Introduction aux méthodes de réduction de modèle pour les EDP

Marie Billaud-Friess

Centrale Méditerranée, Aix-Marseille Université
Institut de Mathématiques de Marseille

Master Class - M2 ANADEAL

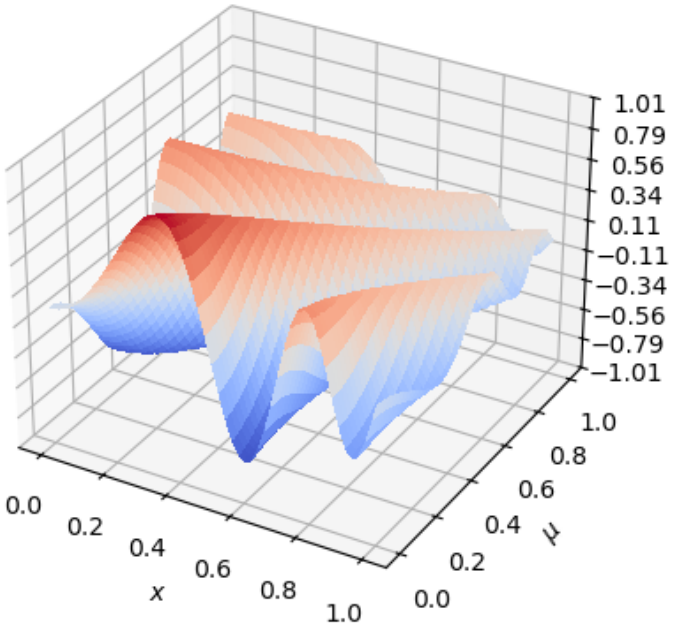
20 Juin 2025

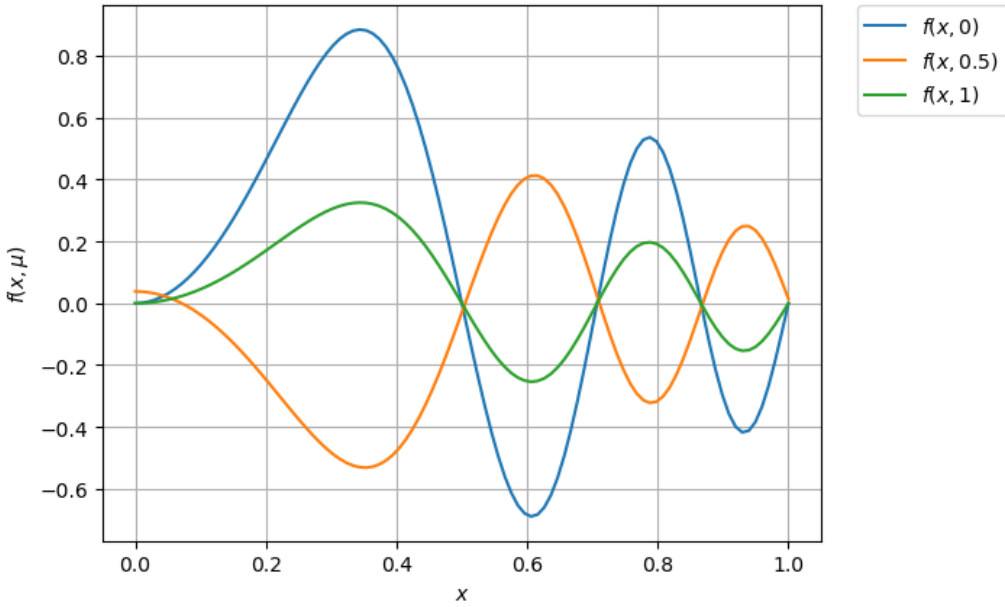
Plan

Exemple jouet ① : fonction à variables séparées

Considérons

$$f(x,\mu) = \exp(-x^2 - \mu)\sin(4\pi(x^2 - \mu^2)), (x,\mu) \in [0,1]^2$$

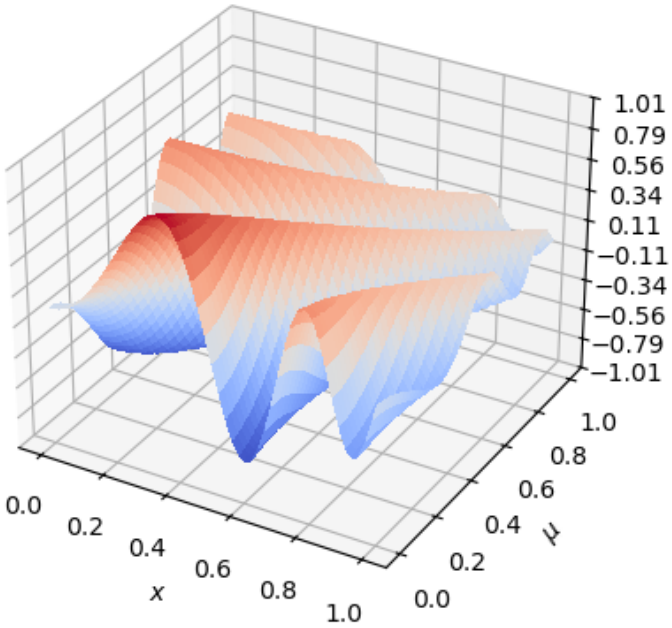


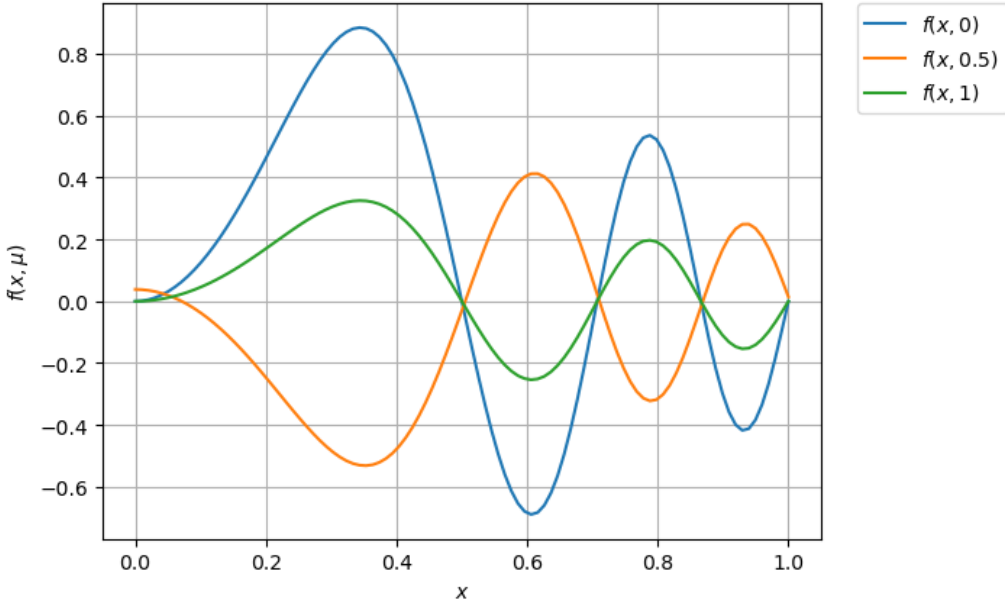


Exemple jouet ① : fonction à variables séparées

Considérons

$$f(x,\mu) = \exp(-x^2 - \mu)\sin(4\pi(x^2 - \mu^2)), (x,\mu) \in [0,1]^2$$





Question. Peut-on écrire la fonction "compliquée" f comme une somme de produits de fonctions "simples" dépendant de x et μ séparément ?

Exemple jouet ① : fonction à variables séparées

Réponse. Oui, après développement :

$$f(x,\mu) = \underbrace{\exp(-x^2)\sin(4\pi x^2)}_{\varphi_1(x)}\underbrace{\exp(-\mu)\cos(\mu^2)}_{\alpha_1(\mu)} + \underbrace{(-\exp(-x^2)\cos(4\pi x^2))}_{\varphi_2(x)}\underbrace{\exp(-\mu)\sin(\mu^2)}_{\alpha_2(\mu)},$$

c'est à dire

$$f(x,\mu) = \sum_{i=1}^{2} \alpha_i(\mu)\varphi_i(x).$$

La fonction f est dite de rang 2.

Exemple jouet ① : fonction à variables séparées

Réponse. Oui, après développement :

$$f(x,\mu) = \underbrace{\exp(-x^2)\sin(4\pi x^2)}_{\varphi_1(x)}\underbrace{\exp(-\mu)\cos(\mu^2)}_{\alpha_1(\mu)} + \underbrace{(-\exp(-x^2)\cos(4\pi x^2))}_{\varphi_2(x)}\underbrace{\exp(-\mu)\sin(\mu^2)}_{\alpha_2(\mu)},$$

c'est à dire

$$f(x,\mu) = \sum_{i=1}^{2} \alpha_i(\mu)\varphi_i(x).$$

La fonction f est dite de rang 2.

En pratique. Soient $\{x_i\}_{i=1}^n, \{\mu_i\}_{i=1}^m \subset [0,1]$. Il s'agit de calculer

- ✓ seulement $2(m+n)$ évaluations de α_1, α_2 en $\{\mu_i\}_{i=1}^m$ et pour φ_1, φ_2 en $\{x_i\}_{i=1}^n$,
- ✗ contre nm évaluations de f en $\{(x_i,\mu_j)\}_{1\le i\le n, 1\le j\le m}$.

Exemple jouet ② : fonction approchable par une fonction à variables séparées

Approximation. Considérons

$$g(x,\mu) = \sum_{i\geq 1} 10^{-i} \exp(-x^2-\mu)\sin(2^i\pi(x^2-\mu^2)), (x,\mu)\in[0,1]^2$$

Exemple jouet ② : fonction approchable par une fonction à variables séparées

Approximation. Considérons

$$g(x,\mu) = \sum_{i\geq 1} 10^{-i} \exp(-x^2-\mu)\sin(2^i\pi(x^2-\mu^2)), (x,\mu)\in[0,1]^2$$

approchée par une fonction de rang $2k$

$$g_{2k}(x,\mu) = \sum_{i=1}^{k} 10^{-i} \exp(-x^2-\mu)\sin(2^i\pi(x^2-\mu^2)), (x,\mu)\in[0,1]^2$$

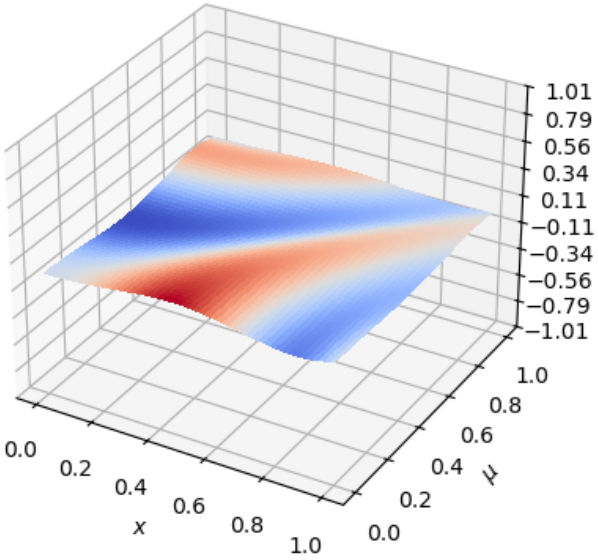
Exemple jouet ② : fonction approchable par une fonction à variables séparées

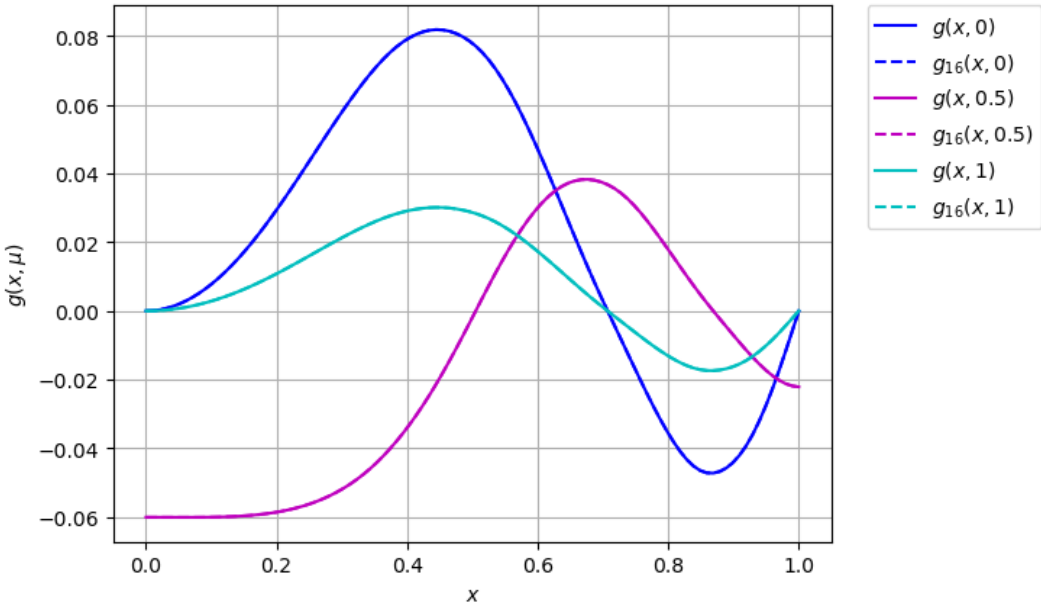
Approximation. Considérons

$$g(x,\mu) = \sum_{i\geq 1} 10^{-i} \exp(-x^2-\mu)\sin(2^i\pi(x^2-\mu^2)), (x,\mu)\in[0,1]^2$$

approchée par une fonction de rang $2k$

$$g_{2k}(x,\mu) = \sum_{i=1}^{k} 10^{-i} \exp(-x^2-\mu)\sin(2^i\pi(x^2-\mu^2)), (x,\mu)\in[0,1]^2$$





$k=1$	rang $=2$	$err_2 = 0.347220406696138$
$k=8$	rang $=16$	$err_2 = 3.616671153088143e-08$
$k=15$	rang $=30$	$err_2 = 3.619439553869622e-15$

Bilan

① Certaines fonctions **admettent** une représentation séparée de faible rang.

ex. : La fonction f.

② D'autres peuvent **être approchées** par des fonctions qui admettent une représentation séparée de faible rang.

ex. : La fonction g.

$\rightsquigarrow$ Une fonction sous format séparé de faible rang est "moins" coûteuse à évaluer.

But de l'exposé

Problèmes complexes issus de la physique, biologie, finance, chimie, ...

But de l'exposé

Problèmes complexes issus de la physique, biologie, finance, chimie, ...

1. formulation mathématique : équation aux dérivées partielles $\mathcal{F}(u) = 0$

Equation de la chaleur

$$\partial_t u - \triangle u = 0$$

Equation d'advection diffusion

$$-\kappa \triangle u + a \cdot \nabla u = f$$

But de l'exposé

Problèmes complexes issus de la physique, biologie, finance, chimie, ...

