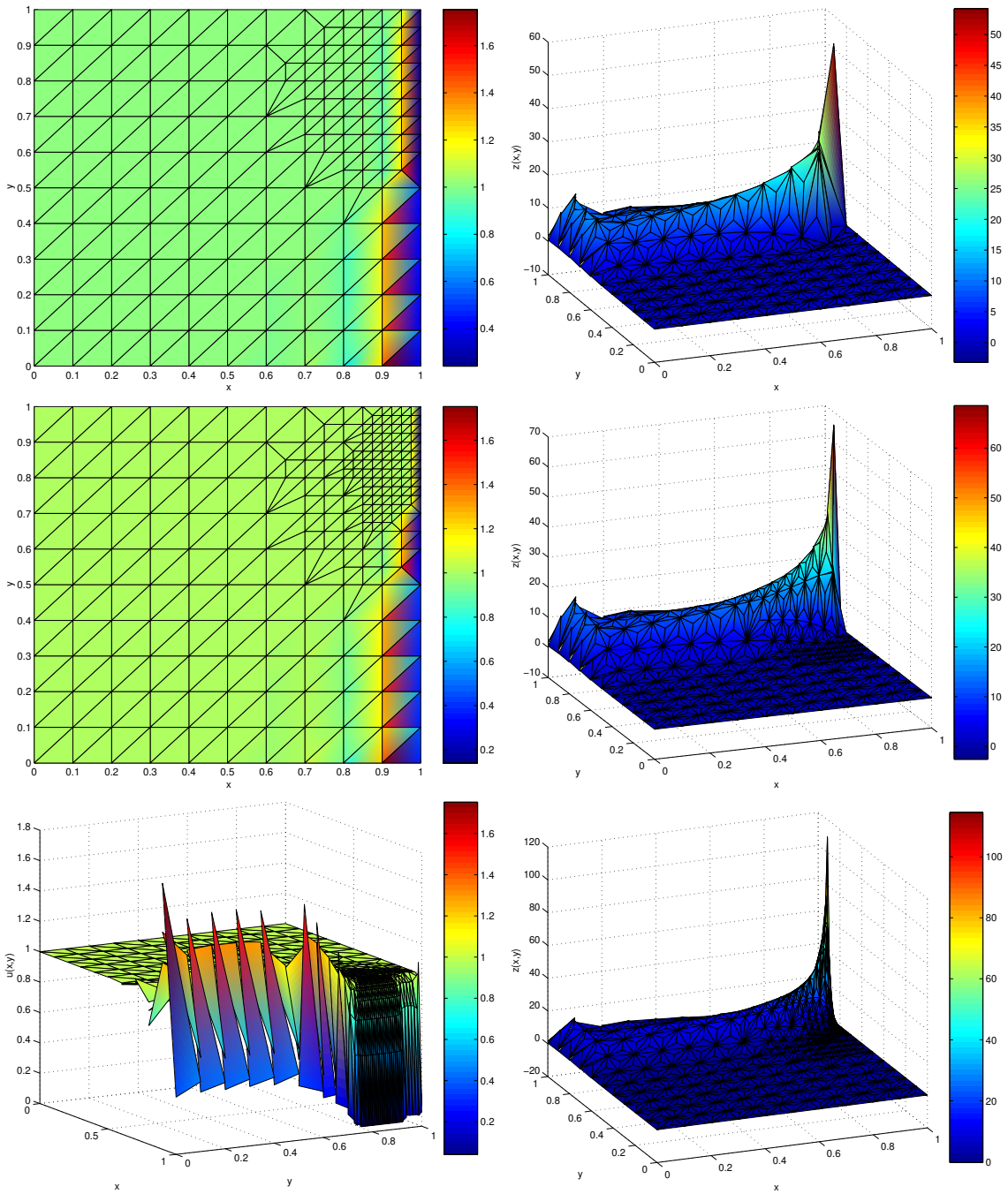



FIG. 7.1 – A gauche : solutions u(x,y) au Problème 3 sur le maillage calculé par l'intermédiaire du problème dual (à droite) et ce après 1,2 et 4 boucles

