# MTH8207 – Mathématiques des éléments finis

**Serge Prudhomme**

Professeur
Département de mathématiques
Polytechnique Montréal

Cours 6

# Sommaire du cours #6

- Erreur dans les approximations EF.
- Existence et unicité des solutions de problèmes aux conditions aux limites :
  - Théorème de Lax-Milgram.
  - Exemples.
- Existence et unicité des approximations EF.
- Estimation a priori des erreurs d'approximation
- Erreurs d'interpolation
- Estimation du taux de convergence
- Estimation des erreurs a posteriori

# Erreur dans les approximations EF

Problème faible :

$$\boxed{\text{Trouver } u \in V \text{ telle que} \quad B(u,v) = F(v), \quad \forall v \in V}$$

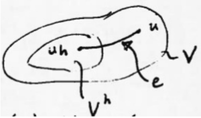
Problème éléments finis : Soit $V^h \subset V$ un espace EF conforme.

$$\boxed{\text{Trouver } u_h \in V^h \text{ telle que} \quad B(u_h,v_h) = F(v_h), \quad \forall v_h \in V^h}$$

Erreur d'approximation : $\boxed{e = u - u_h \in V}$

On pose $u = u_h + e$, soit $B(u,v) = B(u_h + e, v) = B(e,v) + B(u_h,v) = F(v)$.
D'où le problème de l'erreur :

$$\boxed{\text{Trouver } e \in V \text{ telle que} \quad B(e,v) = \mathcal{R}_h(v), \quad \forall v \in V}$$

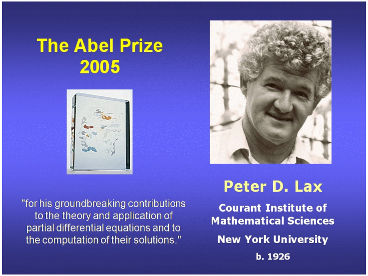



où $\mathcal{R}_h(v) = F(v) - B(u_h,v)$ est appelé le résiduel.

Propriété d'orthogonalité: $\mathcal{R}_h(v_h) = B(e,v_h) = 0, \forall v_h \in V^h$.

# Théorème de Lax-Milgram (1954)

Peter Lax (1926-, né en Hongrie)
Arthur Milgram (1912-1961)




Le théorème donne des conditions suffisantes pour déterminer l'existence et l'unicité de la solution $u \in V$ du problème $B(u, v) = F(v)$, $\forall v \in V$, lorsque :

$$\begin{aligned} V &= \text{espace de Hilbert,} \\ B &= \text{forme bilinéaire,} \\ F &= \text{forme linéaire.} \end{aligned}$$

# Théorème de Lax-Milgram

Théorème :

Soit $V$ un espace de Hilbert muni de la norme $\|\cdot\|_V$.
Si

1) $\exists M > 0$ t.q. $|B(u,v)| \leq M\|u\|_V\|v\|_V, \quad \forall u,v \in V$ (continuité de $B$)
2) $\exists C > 0$ t.q. $|F(v)| \leq C\|v\|_V \quad \forall v \in V$ (continuité de $F$)
3) $\exists \alpha > 0$ t.q. $B(u,u) \geq \alpha\|u\|_V^2 \quad \forall u \in V$ (coercivité de $B$)

alors il existe une unique solution $u \in V$ telle que $B(u,v) = F(v), \forall v \in V$.

De plus,

$$\alpha\|u\|_V^2 \leq B(u,u) = F(u) \leq C\|u\|_V \quad \Rightarrow \quad \boxed{\|u\|_V \leq C/\alpha}$$

ce qui permet de conclure que la solution $u$ dépend continuellement des données du problème.

On dit alors que le problème est bien posé.

# Exemple : $V$ = Espace vectoriel de dimension finie

Soit une matrice $A \in \mathbb{R}^{n \times n}$ de norme induite

$$\|A\| = \sup_{v \in V,\ v \neq 0} \frac{\|Av\|_V}{\|v\|_V}$$

On pose $B(u, v) = v^T Au$.

Continuité de $A$ : $|B(u, v)| \leq \|v\|_V \|Au\|_V \leq \|A\| \|u\|_V \|v\|_V$, donc $M = \|A\|$.