1. formulation mathématique : équation aux dérivées partielles $\mathcal{F}(u) = 0$

Equation de la chaleur

$$\partial_t u - \triangle u = 0$$

Equation d'advection diffusion

$$-\kappa \triangle u + a \cdot \nabla u = f$$

2. non résolvable analytiquement : résolution numérique $F_h(\boldsymbol{u}_h) = 0$

But de l'exposé

Problèmes complexes issus de la physique, biologie, finance, chimie, ...

1. formulation mathématique : équation aux dérivées partielles $\mathcal{F}(u) = 0$

Equation de la chaleur

$$\partial_t u - \triangle u = 0$$

Equation d'advection diffusion

$$-\kappa \triangle u + a \cdot \nabla u = f$$

2. non résolvable analytiquement : résolution numérique $F_h(\boldsymbol{u}_h) = 0$
3. résolution numérique trop coûteuse : réduction de modèle $F_n(\boldsymbol{u}_n) = 0$

But de l'exposé

Problèmes complexes issus de la physique, biologie, finance, chimie, ...

1. formulation mathématique : équation aux dérivées partielles $\mathcal{F}(u) = 0$

Equation de la chaleur

$$\partial_t u - \triangle u = 0$$

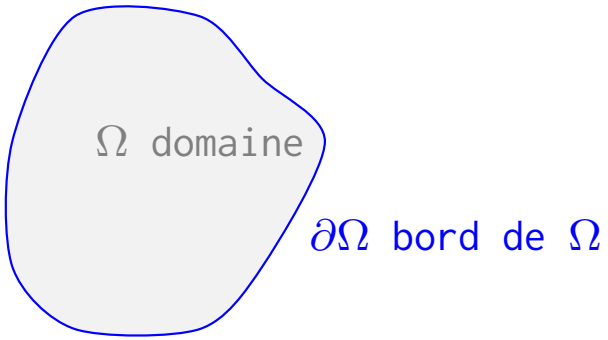
Equation d'advection diffusion

$$-\kappa \triangle u + a \cdot \nabla u = f$$

2. non résolvable analytiquement : résolution numérique $F_h(\boldsymbol{u}_h) = 0$
3. résolution numérique trop coûteuse : réduction de modèle $F_n(\boldsymbol{u}_n) = 0$

Les méthodes de réduction de modèle

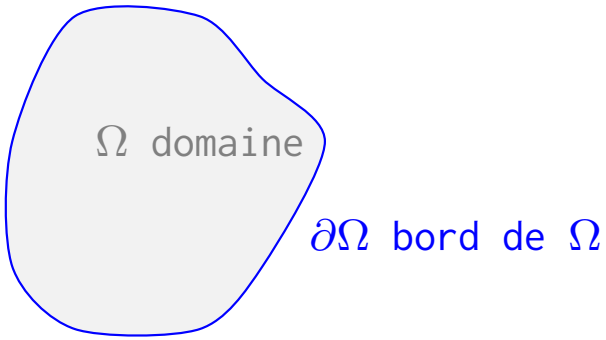
Approches qui utilisent des représentations sous **format séparé de faible rang** pour alléger les coûts de calculs lors de la **résolution numérique** d'équations aux dérivées partielles.



Problème modèle (simple).

On cherche $u : \Omega \to \mathbb{R}$ solution d'une EDP (équation aux dérivées partielles)

$$\begin{cases} \mathcal{L}u = f, \text{ dans } \Omega, \\ + \text{ conditions aux limites sur } \partial\Omega. \end{cases} \tag{1}$$



Problème modèle (simple).

On cherche $u : \Omega \to \mathbb{R}$ solution d'une EDP (équation aux dérivées partielles)

$$\begin{cases} \mathcal{L}u = f, \text{ dans } \Omega, \\ + \text{ conditions aux limites sur } \partial\Omega. \end{cases} \tag{1}$$

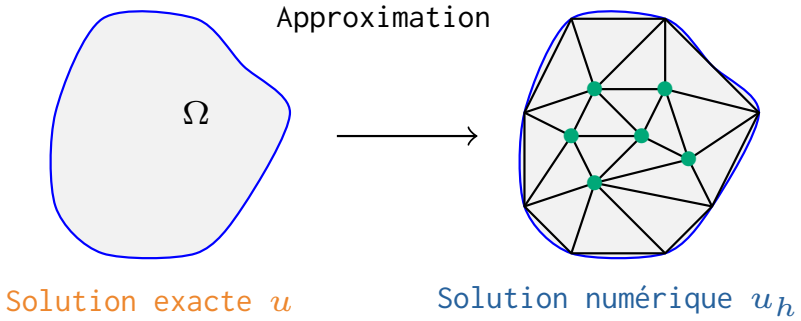
avec

- $f : \Omega \to \mathbb{R}$ est un terme source connu
- $\Omega \subset \mathbb{R}^d$ le domaine spatial
- $\mathcal{L}$ un opérateur différentiel linéaire

ex. : Equation d'advection diffusion $\mathcal{L}u = -\kappa\triangle u + a \cdot \nabla u$

Modèle numérique $F_h(\boldsymbol{u}_h) = 0$

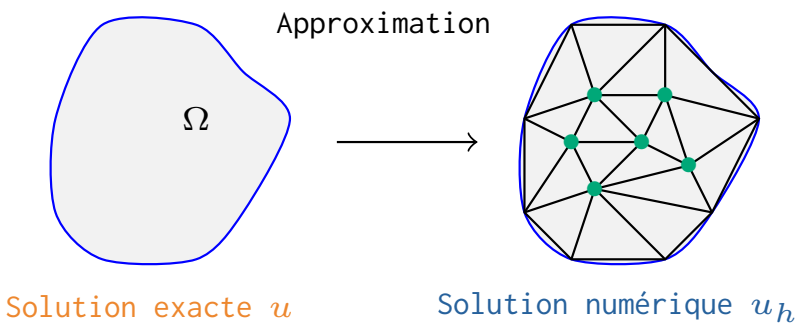
Approximation.



On veut calculer numériquement une approximation u_h de u.

ex. : Méthode éléments finis, différences finies etc.

Approximation.



On veut calculer numériquement une approximation u_h de u.

ex. : Méthode éléments finis, différences finies etc.

Système algébrique de taille N.

En pratique, pour obtenir $u_h \approx u$, il faut calculer le vecteur des coefficients

$$\boldsymbol{u}_h = (\mathsf{u}_1, \ldots, \mathsf{u}_N)^T \in \mathbb{R}^N$$

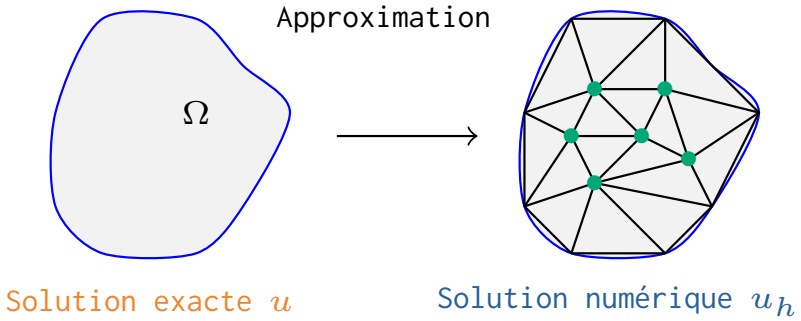
solution de

$$L_h \boldsymbol{u}_h = \boldsymbol{f}_h, \text{ dans } \mathbb{R}^N \tag{2}$$

avec $L_h \in \mathbb{R}^{N \times N}$ et $\boldsymbol{f}_h \in \mathbb{R}^N$ opérateur et terme source discrets.

ex. : Pour les éléments finis $\mathcal{L} = -\frac{d^2}{dx^2} \Rightarrow L_h = \frac{1}{h}\mathrm{tridiag}(-1, 2, -1)$

Approximation.



On veut calculer numériquement une approximation u_h de u.

ex. : Méthode éléments finis, différences finies etc.

Système algébrique de taille N.

En pratique, pour obtenir $u_h \approx u$, il faut calculer le vecteur des coefficients

$$\boldsymbol{u}_h = (\mathsf{u}_1, \ldots, \mathsf{u}_N)^T \in \mathbb{R}^N$$

solution de

$$L_h \boldsymbol{u}_h = \boldsymbol{f}_h, \text{ dans } \mathbb{R}^N \tag{2}$$

avec $L_h \in \mathbb{R}^{N \times N}$ et $\boldsymbol{f}_h \in \mathbb{R}^N$ opérateur et terme source discrets.

ex. : Pour les éléments finis $\mathcal{L} = -\frac{d^2}{dx^2} \Rightarrow L_h = \frac{1}{h}\mathrm{tridiag}(-1, 2, -1)$

Intérêt ? Ce système linéaire peut être résolu efficacement sur ordinateur.

Combien ça coûte ?

Pour mesurer le coût de calcul on parle de complexité algorithmique.

- C'est le nombre d'opérations $(+, -, \times, /)$ pour résoudre ce système linéaire.
- Elle dépend de N !

Combien ça coûte ?

Pour mesurer le coût de calcul on parle de complexité algorithmique.

- C'est le nombre d'opérations $(+, -, \times, /)$ pour résoudre ce système linéaire.
- Elle dépend de N !

Quelques algorithmes

✗ Méthode de Cramer $\mathcal{O}((N+1)!)$ Si $N = 50$ il faut $51! \approx 1.5 \cdot 10^{66}$ opérations. $\approx 5 \cdot 10^{49}$ années sur une machine équipée d'un processeur à $1G.s^{-1}$

✓ Méthode du pivot de Gauss $\mathcal{O}(N^3)$

✓ Méthode Jacobi, Gauss-Seidel $\mathcal{O}(K \cdot N^2)$

Combien ça coûte ?

Pour mesurer le coût de calcul on parle de complexité algorithmique.

- C'est le nombre d'opérations $(+, -, \times, /)$ pour résoudre ce système linéaire.
- Elle dépend de N !

Quelques algorithmes

✗ Méthode de Cramer $\mathcal{O}((N+1)!)$ Si $N = 50$ il faut $51! \approx 1.5 \cdot 10^{66}$ opérations. $\approx 5 \cdot 10^{49}$ années sur une machine équipée d'un processeur à $1G.s^{-1}$

✓ Méthode du pivot de Gauss $\mathcal{O}(N^3)$

✓ Méthode Jacobi, Gauss-Seidel $\mathcal{O}(K \cdot N^2)$

$\Rightarrow$ Pour des **modèles numériques complexes** $N >> 1$, des algorithmes (solveurs) performants et parallélisés sur des machines très puissantes sont utilisés.

Et si on rajoutait des paramètres ?

Problème paramétré. On cherche $u : \mu \mapsto u(\mu)$ avec $\mu \in \mathcal{P} \subset \mathbb{R}^p$ des paramètres.

Modèle continu
$\mathcal{F}(u(\mu), \mu) = 0$

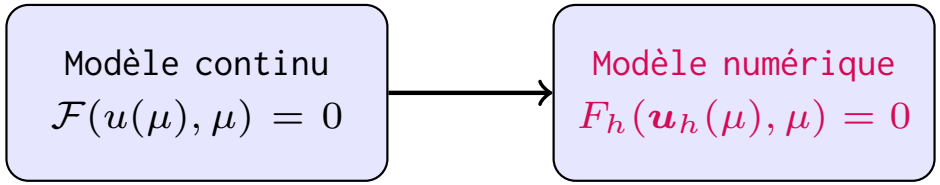
$\longrightarrow$

Modèle numérique
$F_h(\boldsymbol{u}_h(\mu), \mu) = 0$

ex. : $-\kappa(\mu)\triangle u(\mu) + a(\mu) \cdot \nabla u(\mu) = f(\mu)$

Et si on rajoutait des paramètres ?

Problème paramétré. On cherche $u : \mu \mapsto u(\mu)$ avec $\mu \in \mathcal{P} \subset \mathbb{R}^p$ des paramètres.



ex. : $-\kappa(\mu)\triangle u(\mu) + a(\mu) \cdot \nabla u(\mu) = f(\mu)$

Problèmes directs.

- Calculer efficacement $\boldsymbol{u}_h(\mu)$ pour de nombreuses valeurs de $\mu \in \mathcal{P}$.
- Modèle numérique $N >> 1$ $\rightsquigarrow$ Trop coûteux.

Peut-on remplacer le modèle numérique (N) par autre modèle dont le coût de résolution est réduit (n) ?

✗ Modèle numérique (taille N)

$$F_h(\boldsymbol{u}_h(\mu), \mu) = 0$$

ex. : Jacobi, Complexité $\mathcal{O}(N^2)$

$\xrightarrow[\boldsymbol{u}_h(\mu) \approx \boldsymbol{u}_n(\mu)]{\text{Approximation}}$

✓ Modèle réduit (taille n)

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0$$

ex. : Jacobi, Complexité $\mathcal{O}(n^2)$

1. Berkooz, G; Holmes, P; Lumley, J L (January 1993). "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows". Annual Review of Fluid Mechanics. 25 (1): 539–575.

2. A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. AIAA Journal, 18(4) :455–462, April 1980.

Peut-on remplacer le modèle numérique (N) par autre modèle dont le coût de résolution est réduit (n) ?

✗ Modèle numérique (taille N)

$$F_h(\boldsymbol{u}_h(\mu), \mu) = 0$$

ex. : Jacobi, Complexité $\mathcal{O}(N^2)$

$\xrightarrow[\boldsymbol{u}_h(\mu) \approx \boldsymbol{u}_n(\mu)]{\text{Approximation}}$

✓ Modèle réduit (taille n)

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0$$

ex. : Jacobi, Complexité $\mathcal{O}(n^2)$

Méthodes de réduction de modèle.

1. **Offline :** Effectuer des (pré)calculs pour construire le modèle réduit.
2. **Online :** On résoud le modèle réduit pour calculer des solutions.

1. Berkooz, G; Holmes, P; Lumley, J L (January 1993). "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows". Annual Review of Fluid Mechanics. 25 (1): 539–575.

2. A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. AIAA Journal, 18(4) :455–462, April 1980.

Peut-on remplacer le modèle numérique (N) par autre modèle dont le coût de résolution est réduit (n) ?

✗ Modèle numérique (taille N)

$$F_h(\boldsymbol{u}_h(\mu), \mu) = 0$$

ex. : Jacobi, Complexité $\mathcal{O}(N^2)$

$\xrightarrow[\boldsymbol{u}_h(\mu) \approx \boldsymbol{u}_n(\mu)]{\text{Approximation}}$

✓ Modèle réduit (taille n)

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0$$

ex. : Jacobi, Complexité $\mathcal{O}(n^2)$

Méthodes de réduction de modèle.

1. **Offline :** Effectuer des (pré)calculs pour construire le modèle réduit.
2. **Online :** On résoud le modèle réduit pour calculer des solutions.

$\Rightarrow$ Quelques méthodes de réduction de modèle connues :

- Proper Orthogonal Decomposition[1] (POD)
- Bases réduites[2] (BR)

1. Berkooz, G; Holmes, P; Lumley, J L (January 1993). "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows". Annual Review of Fluid Mechanics. 25 (1): 539–575.

2. A. K. Noor and J. M. Peters. Reduced basis technique for nonlinear analysis of structures. AIAA Journal, 18(4) :455–462, April 1980.