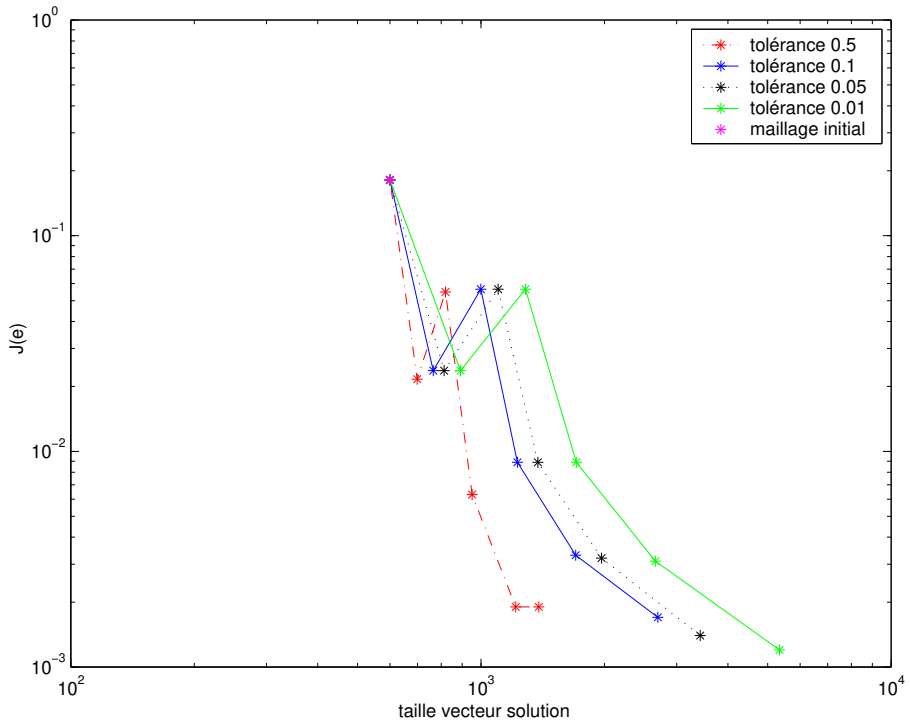



FIG. 7.2 – Convergence de J(e) en fonction de la taille du vecteur solution de $(P)$

Une question qui vient également à l'esprit est de savoir si le système va toujours raffiner aux alentours du centre du disque sur lequel on calcule $J(u)$.

L'exemple qui suit (Figure 7.3) montre qu'on peut également raffiner en $\Gamma_-$ si les conditions de Dirichlet y sont discontinues. En effet, si l'on refait les calculs précédents sur notre Problème 2, on s'aperçoit que le système raffine autour de (0,0.5),là où il y a discontinuité dans les conditions de Dirichlet. De plus, on raffine le long de la *caractéristique* sur laquelle se trouve notre disque d'intérêt.

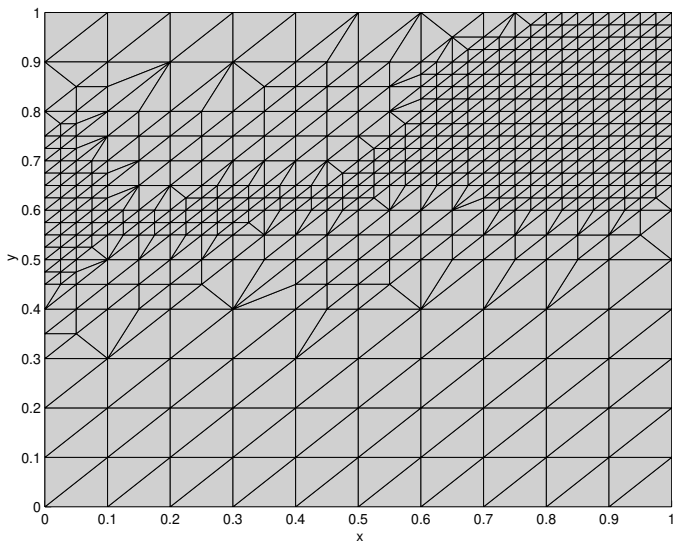



FIG. 7.3 – Maillage calculé après deux boucles, Problème 2

Mais nous pouvons observer un phénomène intéressant...Introduisons un nouveau problème :

**Problème 4**

$$-\epsilon\Delta u + \beta \cdot \nabla u = S \qquad avec\ max(x,y) \leq 1\ et\ min(x,y) \geq 0, \tag{7.21}$$

*avec $\epsilon$ non nul et $\beta = (1,0)'$.*
*De plus, les conditions de Neumann sont nulles en (1,y) de même que les conditions de Dirichlet sur $\partial\Omega \setminus \Gamma_N$.*
*Enfin, le terme source $S$ est calculé de telle sorte que*

$$u(x,y) = (\frac{1}{2} + \frac{1}{2}\tanh(20x - 10))\sin^2(\pi y). \tag{7.22}$$

On a positionné le centre du disque d'intérêt en (0.8667,0.5333) et calculé la solution après trois boucles (Figure 7.4).

On se rend compte que l'on a raffiné aux alentours du point comme auparavant mais également là où la courbure est importante et sur la caractéristique, plus particulièrement en (0.55,0.5333) et en (0.45,0.5333).