Coercivité de $A$ : On considère le cas où $A$ est diagonalisable et les valeurs propres sont toutes réelles. Soit $(\lambda_i, u_i)$ une valeur propre et un vecteur propre associé. Alors

$$\frac{B(u_i, u_i)}{\|u_i\|_V^2} = \frac{u_i^T (Au_i)}{\|u_i\|_V^2} = \frac{u_i^T (\lambda_i u_i)}{\|u_i\|_V^2} = \lambda_i \frac{u_i^T u_i}{\|u_i\|_V^2} = \lambda_i$$

On en conclut que $B$ est coercive si $\lambda_i > 0$, $i = 1, \ldots, n$ et donc $\alpha = \min_i \lambda_i$. Cela garantit que la matrice $A$ est inversible puisque $\det A = \prod \lambda_i \neq 0$.

# Exemple : Problème de conditions aux limites

Problème fort : Trouver $u \in C^2(0,1)$ telle que :

$$\boxed{\begin{aligned} -au'' + bu' + cu = f, \quad \forall x \in \Omega = (0,1) \\ u(0) = 0 \\ au'(1) + \gamma u(1) = g \end{aligned}}$$

où $f \in L^2(\Omega)$, $a$, $c$ sont des fonctions continues positives sur $\overline{\Omega}$, et $b$, $\gamma$, $g \in \mathbb{R}$.

Formulation faible : Soit $V = \{v \in H^1(\Omega);\ v(0) = 0\}$.

$$\boxed{\text{Trouver } u \in V \text{ telle que} \quad B(u,v) = F(v), \quad \forall v \in V}$$

où

$$B(u,v) = \int_0^1 au'v' + bu'v + cuv\ dx + \gamma u(1)v(1)$$

$$F(v) = \int_0^1 fv\ dx + gv(1)$$

# Exemple : Problème de conditions aux limites

De la continuité de $a$ et $c$, on a :

$$0 < a_{\min} \leq a(x) \leq a_{\max}$$
$$0 < c_{\min} \leq c(x) \leq c_{\max}$$

L'espace vectoriel $V$ est un sous-espace fermé de l'espace de Hilbert $H^1(\Omega)$ et de ce fait est aussi un espace de Hilbert, muni de la norme :

$$\|v\|_V = \sqrt{\int_0^1 v^2 + v'^2 dx} = \sqrt{\|v\|_{L^2}^2 + \|v'\|_{L^2}^2}$$

De plus :

$$\|v\|_{L^2} \leq \|v\|_{H^1}$$
$$\|v'\|_{L^2} \leq \|v\|_{H^1}$$

On rappelle aussi l'inégalité de Poincaré :

$$\|v\|_{L^2} \leq C_p \|v'\|_{L^2}, \quad \forall v \in V$$

# Exemple : Problème de conditions aux limites

$B(\cdot,\cdot)$ définie sur $V \times V$ est une forme bilinéaire et $F(\cdot)$ définie sur $V$ est une forme linéaire.

Il reste donc à montrer :

- la continuité de B;
- la continuité de F;
- la coercivité de B.

Remarque :
De la continuité et de la coercivité de la forme bilinéaire $B(\cdot,\cdot)$ on déduit que :

$$\alpha\|u\|_V^2 \leq B(u,u) \leq M\|u\|_V^2$$

et donc nécessairement

$$\boxed{\alpha \leq M}$$

# Existence et unicité des approximations EF

- $V^h \subset V$ est un sous-espace de dimension finie de $V$. Donc $V$ est un espace de Hilbert de norme $\|\cdot\|_V$.
- $B$ est continue et coercive sur $V^h$ puisque $V^h \subset V$. On a donc :

$$
\begin{aligned}
|B(u_h, v_h)| &\leq M \|u_h\|_V \|v_h\|_V, \quad \forall u_h, v_h \in V^h \\
B(u_h, u_h) &\geq \alpha \|u_h\|_V^2, \quad \forall u_h \in V^h
\end{aligned}
$$

  En fait, on pourrait trouver $M_h \leq M$ et $\alpha_h \geq \alpha$ telles que :

$$
\begin{aligned}
|B(u_h, v_h)| &\leq M_h \|u_h\|_V \|v_h\|_V, \quad \forall u_h, v_h \in V^h \\
B(u_h, u_h) &\geq \alpha_h \|u_h\|_V^2, \quad \forall u_h \in V^h
\end{aligned}
$$

- $F$ est continue sur $V^h$ : $F(v_h) \leq C \|v_h\|_V, \quad \forall v_h \in V^h$

On en conclut que le théorème de Lax-Milgram s'applique et $\exists ! u_h \in V^h$ telle que $\|u_h\|_V \leq C/\alpha_h$.