Approximation linéaire

Les méthodes de réduction de modèle classiques (POD,BR) sont des **méthodes d'approximation linéaire**.

Approximation linéaire

Les méthodes de réduction de modèle classiques (POD,BR) sont des **méthodes d'approximation linéaire**.

0n cherche une approximation $\boldsymbol{u}_n(\mu) \in V_n$ de $\boldsymbol{u}_h(\mu) \in \mathbb{R}^N$

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i \in V_n$$

avec $V_n = \mathrm{vect}\{\varphi_1, \ldots, \varphi_n\} \subset \mathbb{R}^N$ et $\dim V_n = n$.

Approximation linéaire

Les méthodes de réduction de modèle classiques (POD,BR) sont des **méthodes d'approximation linéaire**.

0n cherche une approximation $\boldsymbol{u}_n(\mu) \in V_n$ de $\boldsymbol{u}_h(\mu) \in \mathbb{R}^N$

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i \in V_n$$

avec $V_n = \mathrm{vect}\{\varphi_1, \ldots, \varphi_n\} \subset \mathbb{R}^N$ et $\dim V_n = n$.

La fonction $\boldsymbol{u}_n$ est une **approximation de faible rang** de $\boldsymbol{u}_h$.

① Etape offline : espace réduit

✗ Cette étape peut être coûteuse mais n'est réalisée qu'une seule fois !

① Etape offline : espace réduit

✗ Cette étape peut être coûteuse mais n'est réalisée qu'une seule fois !

Objectif. On veut construire un espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \ldots, \varphi_n\}.$$

① Etape offline : espace réduit

✗ Cette étape peut être coûteuse mais n'est réalisée qu'une seule fois !

Objectif. On veut construire un espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \ldots, \varphi_n\}.$$

Pour construire V_n on effectue des "pré-calculs" en résolvant M fois

$$F_h(\boldsymbol{u}_h(\mu_i), \mu_i) = 0, \quad i = 1, \ldots, M,$$

pour M valeurs du paramètre $\{\mu_1, \ldots, \mu_M\}$.

On forme la **matrice des snapshots**

$$A = [\boldsymbol{u}_h(\mu_1), \ldots, \boldsymbol{u}_h(\mu_M)] \in \mathbb{R}^{N \times M}.$$

① Etape offline : espace réduit

✗ Cette étape peut être coûteuse mais n'est réalisée qu'une seule fois !

Objectif. On veut construire un espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \dots, \varphi_n\}.$$

Pour construire V_n on effectue des "pré-calculs" en résolvant M fois

$$F_h(\boldsymbol{u}_h(\mu_i), \mu_i) = 0, \quad i = 1, \dots, M,$$

pour M valeurs du paramètre $\{\mu_1, \dots, \mu_M\}$.

On forme la **matrice des snapshots**

$$A = [\boldsymbol{u}_h(\mu_1), \dots, \boldsymbol{u}_h(\mu_M)] \in \mathbb{R}^{N \times M}.$$

On calcule la **décomposition en valeurs singulières** tronquée à n termes de A.

Wait a minute !

✗ Cette étape peut être coûteuse mais n'est réalisée qu'une seule fois !

Objectif. On veut construire un espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \dots, \varphi_n\}.$$

Pour construire V_n on effectue des "pré-calculs" en résolvant M fois

$$F_h(\boldsymbol{u}_h(\mu_i), \mu_i) = 0, \quad i = 1, \dots, M,$$

pour M valeurs du paramètre $\{\mu_1, \dots, \mu_M\}$.

On forme la **matrice des snapshots**

$$A = [\boldsymbol{u}_h(\mu_1), \dots, \boldsymbol{u}_h(\mu_M)] \in \mathbb{R}^{N \times M}.$$

On calcule la **décomposition en valeurs singulières** tronquée à n termes de A.

Wait a minute !

$\Rightarrow V_n$ est l'espace engendré par les n premiers **vecteurs singuliers** à gauche de A.

C'est quoi la décomposition en valeurs singulières ?

On veut représenter $A \in \mathbb{R}^{N \times M}$ sous la forme d'un produit de trois matrices

$$A = \mathbf{\Phi \Sigma \Psi}^T = \sum_{i=1}^{r} \sigma_i \varphi_i \psi_i^T .$$

avec

- les vect. singuliers (orthonormaux) : à gauche $\varphi_i \in \mathbb{R}^N$, à droite $\psi_i \in \mathbb{R}^M$,
- les valeurs singulières $\sigma_i \in \mathbb{R}_+$,
- le rang $r = \mathrm{rang}(A) \leq \min(M, N)$.

C'est quoi la décomposition en valeurs singulières ?

On veut représenter $A \in \mathbb{R}^{N \times M}$ sous la forme d'un produit de trois matrices

$$A = \mathbf{\Phi \Sigma \Psi}^T = \sum_{i=1}^{r} \sigma_i \varphi_i \psi_i^T .$$

avec

- les vect. singuliers (orthonormaux) : à gauche $\varphi_i \in \mathbb{R}^N$, à droite $\psi_i \in \mathbb{R}^M$,
- les valeurs singulières $\sigma_i \in \mathbb{R}_+$,
- le rang $r = \mathrm{rang}(A) \leq \min(M, N)$.

Exemple. Pour $N = 6, M = 5$.

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} \end{pmatrix}}_{A \in \mathbb{R}^{6 \times 5}} = \underbrace{\begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 & \varphi_4 & \varphi_5 & \varphi_6 \end{pmatrix}}_{\mathbf{\Phi} \in \mathbb{R}^{6 \times 6}} \underbrace{\begin{pmatrix} \sigma_{11} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{22} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{33} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{44} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{55} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{\mathbf{\Sigma} \in \mathbb{R}^{6 \times 5}} \underbrace{\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \psi_5 \end{pmatrix}}_{\mathbf{\Psi}^T \in \mathbb{R}^{5 \times 5}}$$

C'est quoi la décomposition en valeurs singulières ?

On veut représenter $A \in \mathbb{R}^{N \times M}$ sous la forme d'un produit de trois matrices

$$A = \mathbf{\Phi \Sigma \Psi}^T = \sum_{i=1}^{r} \sigma_i \varphi_i \psi_i^T .$$

avec

- les vect. singuliers (orthonormaux) : à gauche $\varphi_i \in \mathbb{R}^N$, à droite $\psi_i \in \mathbb{R}^M$,
- les valeurs singulières $\sigma_i \in \mathbb{R}_+$,
- le rang $r = \mathrm{rang}(A) \leq \min(M, N)$.

Exemple. Pour $N = 6, M = 5$.

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \\ a_{61} & a_{62} & a_{63} & a_{64} & a_{65} \end{pmatrix}}_{A \in \mathbb{R}^{6 \times 5}} = \underbrace{\begin{pmatrix} \varphi_1 & \varphi_2 & \varphi_3 & \varphi_4 & \varphi_5 & \varphi_6 \end{pmatrix}}_{\mathbf{\Phi} \in \mathbb{R}^{6 \times 6}} \underbrace{\begin{pmatrix} \sigma_{11} & 0 & 0 & 0 & 0 \\ 0 & \sigma_{22} & 0 & 0 & 0 \\ 0 & 0 & \sigma_{33} & 0 & 0 \\ 0 & 0 & 0 & \sigma_{44} & 0 \\ 0 & 0 & 0 & 0 & \sigma_{55} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}}_{\mathbf{\Sigma} \in \mathbb{R}^{6 \times 5}} \underbrace{\begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \\ \psi_5 \end{pmatrix}}_{\mathbf{\Psi}^T \in \mathbb{R}^{5 \times 5}}$$

Les trois premiers vecteurs singuliers à gauche forment un sous-espace réduit de $\mathbb{R}^6$ de dimension $n = 3$:

$$V_n := \mathrm{vect}\{\varphi_1, \varphi_2, \varphi_3\}.$$

Objectif. On veut construire le modèle réduit

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0.$$

① Etape offline

Objectif. On veut construire le modèle réduit

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0.$$

Ansatz. On cherche une approximation $\boldsymbol{u}_n(\mu)$ de $\boldsymbol{u}_h(\mu)$ dans V_n i.e.

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i = \boldsymbol{V}\boldsymbol{\alpha}_n(\mu) \in V_n,$$

avec $\boldsymbol{V} = [\varphi_1, \ldots, \varphi_n] \in \mathbb{R}^{N\times n}$ et $\boldsymbol{\alpha}_n(\mu) = (\alpha_1(\mu), \ldots, \alpha_n(\mu))^T \in \mathbb{R}^n$.

Objectif. On veut construire le modèle réduit

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0.$$

Ansatz. On cherche une approximation $\boldsymbol{u}_n(\mu)$ de $\boldsymbol{u}_h(\mu)$ dans V_n i.e.

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i = \boldsymbol{V}\boldsymbol{\alpha}_n(\mu) \in V_n,$$

avec $\boldsymbol{V} = [\varphi_1, \ldots, \varphi_n] \in \mathbb{R}^{N\times n}$ et $\boldsymbol{\alpha}_n(\mu) = (\alpha_1(\mu), \ldots, \alpha_n(\mu))^T \in \mathbb{R}^n$.

Projection. Le modèle réduit peut être obtenu par **projection de Galerkin** dans V_n

$$\langle L_h(\mu)\boldsymbol{u}_n(\mu), \boldsymbol{v}\rangle_2 = \langle \boldsymbol{f}_h(\mu), \boldsymbol{v}\rangle_2, \quad \forall \boldsymbol{v} \in V_n,$$

Objectif. On veut construire le modèle réduit

$$F_n(\boldsymbol{u}_n(\mu), \mu) = 0.$$

Ansatz. On cherche une approximation $\boldsymbol{u}_n(\mu)$ de $\boldsymbol{u}_h(\mu)$ dans V_n i.e.

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i = \boldsymbol{V}\boldsymbol{\alpha}_n(\mu) \in V_n,$$

avec $\boldsymbol{V} = [\varphi_1, \ldots, \varphi_n] \in \mathbb{R}^{N\times n}$ et $\boldsymbol{\alpha}_n(\mu) = (\alpha_1(\mu), \ldots, \alpha_n(\mu))^T \in \mathbb{R}^n$.

Projection. Le modèle réduit peut être obtenu par **projection de Galerkin** dans V_n

$$\langle L_h(\mu)\boldsymbol{u}_n(\mu), \boldsymbol{v}\rangle_2 = \langle \boldsymbol{f}_h(\mu), \boldsymbol{v}\rangle_2, \quad \forall \boldsymbol{v} \in V_n,$$

$$\Leftrightarrow \underbrace{\boldsymbol{V}^T L_h(\mu)\boldsymbol{V}}_{L_n(\mu)\in\mathbb{R}^{n\times n}}\boldsymbol{\alpha}_n(\mu) = \underbrace{\boldsymbol{V}^T \boldsymbol{f}_h(\mu)}_{\boldsymbol{f}_n(\mu)\in\mathbb{R}^{n\times n}}.$$

On a obtenu un système linéaire de taille $n \times n$!

② Etape online : résolution de $F_n(u_n(\mu)) = 0$

Système algébrique pour $n << N$

Pour n'importe quelle valeur de $\mu \in \mathcal{P}$ on cherche $\boldsymbol{\alpha}_n(\mu) \in \mathbb{R}^n$ t.q.

$$L_n(\mu)\boldsymbol{\alpha}_n(\mu) = \boldsymbol{f}_n(\mu) \text{ dans } \mathbb{R}^n.$$

pour obtenir l'approximation $\boldsymbol{u}_n(\mu) = \boldsymbol{V}\boldsymbol{\alpha}_n(\mu)$ de $\boldsymbol{u}_h(\mu)$.

② Etape online : résolution de $F_n(u_n(\mu)) = 0$

Système algébrique pour $n << N$

Pour n'importe quelle valeur de $\mu \in \mathcal{P}$ on cherche $\boldsymbol{\alpha}_n(\mu) \in \mathbb{R}^n$ t.q.

$$L_n(\mu)\boldsymbol{\alpha}_n(\mu) = \boldsymbol{f}_n(\mu) \text{ dans } \mathbb{R}^n.$$

pour obtenir l'approximation $\boldsymbol{u}_n(\mu) = \boldsymbol{V}\boldsymbol{\alpha}_n(\mu)$ de $\boldsymbol{u}_h(\mu)$.

✓ Cette étape est "cheap" et peut être réalisée autant de fois que possible !

Estimation d'erreur a priori

Questions.

1. Quelle est l'erreur commise quand $\boldsymbol{u}_h(\mu) \approx \boldsymbol{u}_n(\mu) \in V_n$: $\|\boldsymbol{u}_h(\mu) - \boldsymbol{u}_n(\mu)\|_2$?
2. Peut-on la comparer à l'idéal ?

Questions.

1. Quelle est l'erreur commise quand $\boldsymbol{u}_h(\mu) \approx \boldsymbol{u}_n(\mu) \in V_n$: $\|\boldsymbol{u}_h(\mu) - \boldsymbol{u}_n(\mu)\|_2$?
2. Peut-on la comparer à l'idéal ?

C-à-d la **meilleure approximation** de $\boldsymbol{u}_h(\mu)$ pour $\|\cdot\|_2$ dans V_n

$$\|\boldsymbol{u}_h(\mu) - \mathrm{P}_{V_n}\boldsymbol{u}_h(\mu)\|_2 = \min_{\boldsymbol{v}\in V_n} \|\boldsymbol{u}_h(\mu) - v\|_2.$$

qui est la donnée par la **projection orthogonale** avec $\mathrm{P}_{V_n} = \boldsymbol{V}\boldsymbol{V}^T$.