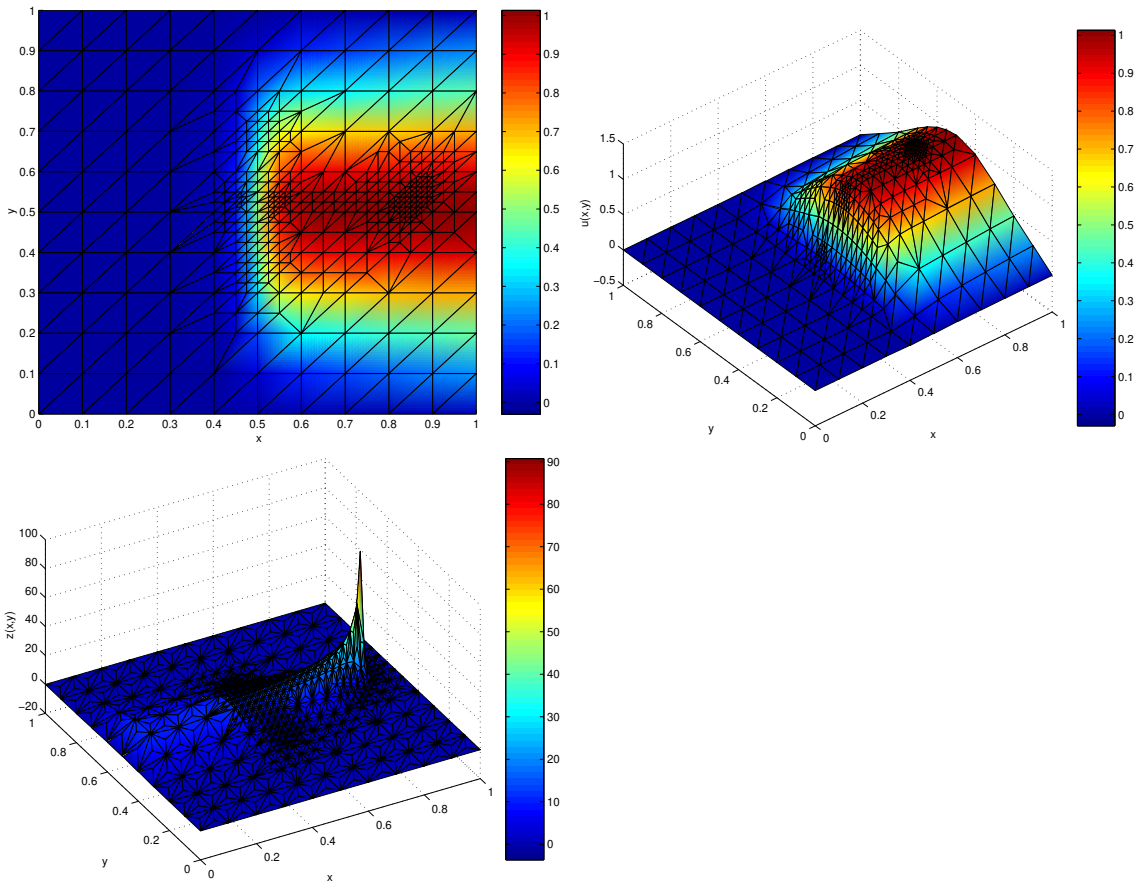



FIG. 7.4 – Solution u(x,y) du Problème 4 après 3 boucles et problème dual correspondant

# Chapitre 8

# Conclusions et perspectives

Dans ce travail, nous avons abordé les méthodes de Galerkin discontinues. On a pu se rendre compte, après implémentation de ces dernières, qu'elles approximaient mieux les couches limites que les méthodes continues qui produisaient des oscillations. Ces méthodes discontinues ont recours à ce qu'on appelle un *crime variationnel* car l'espace de discrétisation n'est pas inclus dans l'espace dans lequel on cherche la solution du problème faible.

Nous avons également déterminé la valeur optimale du paramètre de stabilisation qui permet à la solution numérique d'approximer au mieux, selon une certaine norme, la solution analytique et ce, pour un problème type.

On a ensuite étudié le comportement des valeurs propres des matrices de rigidité pour conclure que les méthodes explicites de résolution du problème d'évolution ne sont pas les plus appropriées lorsqu'elles sont couplées aux méthodes de Galerkin discontinues.

Dans la deuxième partie consacrée aux estimations d'erreur, nous avons tout d'abord étudié la convergence de la norme *énergétique*. Le calcul expérimental nous indique que la méthode est, pour un problème *quasi-hyperbolique*, d'ordre 1/2, ce qui rejoint la théorie. Nous avons également étudié la convergence d'autres normes classiques.

Ensuite, nous nous sommes attachés à mettre au point une méthode permettant le raffinement du maillage afin de diminuer l'erreur commise au sens de la norme énergétique.

Enfin, nous avons abordé les méthodes adjointes qui permettent, et c'est un progrès par rapport au point précédent, de raffiner le maillage dans le but de calculer à une tolérance près une fonctionnelle linéaire (ici, l'intégrale de la solution sur un petit disque). L'implémentation du problème dual nous permet bien de voir comment la solution de ce dernier détermine la zone d'influence de ladite fontionnelle.

Une amélioration possible serait de voir comment toutes ces méthodes se comportent si, en lieu et place de se donner un coefficient de diffusion $\epsilon$ et une vitesse $\beta$ constants, on choisit de travailler avec une matrice de diffusion et une vitesse non-constants.

On peut également s'interroger sur la manière d'implémenter ces méthodes dans le cas d'équations non-linéaires comme, par exemple, les équations de Navier-Stokes.

Enfin, on pourrait également étudier une méthode efficace de résolution des systèmes matriciels obtenus lorsqu'on discrétise les formulations faibles.