# Généralisation du théorème de Lax-Milgram

Le théorème de Lax-Milgram ne s'applique pas si :

- $U \neq V$ (la notion de coercivité n'est pas applicable dans ce cas).
- $U$ et $V$ ne sont pas des espaces de Hilbert.
- $U = V$ mais la coercivité n'est pas vérifiée (e.g. problème de Stokes).

On a cependant le théorème de Lax-Milgram généralisé (1971) :

- $U$ et $V$ sont des espaces de Banach.
- La coercivité est alors remplacée par la condition inf-sup, appelée aussi condition LBB (Ladyzenskaya-Babuška-Brezzi).

# Erreur dans les approximations EF

Problème faible :

$$\boxed{\text{Trouver } u \in V \text{ telle que } \quad B(u, v) = F(v), \quad \forall v \in V}$$

Problème éléments finis : Soit $V^h \subset V$ un espace EF conforme.

$$\boxed{\text{Trouver } u_h \in V^h \text{ telle que } \quad B(u_h, v_h) = F(v_h), \quad \forall v_h \in V^h}$$

Erreur d'approximation : $e = u - u_h \in V$

$$\boxed{\text{Trouver } e \in V \text{ telle que } \quad B(e, v) = \mathcal{R}_h(v), \quad \forall v \in V}$$

où

$$\mathcal{R}_h(v) = F(v) - B(u_h, v) = \text{le résiduel}$$

Propriété d'orthogonalité: $\mathcal{R}_h(v_h) = B(e, v_h) = 0, \quad \forall v_h \in V^h$.

# Théorème de Lax-Milgram

Théorème :

Soit $V$ un espace de Hilbert muni de la norme $\|\cdot\|_V$.
Si

1) $\exists M > 0$ t.q. $|B(u,v)| \leq M\|u\|_V\|v\|_V, \quad \forall u,v \in V$ (continuité de $B$)
2) $\exists C > 0$ t.q. $|F(v)| \leq C\|v\|_V \quad \forall v \in V$ (continuité de $F$)
3) $\exists \alpha > 0$ t.q. $B(u,u) \geq \alpha\|u\|_V^2 \quad \forall u \in V$ (coercivité de $B$)

alors il existe une unique solution $u \in V$ telle que $B(u,v) = F(v), \forall v \in V$.

De plus,

$$\alpha\|u\|_V^2 \leq B(u,u) = F(u) \leq C\|u\|_V \quad \Rightarrow \quad \boxed{\|u\|_V \leq C/\alpha}$$

ce qui permet de conclure que la solution $u$ dépend continuellement des données du problème.

On dit alors que le problème est bien posé.

# Estimation a priori des erreurs d'approximation

Par la coercivité et la continuité de $B$, ainsi que la propriété d'orthogonalité, on a avec $w_h \in V^h$ arbitraire :

$$\begin{aligned}\alpha \|e\|_V^2 &\leq B(e,e) = B(e, u-u_h) = B(e, u-w_h+w_h-u_h) \\ &= B(e, u-w_h) + B(e, w_h-u_h) = B(e, u-w_h) \\ &\leq M\|e\|_V \|u-w_h\|_V\end{aligned}$$

Donc, avec $\alpha > 0$ :

$$\boxed{\|e\|_V \leq \frac{M}{\alpha}\|u-w_h\|_V}, \quad \forall w_h \in V^h$$

Puisque l'inégalité est valable pour tout $w_h \in V^h$, on peut écrire :

$$\boxed{\|e\|_V \leq \frac{M}{\alpha} \inf_{w_h \in V^h} \|u-w_h\|_V}$$

# Estimation a priori des erreurs d'approximation

**Remarques :**

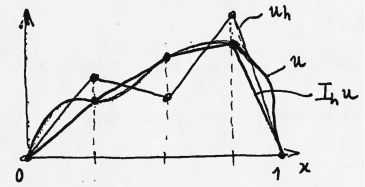
- $M/\alpha \geq 1$.
- Si $B$ est symétrique, on peut améliorer le résultat :

$$\|e\|_V \leq \sqrt{\frac{M}{\alpha}}\|u - w_h\|_V, \quad \forall w_h \in V^h$$

**Interpolant :**

On choisit $w_h = I_h^p u$, où $I_h^p u$ est l'interpolant de $u$ dans $V^h$. Soit $I_h^p u(x) = \sum_{i=1}^N \sigma_i(u)\phi_i(x)$, où les $\sigma_i$ sont les degrés de liberté et $p$ est le degré des polynômes de l'élément fini considéré. On a alors :

$$\|e\|_V \leq \frac{M}{\alpha}\|u - I_h^p u\|_V$$



# Estimation des erreurs d'interpolation

Erreurs d'interpolation dans le cas où $V = H^1$.