Estimation d'erreur a priori

On suppose qu'il existe $0 < \alpha < \beta$ t.q. $L_h(\mu) \in \mathbb{R}^{N \times N}$ vérifie

$$\alpha \|\boldsymbol{v}\|_2^2 \leq \langle L_h(\mu)\boldsymbol{v}, \boldsymbol{v}\rangle_2, \quad \forall \boldsymbol{v} \in \mathbb{R}^N. \tag{3}$$

et

$$\langle L_h(\mu)\boldsymbol{u}, \boldsymbol{v}\rangle_2 \leq \beta \|\boldsymbol{u}\|_2 \|\boldsymbol{v}\|_2, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^N. \tag{4}$$

Estimation d'erreur a priori

On suppose qu'il existe $0 < \alpha < \beta$ t.q. $L_h(\mu) \in \mathbb{R}^{N \times N}$ vérifie

$$\alpha \|\boldsymbol{v}\|_2^2 \leq \langle L_h(\mu)\boldsymbol{v}, \boldsymbol{v}\rangle_2, \quad \forall \boldsymbol{v} \in \mathbb{R}^N. \tag{3}$$

et

$$\langle L_h(\mu)\boldsymbol{u}, \boldsymbol{v}\rangle_2 \leq \beta \|\boldsymbol{u}\|_2 \|\boldsymbol{v}\|_2, \quad \forall \boldsymbol{u}, \boldsymbol{v} \in \mathbb{R}^N. \tag{4}$$

Quasi-optimalité

Il existe $\gamma > 0$ (indépendante de n) t.q.

$$\|\boldsymbol{u}_h(\mu) - \boldsymbol{u}_n(\mu)\|_2 = \gamma \min_{\boldsymbol{v} \in V_n} \|\boldsymbol{u}_h(\mu) - v\|_2. \tag{5}$$

Preuve. $\boldsymbol{u}_h$ et $\boldsymbol{u}_n$ satisfont la **l'orthogonalité de Galerkin**

$$\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_n), \boldsymbol{v}\rangle_2 = 0, \quad \forall \boldsymbol{v} \in V_n.$$

Preuve. $\boldsymbol{u}_h$ et $\boldsymbol{u}_n$ satisfont la **l'orthogonalité de Galerkin**

$$\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_n), \boldsymbol{v}\rangle_2 = 0, \quad \forall \boldsymbol{v} \in V_n.$$

On a

$$\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(3)}{\leq} \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \boldsymbol{u}_n - \boldsymbol{u}_h\rangle_2$$

Preuve. $\boldsymbol{u}_h$ et $\boldsymbol{u}_n$ satisfont la **l'orthogonalité de Galerkin**

$$\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_n), \boldsymbol{v}\rangle_2 = 0, \quad \forall \boldsymbol{v} \in V_n.$$

On a

$$\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(3)}{\leq} \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \boldsymbol{u}_n - \boldsymbol{u}_h\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h + \underbrace{\boldsymbol{u}_n - \mathrm{P}_{V_n}\boldsymbol{u}_h}_{\in V_n}\rangle_2$$

Preuve. $\boldsymbol{u}_h$ et $\boldsymbol{u}_n$ satisfont la **l'orthogonalité de Galerkin**

$$\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_n), \boldsymbol{v}\rangle_2 = 0, \quad \forall \boldsymbol{v} \in V_n.$$

On a

$$\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(3)}{\leq} \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \boldsymbol{u}_n - \boldsymbol{u}_h\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h + \underbrace{\boldsymbol{u}_n - \mathrm{P}_{V_n}\boldsymbol{u}_h}_{\in V_n}\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h\rangle_2$$

Preuve. $\boldsymbol{u}_h$ et $\boldsymbol{u}_n$ satisfont la **l'orthogonalité de Galerkin**

$$\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_n), \boldsymbol{v}\rangle_2 = 0, \quad \forall \boldsymbol{v} \in V_n.$$

On a

$$\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(3)}{\leq} \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \boldsymbol{u}_n - \boldsymbol{u}_h\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h + \underbrace{\boldsymbol{u}_n - \mathrm{P}_{V_n}\boldsymbol{u}_h}_{\in V_n}\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(4)}{\leq} \frac{\beta}{\alpha}\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2 \ \|\mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h\|_2$$

Preuve. $\boldsymbol{u}_h$ et $\boldsymbol{u}_n$ satisfont la **l'orthogonalité de Galerkin**

$$\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_n), \boldsymbol{v}\rangle_2 = 0, \quad \forall \boldsymbol{v} \in V_n.$$

On a

$$\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(3)}{\leq} \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \boldsymbol{u}_n - \boldsymbol{u}_h\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h + \underbrace{\boldsymbol{u}_n - \mathrm{P}_{V_n}\boldsymbol{u}_h}_{\in V_n}\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \leq \frac{1}{\alpha}\langle L_h(\boldsymbol{u}_n - \boldsymbol{u}_h), \mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h\rangle_2$$

$$\Rightarrow \|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2^2 \overset{(4)}{\leq} \frac{\beta}{\alpha}\|\boldsymbol{u}_n - \boldsymbol{u}_h\|_2 \ \|\mathrm{P}_{V_n}\boldsymbol{u}_h - \boldsymbol{u}_h\|_2$$

D'où le résultat (5) en simplifiant et posant $\gamma = \frac{\beta}{\alpha}$.

Illustration

On cherche $u(\mu) : \mathbb{R} \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} -\mu_2 u''(x;\mu) + \mu_1 u'(x;\mu) & = & f(x), \quad x \in (0,1), \\ u(1;\mu) = u(0;\mu) & = & 0. \end{array} \right. \tag{6}$$

On a $\mu = (\mu_1, \mu_2) \in \mathcal{P}$ avec $\mathcal{P} = [1, 10] \times [0.01, 1]$ et $f(x) = 1$. Ici $N = 1000$.

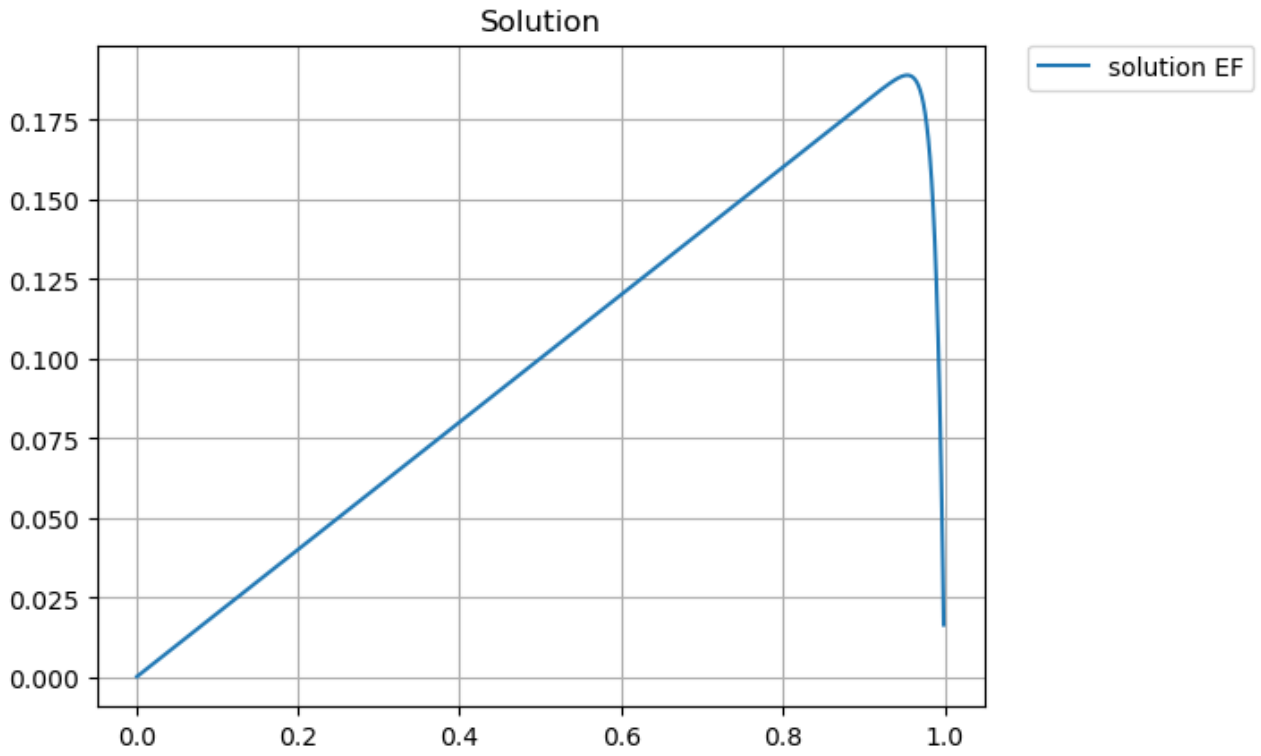


Figure – Solution numérique pour $\mu = (5, 0.05)$

Illustration

On cherche $u(\mu) : \mathbb{R} \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} -\mu_2 u''(x;\mu) + \mu_1 u'(x;\mu) & = & f(x), \quad x \in (0,1), \\ u(1;\mu) = u(0;\mu) & = & 0. \end{array} \right. \tag{6}$$

On a $\mu = (\mu_1, \mu_2) \in \mathcal{P}$ avec $\mathcal{P} = [1, 10] \times [0.01, 1]$ et $f(x) = 1$. Ici $N = 1000$.

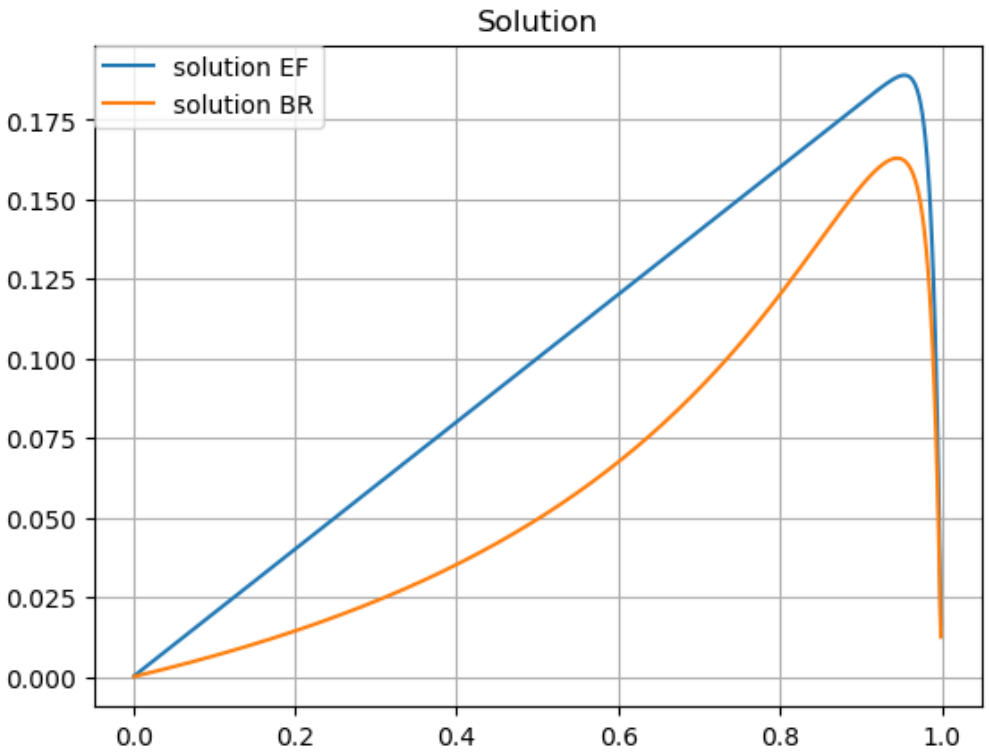


Figure – Base apprise pour $M = 50$ valeurs de μ. Solution calculée pour $n = 2$ et $\mu = (5, 0.05)$

Erreur relative pour $n = 2$: $err_2(\mu) = 0.6663235211802049$
Erreur relative pour $n = 20$: $err_{20}(\mu) = 3.037748791261398e - 07$

Illustration

On cherche $u(\mu) : \mathbb{R} \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} -\mu_2 u''(x;\mu) + \mu_1 u'(x;\mu) & = & f(x), \quad x \in (0,1), \\ u(1;\mu) = u(0;\mu) & = & 0. \end{array} \right. \tag{6}$$

On a $\mu = (\mu_1, \mu_2) \in \mathcal{P}$ avec $\mathcal{P} = [1, 10] \times [0.01, 1]$ et $f(x) = 1$. Ici $N = 1000$.

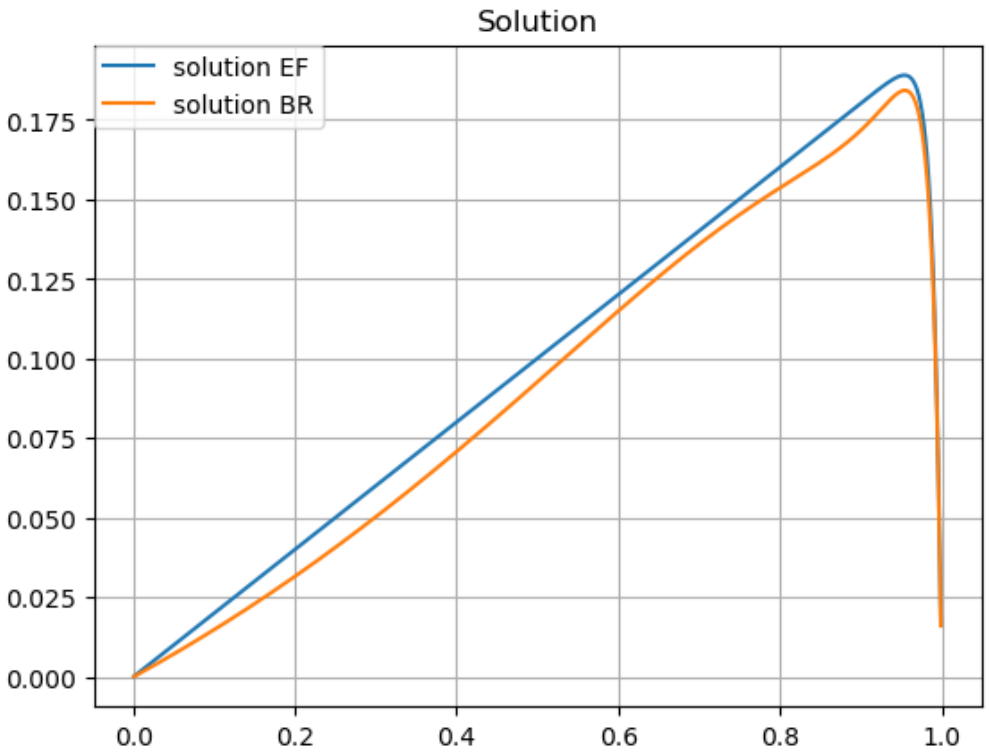


Figure – Base apprise pour $M = 50$ valeurs de μ. Solution calculée pour $n = 5$ et $\mu = (5, 0.05)$

Erreur relative pour $n = 2$: $err_2(\mu) = 0.6663235211802049$
Erreur relative pour $n = 20$: $err_{20}(\mu) = 3.037748791261398e - 07$

Illustration

On cherche $u(\mu) : \mathbb{R} \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} -\mu_2 u''(x;\mu) + \mu_1 u'(x;\mu) & = & f(x), \quad x \in (0,1), \\ u(1;\mu) = u(0;\mu) & = & 0. \end{array} \right. \tag{6}$$

On a $\mu = (\mu_1, \mu_2) \in \mathcal{P}$ avec $\mathcal{P} = [1, 10] \times [0.01, 1]$ et $f(x) = 1$. Ici $N = 1000$.

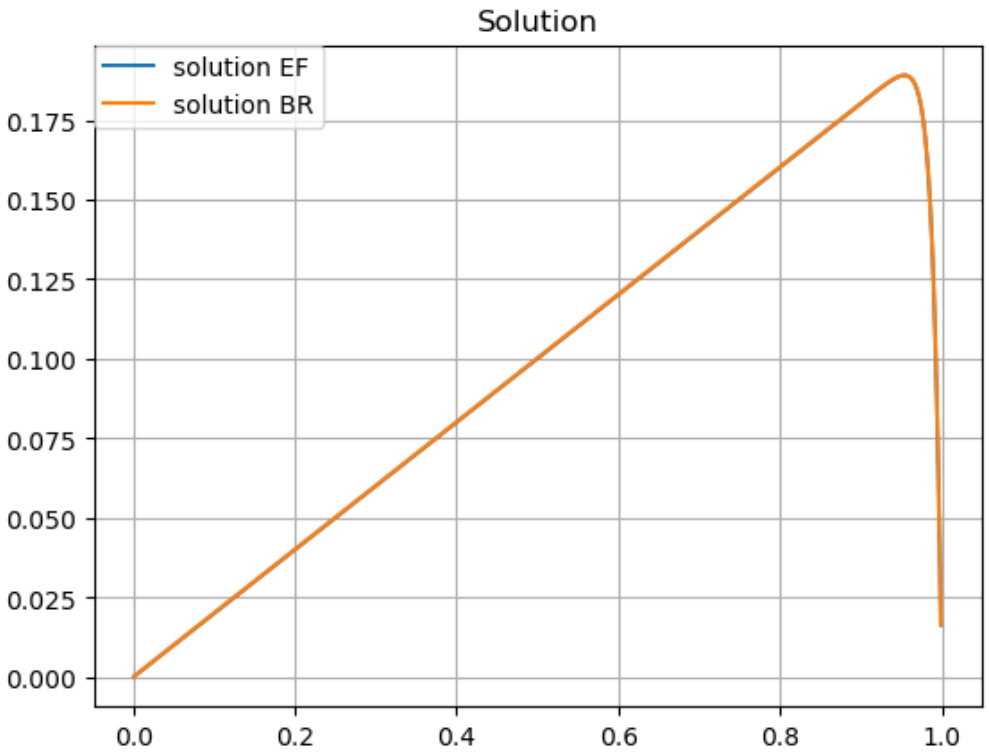


Figure – Base apprise pour $M = 50$ valeurs de μ. Solution calculée pour $n = 10$ et $\mu = (5, 0.05)$

Erreur relative pour $n = 2$: $err_2(\mu) = 0.6663235211802049$
Erreur relative pour $n = 20$: $err_{20}(\mu) = 3.037748791261398e - 07$

Illustration

On cherche $u(\mu) : \mathbb{R} \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} -\mu_2 u''(x;\mu) + \mu_1 u'(x;\mu) & = & f(x), \quad x \in (0,1), \\ u(1;\mu) = u(0;\mu) & = & 0. \end{array} \right. \tag{6}$$

On a $\mu = (\mu_1, \mu_2) \in \mathcal{P}$ avec $\mathcal{P} = [1, 10] \times [0.01, 1]$ et $f(x) = 1$. Ici $N = 1000$.

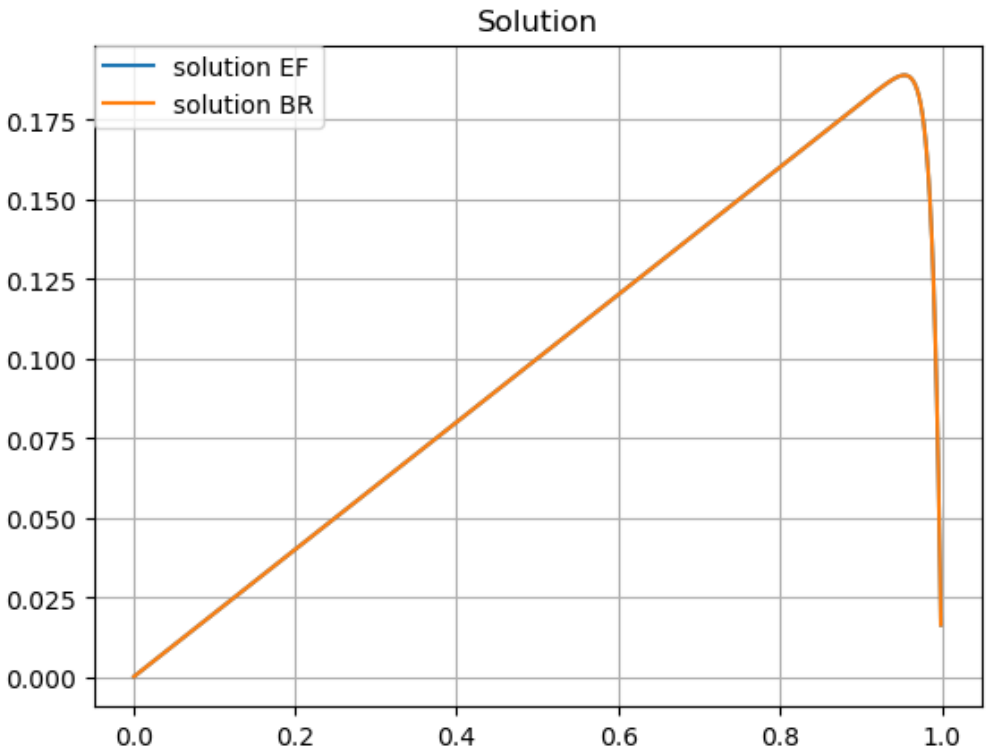


Figure – Base apprise pour $M = 50$ valeurs de μ. Solution calculée pour $n = 20$ et $\mu = (5, 0.05)$

Erreur relative pour $n = 2$: $err_2(\mu) = 0.6663235211802049$
Erreur relative pour $n = 20$: $err_{20}(\mu) = 3.037748791261398e - 07$

Plan

Problème modèle (simple).

On cherche $u : [0, T] \times \Omega \to \mathbb{R}$ solution d'une EDP d'évolution

$$\begin{cases} \partial_t u = \mathcal{L} u, \text{ dans } (0, T] \times \Omega, \\ u(0, \cdot) = u_0, \text{ dans } \Omega, \\ + \text{ conditions aux limites sur } \partial\Omega. \end{cases} \tag{7}$$

Problème modèle (simple).

On cherche $u : [0, T] \times \Omega \to \mathbb{R}$ solution d'une EDP d'évolution

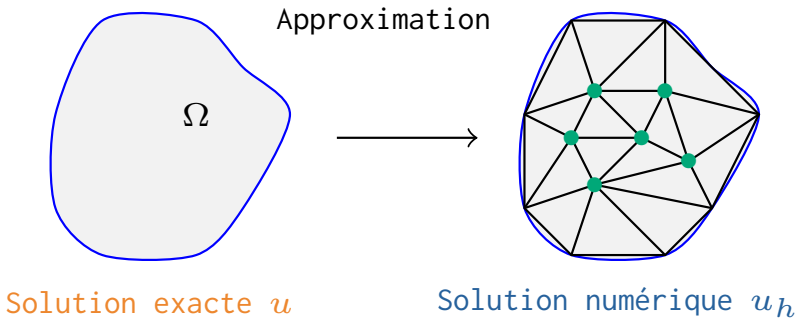
$$\begin{cases} \partial_t u = \mathcal{L}u, \text{ dans } (0, T] \times \Omega, \\ u(0, \cdot) = u_0, \text{ dans } \Omega, \\ + \text{ conditions aux limites sur } \partial\Omega. \end{cases} \tag{7}$$

avec

- $\Omega \subset \mathbb{R}^d$ le domaine spatial
- $\mathcal{L}$ un opérateur différentiel linéaire

ex. : Equation de la chaleur $\mathcal{L}u = \nu \triangle u$

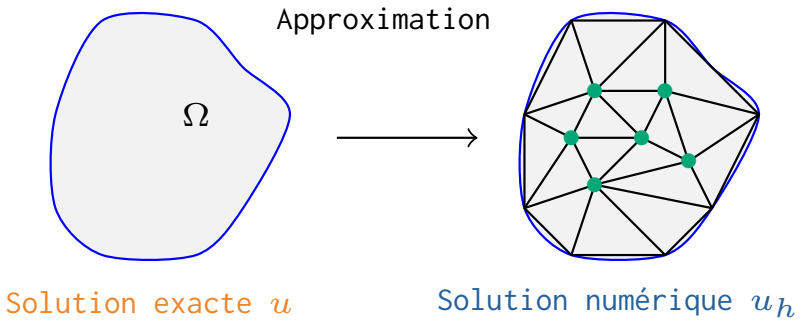
Approximation.



On veut calculer numériquement une approximation $u_h(t)$ de $u(t)$.

ex. : Méthode éléments finis, différences finies etc.

Approximation.



On veut calculer numériquement une approximation $u_h(t)$ de $u(t)$.

ex. : Méthode éléments finis, différences finies etc.

Système d'EDO de taille N.

En pratique, pour obtenir $u_h \approx u$, il faut calculer le vecteur des coefficients

$$\boldsymbol{u}_h : [0, T] \to \mathbb{R}^N$$

solution de

$$\frac{d}{dt}\boldsymbol{u}_h(t) = L_h \boldsymbol{u}_h(t), \text{ dans } \mathbb{R}^N \tag{8}$$

avec $L_h \in \mathbb{R}^{N \times N}$.

ex. : Pour les différences finies : $\mathcal{L} = \frac{\partial^2}{\partial x^2} \Rightarrow L_h = -\frac{1}{h^2}\mathrm{tridiag}(-1, 2, -1)$.

① Etape offline

L'espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \ldots, \varphi_n\}$$

est engendré par les n premiers **vecteurs singuliers** à gauche de

$$A = [\boldsymbol{u}_h(t_1), \ldots, \boldsymbol{u}_h(t_M)] \in \mathbb{R}^{N \times M}.$$

pour M instants $\{t_1, \ldots, t_M\} \subset [0, T]$.

L'espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \ldots, \varphi_n\}$$

est engendré par les n premiers **vecteurs singuliers** à gauche de

$$A = [\boldsymbol{u}_h(t_1), \ldots, \boldsymbol{u}_h(t_M)] \in \mathbb{R}^{N \times M}.$$

pour M instants $\{t_1, \ldots, t_M\} \subset [0, T]$.

On cherche une approximation $\boldsymbol{u}_n(t) = \boldsymbol{V}\boldsymbol{\alpha}_n(t) \in V_n$ solution de

$$F_n(u_n) = 0,$$

L'espace réduit

$$V_n = \mathrm{vect}\{\varphi_1, \dots, \varphi_n\}$$

est engendré par les n premiers **vecteurs singuliers** à gauche de

$$A = [\boldsymbol{u}_h(t_1), \dots, \boldsymbol{u}_h(t_M)] \in \mathbb{R}^{N \times M}.$$

pour M instants $\{t_1, \dots, t_M\} \subset [0, T]$.

On cherche une approximation $\boldsymbol{u}_n(t) = \boldsymbol{V}\boldsymbol{\alpha}_n(t) \in V_n$ solution de

$$F_n(u_n) = 0,$$

obtenu par **projection de Galerkin** dans V_n

$$\frac{d}{dt}\boldsymbol{\alpha}_n(t) = \underbrace{\boldsymbol{V}^T L_h \boldsymbol{V}}_{L_n \in \mathbb{R}^{n \times n}} \boldsymbol{\alpha}_n(t).$$

On a obtenu un système d'EDOs de taille $n \times n$!

② Etape online : résolution de $F_n(u_n(\mu)) = 0$

Système d'EDO de taille $n << N$

On cherche $\boldsymbol{\alpha}_n : [0, T] \to \mathbb{R}^n$ solution de

$$\frac{d}{dt}\boldsymbol{\alpha}_n(t) = L_n \boldsymbol{\alpha}_n(t), t \in (0, T]$$

avec la condition initiale $\boldsymbol{\alpha}(0) = \boldsymbol{V}^T \boldsymbol{u}_{h,0}$.

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Soit ℓ la constante de lipschitz logarithmique pour $L_h(\mu) \in \mathbb{R}^{N\times N}$ t.q.

$$\ell := \sup_{x\neq 0} \frac{\langle L_h \boldsymbol{v}, \boldsymbol{v}\rangle_2}{\|\boldsymbol{v}\|_2} = \lambda_{max}\left(\frac{1}{2}(L_h + L_h^T)\right). \tag{9}$$

Borne de l'erreur

Pour tout $t \in [0, T]$, on a

$$\|\boldsymbol{u}_h(t) - \boldsymbol{u}_n(t)\|_2 \leq e^{\ell t}\underbrace{\|\boldsymbol{u}_n(0) - \boldsymbol{u}_h(0)\|_2}_{①} + \int_0^t e^{\ell(t-s)}\underbrace{\|(\mathrm{I}_N - \mathrm{P}_{V_n})L_h\boldsymbol{u}_n(s)\|_2}_{②}ds, \tag{10}$$

avec les erreurs de projection dans V_n

- ① de la condition initiale : $\boldsymbol{u}_n(0) = \mathrm{P}_{V_n}\boldsymbol{u}_h(0)$,
- ② du "second membre" :$(\mathrm{I}_N - \mathrm{P}_{V_n})L_h\boldsymbol{u}_n$.

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

En soustrayant membre à membre

$$\frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h \boldsymbol{u}_h - \mathrm{P}_{V_n} L_h \boldsymbol{u}_n$$

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

En soustrayant membre à membre

$$\frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h \boldsymbol{u}_h - \mathrm{P}_{V_n} L_h \boldsymbol{u}_n$$

$$\Rightarrow \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h(\boldsymbol{u}_h - \boldsymbol{u}_n) + (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n$$

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

En soustrayant membre à membre

$$\frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h \boldsymbol{u}_h - \mathrm{P}_{V_n} L_h \boldsymbol{u}_n$$

$$\Rightarrow \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h(\boldsymbol{u}_h - \boldsymbol{u}_n) + (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n$$

$$\Rightarrow \langle \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 + \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

En soustrayant membre à membre

$$\frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h \boldsymbol{u}_h - \mathrm{P}_{V_n} L_h \boldsymbol{u}_n$$

$$\Rightarrow \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h(\boldsymbol{u}_h - \boldsymbol{u}_n) + (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n$$

$$\Rightarrow \langle \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 + \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

$$\Rightarrow \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \frac{d}{dt}\|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 - \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

En soustrayant membre à membre

$$\frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h \boldsymbol{u}_h - \mathrm{P}_{V_n} L_h \boldsymbol{u}_n$$

$$\Rightarrow \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h(\boldsymbol{u}_h - \boldsymbol{u}_n) + (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n$$

$$\Rightarrow \langle \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 + \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

$$\Rightarrow \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \frac{d}{dt}\|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 - \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

$$\Rightarrow \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \frac{d}{dt}\|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \overset{(9)}{\leq} \ell \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2^2 + \|(\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n\| \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2.$$

Estimation d'erreur a priori $\boldsymbol{u}_h(t) \approx \boldsymbol{u}_n(t)$

Preuve. Rappelons que pour tout $t \in (0, T]$ on a

$$\frac{d}{dt}\boldsymbol{u}_h = L_h \boldsymbol{u}_h, \qquad \frac{d}{dt}\boldsymbol{u}_n = \mathrm{P}_{V_n} L_h \boldsymbol{u}_n.$$

En soustrayant membre à membre

$$\frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h \boldsymbol{u}_h - \mathrm{P}_{V_n} L_h \boldsymbol{u}_n$$

$$\Rightarrow \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n) = L_h(\boldsymbol{u}_h - \boldsymbol{u}_n) + (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n$$

$$\Rightarrow \langle \frac{d}{dt}(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 + \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

$$\Rightarrow \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \frac{d}{dt}\|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 = \langle L_h(\boldsymbol{u}_h - \boldsymbol{u}_n), \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2 - \langle (\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n, \boldsymbol{u}_h - \boldsymbol{u}_n \rangle_2.$$

$$\Rightarrow \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \frac{d}{dt}\|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \overset{(9)}{\leq} \ell \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2^2 + \|(\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n\| \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2.$$

$$\Rightarrow \frac{d}{dt}\|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 \leq \ell \|\boldsymbol{u}_h - \boldsymbol{u}_n\|_2 + \|(\mathrm{I}_N - \mathrm{P}_{V_n}) L_h \boldsymbol{u}_n\|.$$

On obtient le résultat par le lemme de comparaison.