Si $u \in H^{r+1}$, $r \geq 1$, alors on peut montrer que :

$$\|u - I_h^p u\|_{L^2} \leq C_I h^{\min(p,r)+1} \|u\|_{H^{r+1}}$$
$$\|(u - I_h^p u)'\|_{L^2} \leq C_I h^{\min(p,r)} \|u\|_{H^{r+1}}$$

On en déduit que

$$\|e\|_{H^1} \leq \frac{M}{\alpha} C_I h^{\min(p,r)} \|u\|_{H^{r+1}}$$

En introduisant $C_u = \frac{M}{\alpha} C_I \|u\|_{H^{r+1}}$, une constante indépendante de $h$, on a

$$\boxed{\|e\|_{H^1} \leq C_u h^{\min(p,r)}}$$

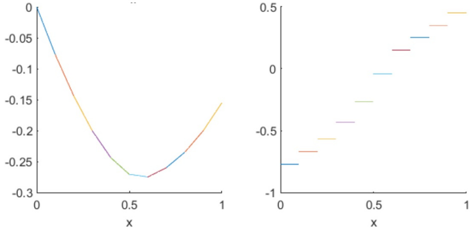
$\beta = \min(p, r)$ est appelé le taux de convergence ou l'ordre de la méthode.

# Estimation des erreurs d'interpolation

Remarque : On peut aussi trouver une estimation de l'erreur dans la norme $L^2$ en utilisant le résultat d'Aubin-Nitsche (on passe les détails), i.e. il existe une constante $C > 0$ indépendante de $u$ et $h$, telle que :

$$\boxed{\|e\|_{L^2} \leq Ch^{\min(p,r)+1}\|u\|_{H^{r+1}}}$$

L'erreur dans la dérivée de la solution $u$ est dominante par rapport à l'erreur dans la solution $u$. Il est donc important, lorsqu'on analyse la solution EF, de regarder les dérivées de $u$ plutôt que $u$ elle-même.



# Exemples

- Supposons que $u \in H^s$ avec $s$ très grand. On a $r = s - 1$ très grand et donc $\beta = \min(p, r) = p$.
  Le taux de convergence $\beta$ augmente alors lorsque l'on augmente le degré $p$ des polynômes pour approcher $u$.
- Supposons que $u \in H^2$. Dans ce cas, $r = s - 1 = 2 - 1 = 1$ et $\beta = \min(p, r) = 1, \forall p = 1, 2,$ etc.
  Il n'y a pas d'amélioration du taux de convergence car celui dépend aussi de la régularité $s$ de la solution du problème.

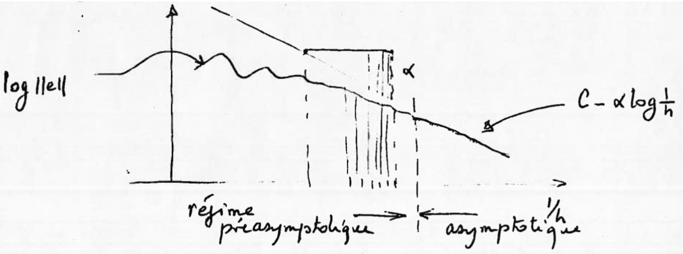
# Estimation du taux de convergence

On suppose que l'erreur se comporte selon la relation :

$$\|e\| \approx Ch^{\beta}$$

où $\|e\|$ est la norme $L^2$, $H^1$, ou tout autre norme.

$$\log \|e\| \approx C + \beta \log h \approx C - \beta \log \frac{1}{h}$$



# En pratique

On considère un problème pour lequel la solution exacte est connue de manière à pouvoir calculer l'erreur $e = u - u_h$.

Soient deux maillages uniformes de tailles $h_1$ et $h_2$, telles que $h_1 > h_2$ (prendre e.g. $h_1 = 2h_2$).

- Calculer :

$$\begin{aligned} \log\|e_1\| &\approx C_h + \beta_h \log h_1 \\ \log\|e_2\| &\approx C_h + \beta_h \log h_2 \end{aligned}$$

- Évaluer $\beta_h$ :

$$\boxed{\beta_h = \frac{\log\|e_1\| - \log\|e_2\|}{\log h_1 - \log h_2} = \frac{\log(\|e_1\|/\|e_2\|)}{\log(h_1/h_2)}}$$

L'idée est alors d'itérer et d'estimer $\beta$ en prenant la limite de $\beta_h$ lorsque $h$ tend vers zéro.

# En pratique

On considère une séquence de maillages uniformes d'éléments de taille $h_k$, $k = 1, 2 \ldots, n$, telle que $h_k > h_{k+1}$ (prendre e.g. $h_k = 2h_{k+1}$).

Pour $k = 1, 2, \ldots, n-1$, évaluer $\beta_k$ :

$$\boxed{\beta_k = \frac{\log(\|e_k\|/\|e_{k+1}\|)}{\log(h_k/h_{k+1})} = \frac{\log(\|e_k\|/\|e_{k+1}\|)}{\log 2}}$$

Tabuler les résultats (représentation n'est pas unique), par exemple :

| $k$ | $h_k$ | $\|e_k\|$ | $\beta_k$ |
|---|---|---|---|
| 1 | 1/2 | 0.10541 | 1.464 |
| 2 | 1/4 | 0.03820 | 1.765 |
| 3 | 1/8 | 0.01123 | 1.926 |
| 4 | 1/16 | 0.00295 | – |

ou

| $h$ | $\|u - u_h\|$ | $\beta$ |
|---|---|---|
| 1/2 | 0.10541 | – |
| 1/4 | 0.03820 | 1.464 |
| 1/8 | 0.01123 | 1.765 |
| 1/16 | 0.00295 | 1.926 |

Si $\beta_k$ converge lorsque $h_k$ tend vers zéro, i.e. $\beta_k \to \beta$, alors $\beta$ est le taux de convergence (= ordre de la méthode).

# Remarques

- Parfois, on préfère représenter les erreurs $\|e\|$ en fonction du nombre total $N$ de degrés de liberté, i.e. afficher les courbes $\log\|e\|$ versus $\log N$. En 1D, $N$ est de l'ordre de $p/h$.
- Dans le cas où la solution exacte n'est pas disponible, on peut en obtenir une approximation précise sur un maillage très fin ("overkill solution" en anglais) et l'utiliser comme solution de référence.
- Le calcul du taux de convergence (ordre de la méthode) est utile pour la comparaison de méthodes ou la vérification des codes de calcul.
- L'estimation a priori des erreurs permet d'estimer les taux de convergence mais ne peut pas être utilisée pour quantifier les erreurs (puisque les estimateurs dépendent de la solution exacte que l'on ne connait pas a priori).

# Estimation des erreurs a posteriori

On veut quantifier l'erreur $e = u - u_h$ où :

$$\text{Trouver } u \in V \text{ telle que} \quad B(u, v) = F(v), \quad \forall v \in V$$

$$\text{Trouver } u_h \in V^h \text{ telle que} \quad B(u_h, v_h) = F(v_h), \quad \forall v_h \in V^h$$

Deux approches fondamentales :

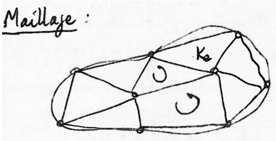
1) Estimation des erreurs en normes globales ($H^1$, $L^2$, ou normes dites d'énergie).
Premiers travaux dès la fin des années 1970.

2) Estimation par rapport à des grandeurs physiques (que l'on appelle aussi quantités d'intérêt), par exemple, la solution en un point donné $x_0 \in \Omega$, la moyenne locale de la solution dans un sous-domaine $\omega \subset \Omega$, le flux à travers certaines parties de la frontière, etc.
Premiers travaux dès le milieu des années 1990.