Schéma d'intégration temporel

Soit $\{t^0, \ldots, t^K\}$ un maillage uniforme de $[0, T]$: $t^k = k\delta t, \delta t = \frac{T}{K}$.

Schéma d'intégration temporel

Soit $\{t^0, \ldots, t^K\}$ un maillage uniforme de $[0, T]$: $t^k = k\delta t, \delta t = \frac{T}{K}$.

On cherche une approximation de la fonction $\boldsymbol{u}_n$ en t^k i.e.

$$\boldsymbol{u}_n(t^k) \approx \boldsymbol{u}_n^k := \boldsymbol{V}\boldsymbol{\alpha}_n^k$$

e.g. par schéma **d'Euler implicite** qui permet de construire $\{\boldsymbol{\alpha}_n^k\}_{k=0}^K$ en résolvant

$$(\mathrm{I}_N - \delta t L_n)\boldsymbol{\alpha}_n^{k+1} = \boldsymbol{\alpha}_n^k, \quad k = 0, \ldots, K-1$$

avec $\boldsymbol{\alpha}^0 = \boldsymbol{\alpha}(0)$.

Schéma d'intégration temporel

Soit $\{t^0, \ldots, t^K\}$ un maillage uniforme de $[0, T]$: $t^k = k\delta t, \delta t = \frac{T}{K}$.

On cherche une approximation de la fonction $\boldsymbol{u}_n$ en t^k i.e.

$$\boldsymbol{u}_n(t^k) \approx \boldsymbol{u}_n^k := \boldsymbol{V}\boldsymbol{\alpha}_n^k$$

e.g. par schéma **d'Euler implicite** qui permet de construire $\{\boldsymbol{\alpha}_n^k\}_{k=0}^K$ en résolvant

$$(\mathrm{I}_N - \delta t L_n)\boldsymbol{\alpha}_n^{k+1} = \boldsymbol{\alpha}_n^k, \quad k = 0, \ldots, K-1$$

avec $\boldsymbol{\alpha}^0 = \boldsymbol{\alpha}(0)$.

✓ On résout une succession de petits problèmes linéaires de taille $n \times n$, K fois !

Illustration

On cherche $u : [0,T] \times \Omega \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} \partial_t u(t,x) & = & \partial^2_{xx} u(t,x), \quad x \in \Omega, \\ u(t,0) & = & u(t,1) = 0, \end{array} \right. \tag{11}$$

avec $u(0,x) = \mathbf{1}_{[0.1,0.3]}(x) - \mathbf{1}_{[0.6,0.8]}(x)$, $T = 0.5$. Ici $K = 200$, $N = 1000$.

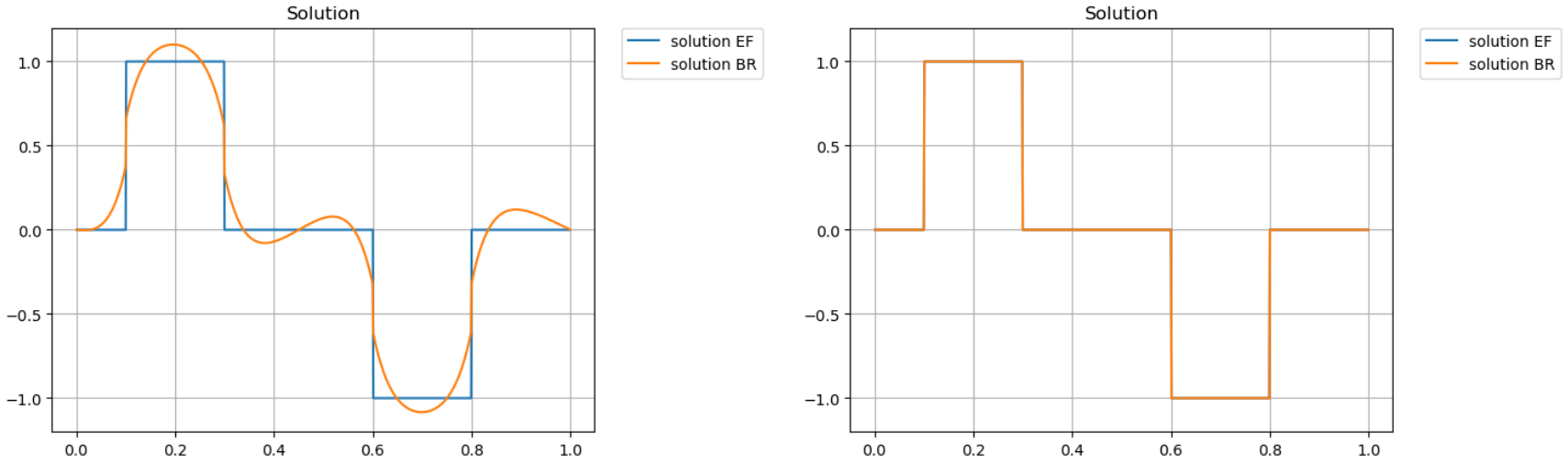


Figure – Base apprise pour $M = 50$. Solution à $t = 0$ pour $n = 2$ (gauche) $n = 20$ (droite).

Erreur relative pour $n = 2$: $err_2 = 0.17475296041661006$
Erreur relative pour $n = 20$: $err_{20} = 1.7279662844846755e - 05$

Illustration

On cherche $u : [0,T] \times \Omega \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} \partial_t u(t,x) & = & \partial^2_{xx} u(t,x), \quad x \in \Omega, \\ u(t,0) & = & u(t,1) = 0, \end{array} \right. \tag{11}$$

avec $u(0,x) = \mathbf{1}_{[0.1,0.3]}(x) - \mathbf{1}_{[0.6,0.8]}(x)$, $T = 0.5$. Ici $K = 200$, $N = 1000$.

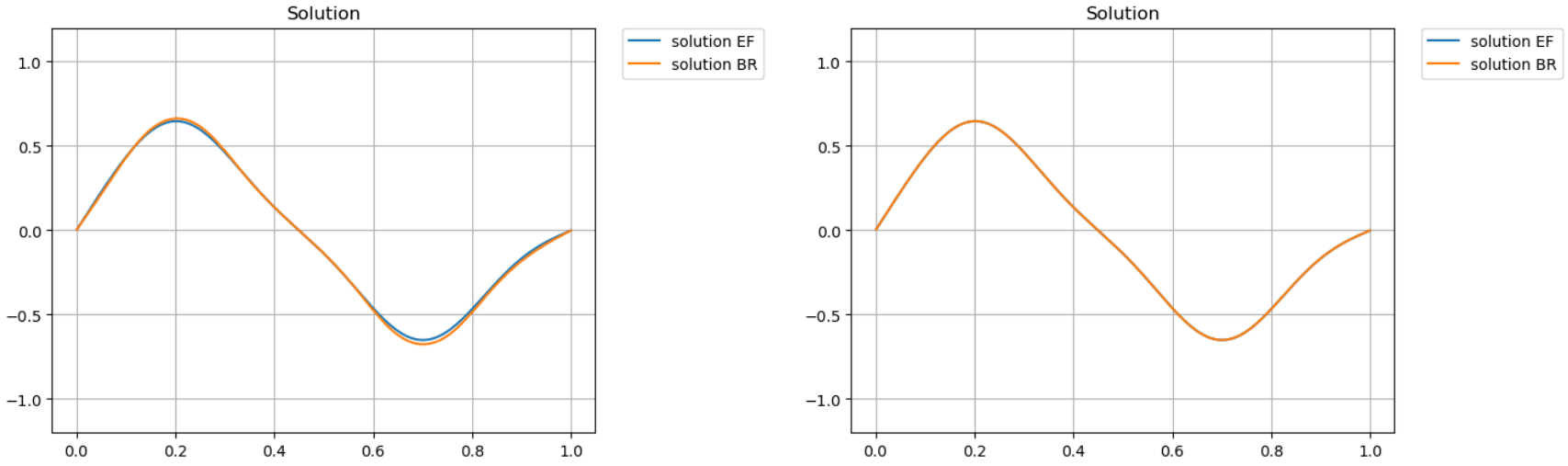


Figure – Base apprise pour $M = 50$. Solution à $t = 0.05$ pour $n = 2$ (gauche) $n = 20$ (droite).

Erreur relative pour $n = 2$: $err_2 = 0.17475296041661006$
Erreur relative pour $n = 20$: $err_{20} = 1.7279662844846755e - 05$

Illustration

On cherche $u : [0,T] \times \Omega \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} \partial_t u(t,x) & = & \partial^2_{xx} u(t,x), \quad x \in \Omega, \\ u(t,0) & = & u(t,1) = 0, \end{array} \right. \tag{11}$$

avec $u(0,x) = \mathbf{1}_{[0.1,0.3]}(x) - \mathbf{1}_{[0.6,0.8]}(x)$, $T = 0.5$. Ici $K = 200$, $N = 1000$.

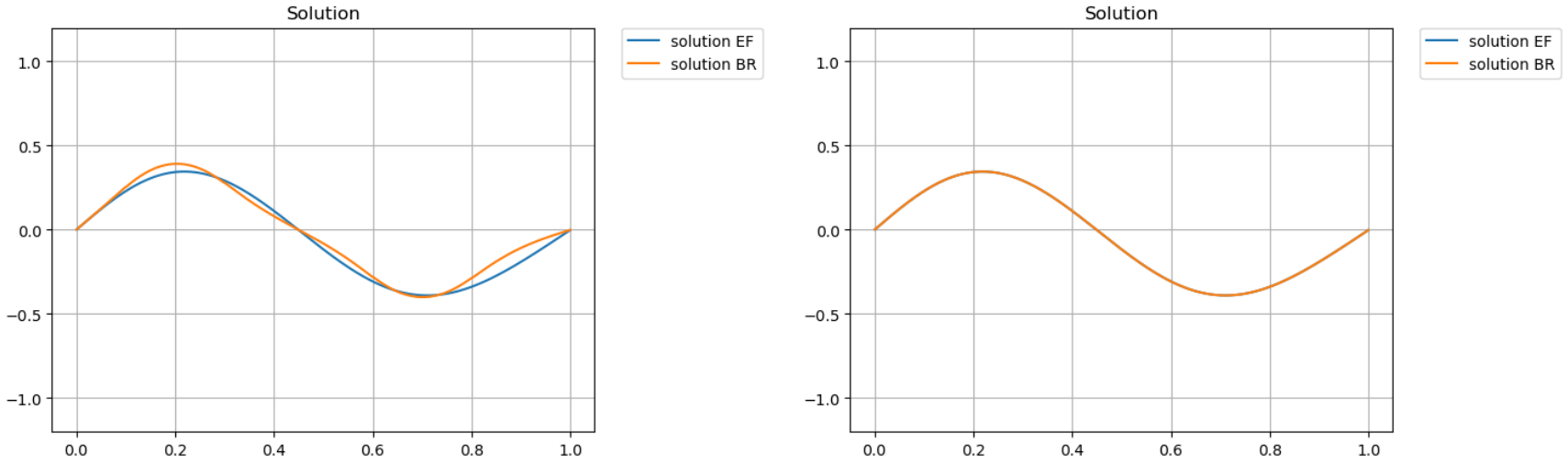


Figure – Base apprise pour $M = 50$. Solution à $t = 0.15$ pour $n = 2$ (gauche) $n = 20$ (droite).

Erreur relative pour $n = 2$: $err_2 = 0.17475296041661006$
Erreur relative pour $n = 20$: $err_{20} = 1.7279662844846755e - 05$

Illustration

On cherche $u : [0,T] \times \Omega \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} \partial_t u(t,x) & = & \partial^2_{xx} u(t,x), \quad x \in \Omega, \\ u(t,0) & = & u(t,1) = 0, \end{array} \right. \tag{11}$$

avec $u(0,x) = \mathbf{1}_{[0.1,0.3]}(x) - \mathbf{1}_{[0.6,0.8]}(x)$, $T = 0.5$. Ici $K = 200$, $N = 1000$.

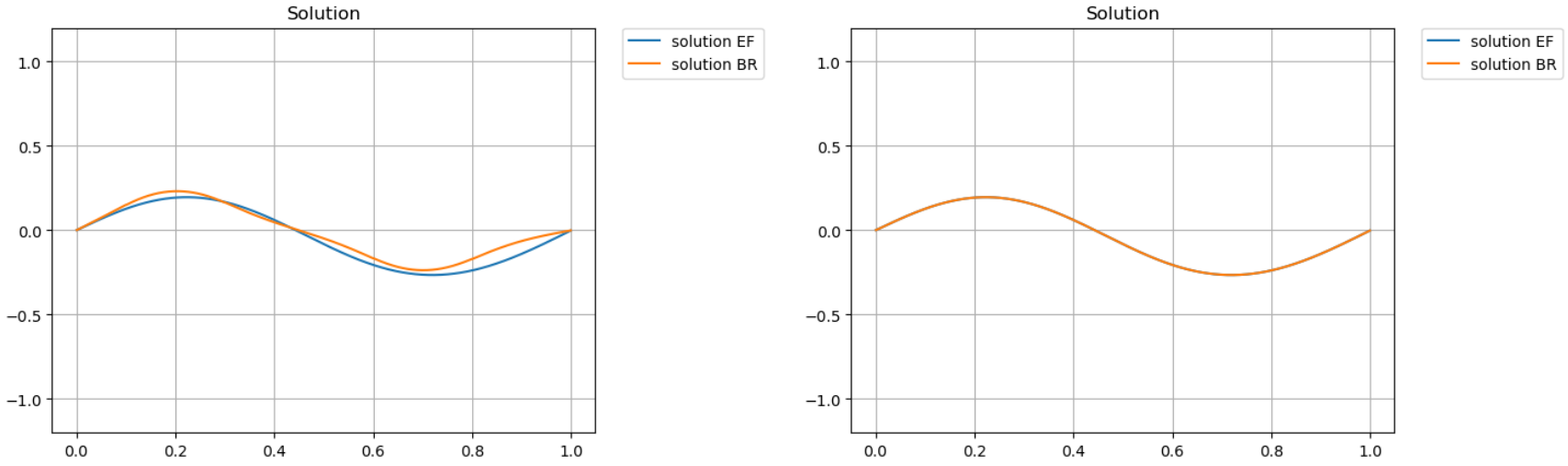


Figure – Base apprise pour $M = 50$. Solution à $t = 0.25$ pour $n = 2$ (gauche) $n = 20$ (droite).