# Sources d'erreur

1) Erreurs de discrétisation du domaine
2) Erreurs de quadrature

$$
\begin{aligned}
B(u_h, v_h) &= \int_\Omega u_h v_h dx \\
&\approx \int_{\Omega_h} u_h v_h dx = \sum_K \int_K u_h v_h dx \\
&\approx \sum_K \sum_g \omega(\xi_g) u_h(\xi_g) v_h(\xi_g) |J(\xi_g)|
\end{aligned}
$$




$\Omega_h = \cup_e K_e \neq \Omega$

3) Erreurs dans la solution du système $KU = F$ (e.g. erreurs d'arrondi si solveur direct ou erreurs de convergence si solveur itératif)
4) Approximation de la solution $u$ par des fonctions $u_h$ continues et polynomiales par morceaux

$\Rightarrow$ On ignore ici les sources d'erreur #1, #2, et #3 et se concentre sur #4!

# Relation entre l'erreur et le résiduel

$$\boxed{\text{Trouver } e \in V \text{ telle que} \quad B(e,v) = \mathcal{R}_h(v) \equiv F(v) - B(u_h, v), \quad \forall v \in V}$$

The résiduel $\mathcal{R}$ est une forme linéaire (fonctionnelle) de $V$ dans $\mathbb{R}$.

$$\|\mathcal{R}\| = \sup_{v \neq 0 \in V} \frac{|\mathcal{R}(v)|}{\|v\|_V} \quad \Rightarrow \quad \frac{|\mathcal{R}(e)|}{\|e\|_V} \leq \|\mathcal{R}\| \quad \Rightarrow \quad \mathcal{R}(e) \leq \|\mathcal{R}\| \|e\|_V$$

Coercivité de B :

$$\alpha \|e\|_V^2 \leq B(e,e) = \mathcal{R}(e) \leq \|\mathcal{R}\| \|e\|_V \quad \Rightarrow \quad \boxed{\alpha \|e\|_V \leq \|\mathcal{R}\|}$$

Continuité de B :

$$\|\mathcal{R}\| = \sup_{v \neq 0 \in V} \frac{|\mathcal{R}(v)|}{\|v\|_V} = \sup_{v \neq 0 \in V} \frac{|B(e,v)|}{\|v\|_V} \leq \sup_{v \neq 0 \in V} \frac{M \|e\|_V \|v\|_V}{\|v\|_V} = M \|e\|_V$$

On a donc :

$$\boxed{\alpha \|e\|_V \leq \|\mathcal{R}\| \leq M \|e\|_V}$$

# Méthodes d'estimation d'erreur a posteriori

$$\boxed{\alpha\|e\|_V \leq \|\mathcal{R}\| \leq M\|e\|_V}$$

Remarques :

- $\|\mathcal{R}\|$ est une bonne mesure de l'erreur $\|e\|_V$ si $\alpha \approx 1$ et $M \approx 1$. En particulier, $\|\mathcal{R}\| = \|e\|_V$ si $\alpha = M = 1$.
- Puisque $\mathcal{R}$ est le terme source dans l'équation d'erreur, c'est le terme à contrôler dans une stratégie d'adaptation de maillage.

L'objectif est alors de trouver une norme $\|\cdot\|_V$ (norme énergie), de sorte que $\alpha$ et $M$ soient proches de 1, et de développer des méthodes efficaces pour estimer la norme $\|\mathcal{R}\|$ :

- Méthodes explicites : aucun problème supplémentaire à résoudre.
- Méthodes implicites : résolution de problèmes auxiliaires sur chaque élément ou sur des sous-domaines de $\Omega$.

# Résumé du cours

- Un outil essentiel pour montrer l'existence et l'unicité de solutions de problèmes de conditions aux limites est le théorème de Lax-Milgram.
- Le théorème est utilisé pour obtenir des estimations a priori et a posteriori des erreurs $e = u - u_h$ dans les approximations EF.
- Les estimateurs d'erreur a priori reposent sur les estimateurs des erreurs d'interpolation entre la solution exacte $u$ et son interpolant dans $V^h \subset V$.
- L'estimation a priori des erreurs permet d'estimer les taux de convergence mais n'est pas utile pour quantifier les erreurs (puisque les estimateurs dépendent de la solution exacte que l'on ne connait pas a priori). On ne peut donc pas s'en servir pour le développement d'une stratégie d'adaptation de maillage.
- Les méthodes d'estimation a posteriori permettent de quantifier les erreurs en estimant par des méthodes efficaces la norme $\|\mathcal{R}\|$ du résiduel, puisque celle-ci est équivalente à la norme de l'erreur $\|e\|_V$.