Erreur relative pour $n = 2$: $err_2 = 0.17475296041661006$
Erreur relative pour $n = 20$: $err_{20} = 1.7279662844846755e - 05$

Illustration

On cherche $u : [0,T] \times \Omega \to \mathbb{R}$ solution de

$$\left\{ \begin{array}{rcl} \partial_t u(t,x) & = & \partial^2_{xx} u(t,x), \quad x \in \Omega, \\ u(t,0) & = & u(t,1) = 0, \end{array} \right. \tag{11}$$

avec $u(0,x) = \mathbf{1}_{[0.1,0.3]}(x) - \mathbf{1}_{[0.6,0.8]}(x)$, $T = 0.5$. Ici $K = 200$, $N = 1000$.

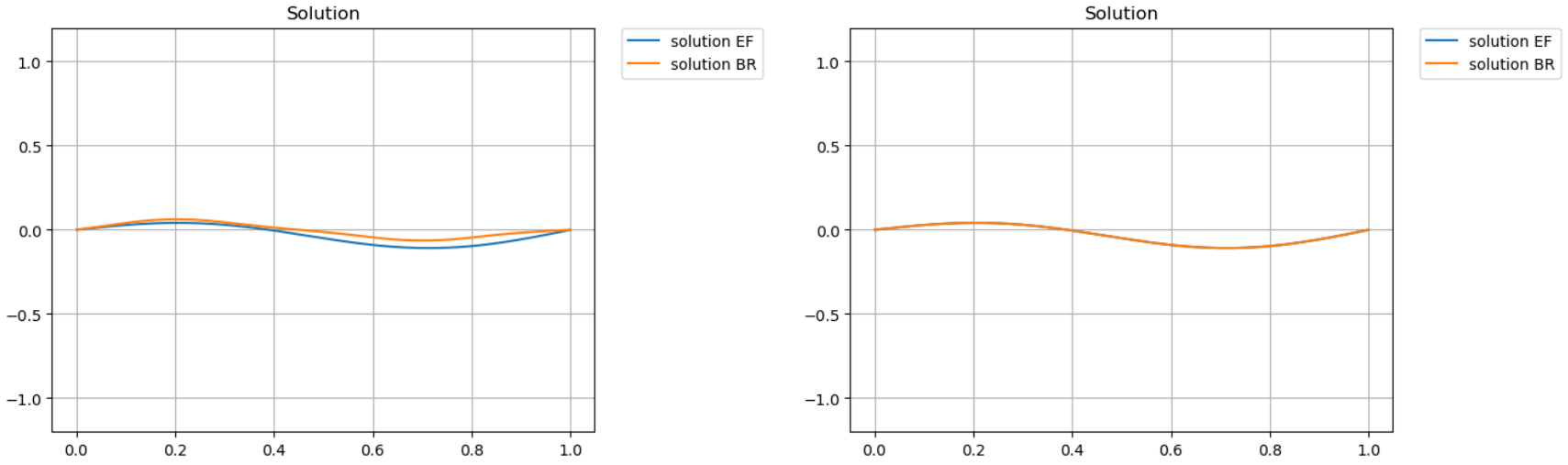


Figure – Base apprise pour $M = 50$. Solution à $t = 0.5$ pour $n = 2$ (gauche) $n = 20$ (droite).

Erreur relative pour $n = 2$: $err_2 = 0.17475296041661006$
Erreur relative pour $n = 20$: $err_{20} = 1.7279662844846755e - 05$

Pourquoi ça marche ?

Bilan.

Equation de la chaleur

$$\boldsymbol{u}_n(t) = \sum_{i=1}^{n} \alpha_i(t)\varphi_i \in V_n$$

Equation d'advection diffusion

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i \in V_n$$

Pourquoi ça marche ?

Bilan.

Equation de la chaleur

$$\boldsymbol{u}_n(t) = \sum_{i=1}^{n} \alpha_i(t)\varphi_i \in V_n$$

Equation d'advection diffusion

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i \in V_n$$

Question. Pourquoi une approximation sous format $\boldsymbol{u}_n \approx \boldsymbol{u}_h$ de rang n "fonctionne" ?

Pourquoi ça marche ?

Bilan.

Equation de la chaleur

$$\boldsymbol{u}_n(t) = \sum_{i=1}^{n} \alpha_i(t)\varphi_i \in V_n$$

Equation d'advection diffusion

$$\boldsymbol{u}_n(\mu) = \sum_{i=1}^{n} \alpha_i(\mu)\varphi_i \in V_n$$

Question. Pourquoi une approximation sous format $\boldsymbol{u}_n \approx \boldsymbol{u}_h$ de rang n "fonctionne" ?

Réponse. On cherche des approximations par le biais de projections dans des espaces V_n (quasi)-optimaux.

Optimalité de la base réduite [QMN, §6.3.1.]

Soit

$$\mathcal{V}_n := \{\boldsymbol{W} \in \mathbb{R}^{N\times n} : \boldsymbol{W}\boldsymbol{W}^T = \mathrm{I}_n\}$$

l'ensemble des matrices orthonormées. Alors

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \boldsymbol{\Phi}_n\boldsymbol{\Phi}_n^T\boldsymbol{u}_h(\mu_i)\|_2^2 = \min_{\mathrm{W}\in\mathcal{V}_n} \sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \boldsymbol{W}\boldsymbol{W}^T\boldsymbol{u}_h(\mu_i)\|_2^2 = \sum_{i=n+1}^{r} \sigma_i^2, \quad (12)$$

Optimalité de la base réduite [QMN, §6.3.1.]

Soit

$$\mathcal{V}_n := \{\boldsymbol{W} \in \mathbb{R}^{N\times n} : \boldsymbol{W}\boldsymbol{W}^T = \mathrm{I}_n\}$$

l'ensemble des matrices orthonormées. Alors

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \boldsymbol{\Phi}_n\boldsymbol{\Phi}_n^T\boldsymbol{u}_h(\mu_i)\|_2^2 = \min_{\mathrm{W}\in\mathcal{V}_n} \sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \boldsymbol{W}\boldsymbol{W}^T\boldsymbol{u}_h(\mu_i)\|_2^2 = \sum_{i=n+1}^{r} \sigma_i^2, \quad (12)$$

avec

- $\boldsymbol{\Phi}_n = [\varphi_1, \ldots, \varphi_n] \in \mathbb{R}^{N\times n}$,
- $\{\sigma_i\}_{i=1}^r \subset \mathbb{R}_+$ les valeurs singulières de
- la matrice de snapshots $A = [\boldsymbol{u}_h(\mu_1), \ldots, \boldsymbol{u}_h(\mu_M)] \in \mathbb{R}^{N\times n}$.

Pourquoi ça marche ?

Interprétation. Pour tout $\boldsymbol{W} \in \mathcal{V}_n$, on a $P_V = \boldsymbol{W}\boldsymbol{W}^T$ pour $V = \mathrm{vectcol}(\boldsymbol{W})$.

Interprétation. Pour tout $\boldsymbol{W} \in \mathcal{V}_n$, on a $P_V = \boldsymbol{W}\boldsymbol{W}^T$ pour $V = \mathrm{vectcol}(\boldsymbol{W})$.

Considérons

$$\mathcal{M}_h^M = \{\boldsymbol{u}_h(\mu_i) : i = 1, \ldots, M\},$$

l'ensemble des solutions associé aux $\{\mu_1, \ldots, \mu_M\} \subset \mathcal{P}$.

Interprétation. Pour tout $\boldsymbol{W} \in \mathcal{V}_n$, on a $P_V = \boldsymbol{W}\boldsymbol{W}^T$ pour $V = \mathrm{vectcol}(\boldsymbol{W})$.

Considérons

$$\mathcal{M}_h^M = \{\boldsymbol{u}_h(\mu_i) : i = 1, \ldots, M\},$$

l'ensemble des solutions associé aux $\{\mu_1, \ldots, \mu_M\} \subset \mathcal{P}$. Ainsi (12) implique

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 = \min_{V \subset \mathbb{R}^N, \dim V = n} \sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - P_V \boldsymbol{u}_h(\mu_i)\|_2^2 = \sum_{i=n+1}^{r} \sigma_i^2.$$

Pourquoi ça marche ?

Interprétation. Pour tout $\boldsymbol{W} \in \mathcal{V}_n$, on a $P_V = \boldsymbol{W}\boldsymbol{W}^T$ pour $V = \mathrm{vectcol}(\boldsymbol{W})$.

Considérons

$$\mathcal{M}_h^M = \{\boldsymbol{u}_h(\mu_i) : i = 1, \dots, M\},$$

l'ensemble des solutions associé aux $\{\mu_1, \dots, \mu_M\} \subset \mathcal{P}$. Ainsi (12) implique

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 = \min_{V \subset \mathbb{R}^N, \dim V = n} \sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - P_V \boldsymbol{u}_h(\mu_i)\|_2^2 = \sum_{i=n+1}^{r} \sigma_i^2.$$

Cela signifie que l'espace $V_n = \mathrm{vectcol}(\Phi) = \mathrm{vect}\{\varphi_1, \dots, \varphi_n\}$

1. est **optimal** i.e. le meilleur sous-espace de dimension n de $\mathbb{R}^N$

Interprétation. Pour tout $\boldsymbol{W} \in \mathcal{V}_n$, on a $P_V = \boldsymbol{W}\boldsymbol{W}^T$ pour $V = \mathrm{vectcol}(\boldsymbol{W})$.

Considérons

$$\mathcal{M}_h^M = \{\boldsymbol{u}_h(\mu_i) : i = 1, \dots, M\},$$

l'ensemble des solutions associé aux $\{\mu_1, \dots, \mu_M\} \subset \mathcal{P}$. Ainsi (12) implique

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 = \min_{V \subset \mathbb{R}^N, \dim V = n} \sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - P_V \boldsymbol{u}_h(\mu_i)\|_2^2 = \sum_{i=n+1}^{r} \sigma_i^2.$$

Cela signifie que l'espace $V_n = \mathrm{vectcol}(\Phi) = \mathrm{vect}\{\varphi_1, \dots, \varphi_n\}$

1. est **optimal** i.e. le meilleur sous-espace de dimension n de $\mathbb{R}^N$
2. qui approche $\mathcal{M}_h^M$ en norme 2.

Interprétation. Pour tout $\boldsymbol{W} \in \mathcal{V}_n$, on a $P_V = \boldsymbol{W}\boldsymbol{W}^T$ pour $V = \mathrm{vectcol}(\boldsymbol{W})$.

Considérons

$$\mathcal{M}_h^M = \{\boldsymbol{u}_h(\mu_i) : i = 1, \dots, M\},$$

l'ensemble des solutions associé aux $\{\mu_1, \dots, \mu_M\} \subset \mathcal{P}$. Ainsi (12) implique

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 = \min_{V \subset \mathbb{R}^N, \dim V = n} \sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - P_V \boldsymbol{u}_h(\mu_i)\|_2^2 = \sum_{i=n+1}^{r} \sigma_i^2.$$

Cela signifie que l'espace $V_n = \mathrm{vectcol}(\Phi) = \mathrm{vect}\{\varphi_1, \dots, \varphi_n\}$

1. est **optimal** i.e. le meilleur sous-espace de dimension n de $\mathbb{R}^N$
2. qui approche $\mathcal{M}_h^M$ en norme 2.
3. L'erreur commise quand on considère $V_n \approx \mathcal{M}_h^M$ est $\sum_{i=n+1}^{r} \sigma_i^2$.

Preuve

Théorème d'Eckart-Young [QMN, §6.3.1.]

Soit $A \in \mathbb{R}^{N \times M}$. On a

$$\|A - \mathbf{\Phi}_n \mathbf{\Phi}_n^T A\|_F = \min_{B \in \mathbb{R}^{N \times M}, \mathrm{rang}(B) \ \leq n} \|A - B\|_F = \sqrt{\sum_{i=n+1}^{r} \sigma_i^2}.$$

Preuve

Théorème d'Eckart-Young [QMN, §6.3.1.]

Soit $A \in \mathbb{R}^{N \times M}$. On a

$$\|A - \mathbf{\Phi}_n \mathbf{\Phi}_n^T A\|_F = \min_{B \in \mathbb{R}^{N \times M}, \mathrm{rang}(B) \ \leq n} \|A - B\|_F = \sqrt{\sum_{i=n+1}^{r} \sigma_i^2}.$$

Eléments de preuve (Optimalité). On a

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathbf{\Phi}_n \mathbf{\Phi}_n^T \boldsymbol{u}_h(\mu_i)\|_2^2 = \|A - \mathbf{\Phi}_n \mathbf{\Phi}_n^T A\|_F^2$$

Preuve

Théorème d'Eckart-Young [QMN, §6.3.1.]

Soit $A \in \mathbb{R}^{N \times M}$. On a

$$\|A - \mathbf{\Phi}_n \mathbf{\Phi}_n^T A\|_F = \min_{B \in \mathbb{R}^{N \times M}, \mathrm{rang}(B) \ \leq n} \|A - B\|_F = \sqrt{\sum_{i=n+1}^{r} \sigma_i^2}.$$

Eléments de preuve (Optimalité). On a

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathbf{\Phi}_n \mathbf{\Phi}_n^T \boldsymbol{u}_h(\mu_i)\|_2^2 = \|A - \mathbf{\Phi}_n \mathbf{\Phi}_n^T A\|_F^2$$

Par Eckart Young, on a pour tout B de rang n

$$\|A - \mathbf{\Phi}_n \mathbf{\Phi}_n^T A\|_F^2 \leq \|A - B\|_F^2$$

Preuve

Théorème d'Eckart-Young [QMN, §6.3.1.]

Soit $A \in \mathbb{R}^{N \times M}$. On a

$$\|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F = \min_{B \in \mathbb{R}^{N \times M}, \mathrm{rang}(B) \ \leq n} \|A - B\|_F = \sqrt{\sum_{i=n+1}^{r} \sigma_i^2}.$$

Eléments de preuve (Optimalité). On a

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T \boldsymbol{u}_h(\mu_i)\|_2^2 = \|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F^2$$

Par Eckart Young, on a pour tout B de rang n

$$\|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F^2 \leq \|A - B\|_F^2$$

en particulier si $B = \boldsymbol{W}\boldsymbol{W}^T A$ avec $W \in \mathcal{V}_n$

$$\|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F^2 \leq \|A - \boldsymbol{W}\boldsymbol{W}^T A\|_F^2$$

Preuve

Théorème d'Eckart-Young [QMN, §6.3.1.]

Soit $A \in \mathbb{R}^{N \times M}$. On a

$$\|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F = \min_{B \in \mathbb{R}^{N \times M}, \mathrm{rang}(B) \ \leq n} \|A - B\|_F = \sqrt{\sum_{i=n+1}^{r} \sigma_i^2}.$$

Eléments de preuve (Optimalité). On a

$$\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T \boldsymbol{u}_h(\mu_i)\|_2^2 = \|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F^2$$

Par Eckart Young, on a pour tout B de rang n

$$\|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F^2 \leq \|A - B\|_F^2$$

en particulier si $B = \boldsymbol{W}\boldsymbol{W}^T A$ avec $W \in \mathcal{V}_n$

$$\|A - \boldsymbol{\Phi}_n \boldsymbol{\Phi}_n^T A\|_F^2 \leq \|A - \boldsymbol{W}\boldsymbol{W}^T A\|_F^2$$

On retrouve (12) en passant au min sur $\mathcal{V}_n$.

Illustration sur les cas tests

1. L'efficacité de la SVD est liée aux valeurs singulières de A.
2. Si elles décroissent vite, un espace V_n de petite dimension n suffit pour bien approcher $\mathcal{M}_h^M$!

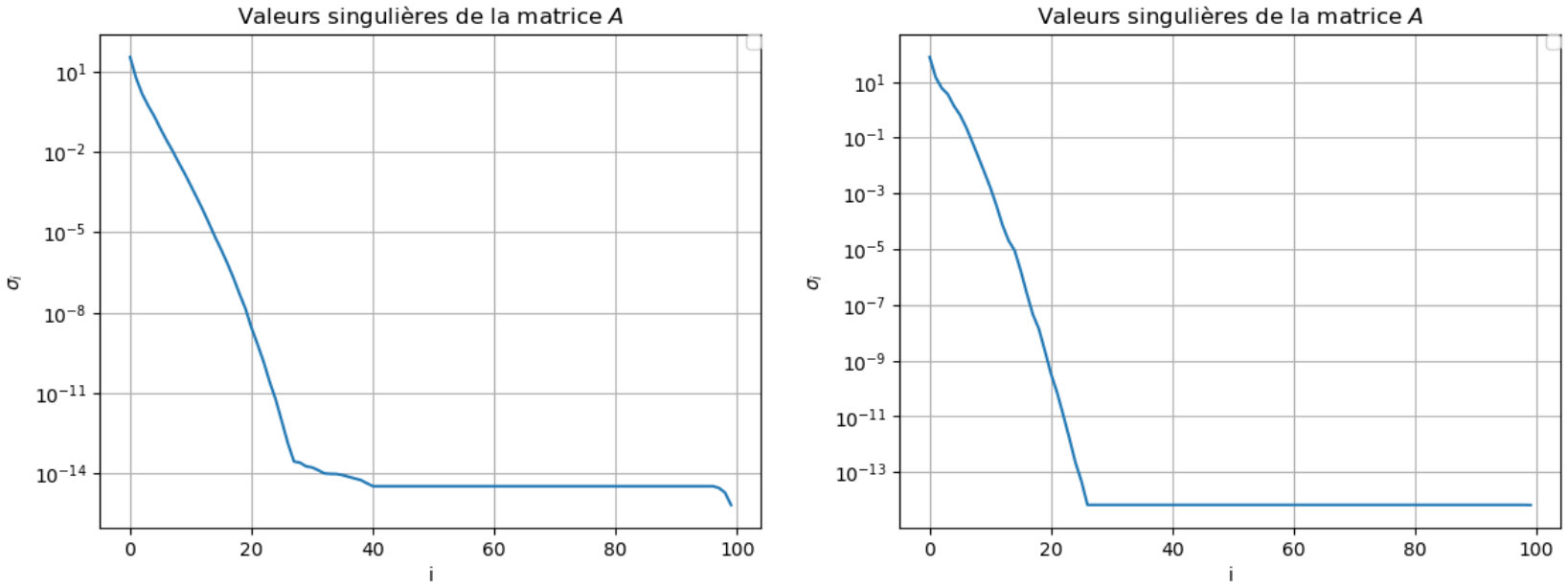


Figure – Valeurs singulières de A pour l'équation d'advection diffusion (gauche) et de la chaleur (droite).

Optimalité en norme 2

$$d_2^M(V_n, \mathcal{M}_h^M) := \left(\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 \right)^{1/2}$$

Optimalité en norme 2

$$d_2^M(V_n, \mathcal{M}_h^M) := \left(\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 \right)^{1/2}$$

Optimalité en norme ∞

On définit la n-**épaisseur de Kolmogorov**

$$d_\infty^M(V_n, \mathcal{M}_h^M) := \min_{V \subset \mathbb{R}^N, \dim V = n} \max_{i=1,M} \|\boldsymbol{u}_h(\mu_i) - P_V \boldsymbol{u}_h(\mu_i)\|_2 \tag{13}$$

Optimalité en norme 2

$$d_2^M(V_n, \mathcal{M}_h^M) := \left(\sum_{i=1}^{M} \|\boldsymbol{u}_h(\mu_i) - \mathrm{P}_{V_n} \boldsymbol{u}_h(\mu_i)\|_2^2 \right)^{1/2}$$

Optimalité en norme ∞

On définit la n-**épaisseur de Kolmogorov**

$$d_\infty^M(V_n, \mathcal{M}_h^M) := \min_{V \subset \mathbb{R}^N, \dim V = n} \max_{i=1,M} \|\boldsymbol{u}_h(\mu_i) - P_V \boldsymbol{u}_h(\mu_i)\|_2 \tag{13}$$

- ☞ Par définition des normes 2 et ∞ sur $\mathbb{R}^M$

$$M^{-1/2} d_2^M(V_n, \mathcal{M}_h^M) \leq d_\infty^M(V_n, \mathcal{M}_h^M) \leq d_2^M(V_n, \mathcal{M}_h^M).$$

- ☞ Résoudre le problème (13) n'est pas trivial en pratique.

Algorithme glouton

Principe. On construit $\{V_\ell\}_{\ell=0}^n \subset \mathbb{R}^N$ t.q. $V_{\ell-1} \subset V_\ell$ de façon **adaptative**.

Soit $V_0 = \{0\}$. Pour $\ell = 1, \ldots, n$ on construit V_ℓ comme il suit.

1) Sélectionner
$$\mu_\ell \in \arg\max_{i=1,\ldots,M} \|\boldsymbol{u}_h(\mu_i) - P_{V_{\ell-1}} \boldsymbol{u}_h(\mu_i)\|_2..$$
2) Calculer le snapshot $\boldsymbol{u}_h(\mu_\ell)$ solution de $F_h(\boldsymbol{u}_h(\mu_\ell), \mu_\ell) = 0$.
3) Poser $V_\ell = \mathrm{vect}\{\boldsymbol{u}_h(\mu_1), \ldots, \boldsymbol{u}_h(\mu_\ell)\}$.

Algorithme glouton

Principe. On construit $\{V_\ell\}_{\ell=0}^n \subset \mathbb{R}^N$ t.q. $V_{\ell-1} \subset V_\ell$ de façon **adaptative**.

Soit $V_0 = \{0\}$. Pour $\ell = 1, \ldots, n$ on construit V_ℓ comme il suit.

1) Sélectionner
$$\mu_\ell \in \arg\max_{i=1,\ldots,M} \|\boldsymbol{u}_h(\mu_i) - P_{V_{\ell-1}} \boldsymbol{u}_h(\mu_i)\|_2..$$
2) Calculer le snapshot $\boldsymbol{u}_h(\mu_\ell)$ solution de $F_h(\boldsymbol{u}_h(\mu_\ell), \mu_\ell) = 0$.
3) Poser $V_\ell = \mathrm{vect}\{\boldsymbol{u}_h(\mu_1), \ldots, \boldsymbol{u}_h(\mu_\ell)\}$.

✗ Calculer l'erreur de projection

$$\|\boldsymbol{u}_h(\mu_i) - P_{V_{\ell-1}} \boldsymbol{u}_h(\mu_i)\|_2$$

nécessite de calculer $\boldsymbol{u}_h(\mu_i)$ pour tout $i = 1, \ldots, M$.

Algorithme glouton

Principe. On construit $\{V_\ell\}_{\ell=0}^n \subset \mathbb{R}^N$ t.q. $V_{\ell-1} \subset V_\ell$ de façon **adaptative**.

Soit $V_0 = \{0\}$. Pour $\ell = 1, \ldots, n$ on construit V_ℓ comme il suit.

1) Sélectionner
$$\mu_\ell \in \arg \max_{i=1,\ldots,M} \Delta(u_{\ell-1}(\xi), \xi).$$
2) Calculer le snapshot $\boldsymbol{u}_h(\mu_\ell)$ solution de $F_h(\boldsymbol{u}_h(\mu_\ell), \mu_\ell) = 0$.
3) Poser $V_\ell = \mathrm{vect}\{\boldsymbol{u}_h(\mu_1), \ldots, \boldsymbol{u}_h(\mu_\ell)\}$.

✗ Calculer l'erreur de projection
$$\|\boldsymbol{u}_h(\mu_i) - P_{V_{\ell-1}} \boldsymbol{u}_h(\mu_i)\|_2$$
nécessite de calculer $\boldsymbol{u}_h(\mu_i)$ pour tout $i = 1, \ldots, M$.

✓ On utilise un estimateur d'erreur e.g.
$$\Delta(u_\ell(\mu_i), \mu_i) := \|L_h(\mu_i)\boldsymbol{u}_\ell(\mu_i) - \boldsymbol{f}_h(\mu_i)\|_2$$
ne nécessite pas de connaitre $\boldsymbol{u}_h(\mu_i)$.

Algorithme glouton

Principe. On construit $\{V_\ell\}_{\ell=0}^n \subset \mathbb{R}^N$ t.q. $V_{\ell-1} \subset V_\ell$ de façon **adaptative**.

Soit $V_0 = \{0\}$. Pour $\ell = 1, \ldots, n$ on construit V_ℓ comme il suit.

1) Sélectionner
$$\mu_\ell \in \arg \max_{i=1,\ldots,M} \Delta(u_{\ell-1}(\xi), \xi).$$
2) Calculer le snapshot $\boldsymbol{u}_h(\mu_\ell)$ solution de $F_h(\boldsymbol{u}_h(\mu_\ell), \mu_\ell) = 0$.
3) Poser $V_\ell = \mathrm{vect}\{\boldsymbol{u}_h(\mu_1), \ldots, \boldsymbol{u}_h(\mu_\ell)\}$.

✗ Calculer l'erreur de projection
$$\|\boldsymbol{u}_h(\mu_i) - P_{V_{\ell-1}} \boldsymbol{u}_h(\mu_i)\|_2$$
nécessite de calculer $\boldsymbol{u}_h(\mu_i)$ pour tout $i = 1, \ldots, M$.

✓ On utilise un estimateur d'erreur e.g.
$$\Delta(u_\ell(\mu_i), \mu_i) := \|L_h(\mu_i)\boldsymbol{u}_\ell(\mu_i) - \boldsymbol{f}_h(\mu_i)\|_2$$
ne nécessite pas de connaitre $\boldsymbol{u}_h(\mu_i)$.

✓ On a besoin de n snapshots contre M dans la méthode basée sur la SVD.

Bilan

Les **méthodes d'approximation linéaire** ou **sous format de faible rang**

✓ sont efficaces pour des problèmes pour lesquels n-épaisseur de Kolmogorov décroit rapidement

ex. : Exponentiellement pour des problèmes de type diffusion

Bilan

Les **méthodes d'approximation linéaire** ou **sous format de faible rang**

✓ sont efficaces pour des problèmes pour lesquels n-épaisseur de Kolmogorov décroit rapidement

ex. : Exponentiellement pour des problèmes de type diffusion

✗ mais montrent leur limite quand la n-épaisseur de Kolmogorov décroit lentement !

ex. : Comme $n^{-1/2}$ pour les problèmes de type transport

Ca ne marche pas toujours...

Soit $u : [0,T] \times \mathbb{R} \to \mathbb{R}$ solution de

$$\partial_t u(t,x) + \partial_x u(t,x) \quad = \quad 0, \quad x \in \mathbb{R} \tag{14}$$

avec $u(0,x) = \dfrac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}$ et $\sigma > 0$. Le problème (14) admet pour solution

$$u(t,x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-t)^2}{2\sigma^2}}.$$

Construisons la matrice des snapshots $A \in \mathbb{R}^{N \times M}$ t.q. $A_{ij} = u(t_j, x_i)$.

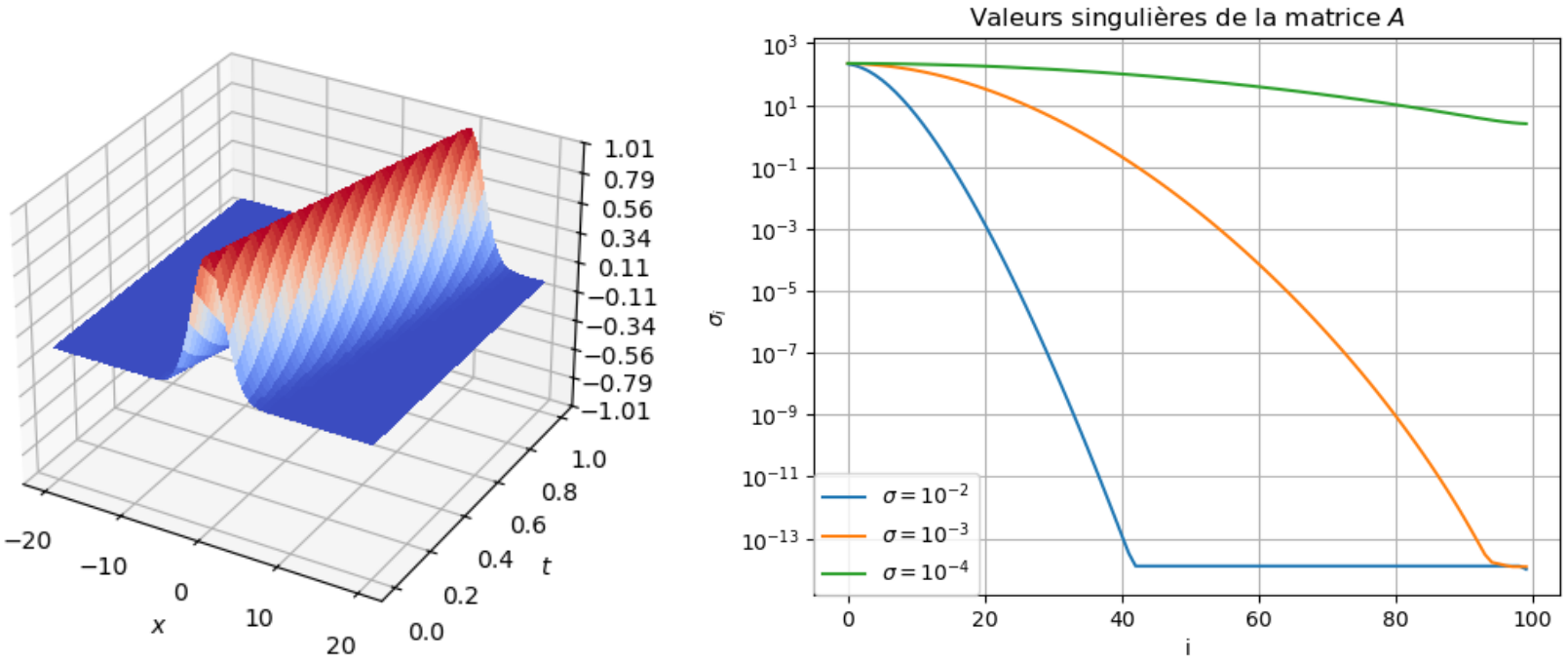


Figure – Gauche : solution u pour $\sigma = 1$. Droite : valeurs singulières de A, $N = 10^4, M = 500$.

Pour aller plus loin

[QMN] Alfio Quarteroni, Andrea Manzoni et Federico Negri. Reduced Basis Methods for Partial Differential Equations. Cham, Switzerland : Springer International Publishing (cf. p. 91, 92, 99–103).