

# Méthodes de discrétisation en fluides

## 8. Systèmes d'équations linéaires

Marc A. Habisreutinger

Ecole Polytechnique Fédérale de Lausanne  
Section de génie mécanique, CH-1015 Lausanne

Jeudi 18 avril 2024

# Contenu

## Introduction

- Equations de Navier–Stokes
- Turbulence en mécanique des fluides
- Simulation numérique directe

## Systèmes d'équations linéaires

- Définitions
- Méthodes de résolution
- Rappels

## Méthodes directes

- Systèmes triangulaires
- Méthode d'élimination de Gauss
- Factorisation LU
- Méthode de Thomas

## Méthodes itératives

- Historique
- Définition
- Méthodes linéaires
- Critères d'arrêt
- Méthodes de Richardson

## Références

# Introduction

## Equations de Navier–Stokes - Formulation dimensionnelle

On considère les équations de Navier–Stokes

$$\rho \underbrace{\left( \frac{\rho V}{T} \right)}_{\left( \frac{\rho V}{T} \right)} \underbrace{\left( \partial_t \mathbf{v} + \mathbf{v} \cdot \nabla \mathbf{v} \right)}_{\left( \frac{\rho V^2}{L} \right)} = - \underbrace{\nabla p}_{\left( \frac{P}{L} \right)} + \underbrace{\mu \nabla^2 \mathbf{v}}_{\left( \frac{\mu V}{L^2} \right)}$$

$$\underbrace{\nabla \cdot \mathbf{v}}_{\left( \frac{V}{L} \right)} = 0$$

$$\mathbf{x} = L \hat{\mathbf{x}}$$

$$t = T \hat{t}$$

$$\mathbf{v} = V \hat{\mathbf{v}}$$

$$p = P \hat{p}$$

En utilisant les variables adimensionnelles, on a

$$\left( \frac{\rho V}{T} \right) \partial_{\hat{t}} \hat{\mathbf{v}} + \left( \frac{\rho V^2}{L} \right) \hat{\mathbf{v}} \cdot \hat{\nabla} \hat{\mathbf{v}} = - \left( \frac{P}{L} \right) \hat{\nabla} \hat{p} + \left( \frac{\mu V}{L^2} \right) \hat{\nabla}^2 \hat{\mathbf{v}}$$

$$\left( \frac{V}{L} \right) \hat{\nabla} \cdot \hat{\mathbf{v}} = 0$$

# Introduction

## Equations de Navier–Stokes - Formulations adimensionnelles

- avec le temps d'advection  $T = \frac{L}{V}$ , on obtient

$$\partial_t \hat{\mathbf{v}} + \hat{\mathbf{v}} \cdot \hat{\nabla} \hat{\mathbf{v}} = -\hat{\nabla} \hat{p} + \text{Re}^{-1} \hat{\nabla}^2 \hat{\mathbf{v}}$$

$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0$$

- avec le temps de diffusion  $T = \frac{\rho L^2}{\mu}$ , on obtient

$$\partial_t \hat{\mathbf{v}} + \text{Re} \hat{\mathbf{v}} \cdot \hat{\nabla} \hat{\mathbf{v}} = -\hat{\nabla} \hat{p} + \hat{\nabla}^2 \hat{\mathbf{v}}$$

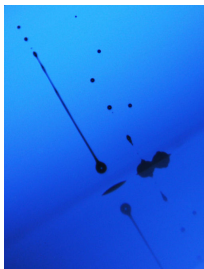
$$\hat{\nabla} \cdot \hat{\mathbf{v}} = 0$$

où le nombre de Reynolds

$$\text{Re} = \frac{\rho V L}{\mu}$$

# Introduction

## Turbulence en mécanique des fluides

 $L$ 
Marmottant *et al.*, 2004
 $L, r$ 


Hoath, 2006

 $L, r_1, r_2$ 
Marmottant *et al.*, 2004
 $L, r_1, r_2, \dots, \lambda$ 


Genève, 1886



# Introduction

## Simulation numérique directe

Lorsqu'on discrétise les équations de Navier–Stokes, on obtient un système d'équations algébriques de taille  $p$

$$\mathbf{A}\mathbf{u} = \mathbf{M}\mathbf{f}$$

- coût de stockage

Matrices pleines (p. ex. méthodes spectrales)

$$\propto p^2 \propto \text{Re}^{18/4}$$

- coût de calcul

Solution avec la méthode d'élimination de Gauss

$$\propto p^3 \propto \text{Re}^{27/4}$$

# Introduction

## Simulation numérique directe - Défis

### 1. Diminution du coût de stockage

- Stockage morse
- Matrices non construites (on calcule seulement l'action d'une matrice sur un vecteur)

### 2. Diminution du coût de calcul

- Méthodes de résolution de systèmes algébriques
- Méthodes de discrétisation spatiale, et d'intégration temporelle

### 3. Diminution du temps de calcul

- Parallélisation des algorithmes
- Amélioration des calculateurs

### 4. Modélisation de la turbulence<sup>1</sup>

- Reynolds-Averaged Numerical Simulation (RANS)
- Large-Eddy Simulation (LES)

<sup>1</sup>i.e. diminution du nombre d'échelles calculées



# Systèmes d'équations linéaires

## Définitions

On considère le système de  $p$  équations à  $p$  inconnues de la forme

$$\sum_{j=1}^p A_{ij}x_j = b_i, \quad i = 1, \dots, p$$

avec les coefficients  $A_{ij}$  et  $b_i$  constants. On peut écrire ce système sous la forme

$$\mathbf{Ax} = \mathbf{b}$$

avec

$$\mathbf{A} \in \mathbb{R}^{p \times p}$$

$$\mathbf{x}, \mathbf{b} \in \mathbb{R}^p$$

# Systèmes d'équations linéaires

## Définitions

La solution  $\mathbf{x}$  existe ssi ces trois conditions équivalentes sont satisfaites

- $\mathbf{A}$  est inversible, *i.e.*

$$\det(\mathbf{A}) \neq 0$$

- Les vecteurs colonne de  $\mathbf{A}$  sont linéairement indépendants, *i.e.*

$$\text{rk}(\mathbf{A}) = p$$

- Le système homogène admet seulement la solution nulle, *i.e.*

$$\ker(\mathbf{A}) = \{\mathbf{x} \in \mathbb{R}^p \mid \mathbf{A}\mathbf{x} = \mathbf{0}\} = \mathbf{0}$$

# Systèmes d'équations linéaires

## Méthodes de résolution

### Méthodes directes

- Solution en un nombre fini d'étapes
- Complexité généralement  $\propto p^3$

### Méthodes itératives

- Solution en un nombre infini d'étapes
- Complexité par étape généralement  $\propto p^2$
- Utilité si le nombre d'étapes pour la convergence  $< p$

# Systèmes d'équations linéaires

## Rappels

On rappelle les définitions suivantes

- Valeurs et vecteurs propres de  $\mathbf{A}$

$$\mathbf{A}\widehat{\mathbf{x}}^{(k)} = \lambda^{(k)}\widehat{\mathbf{x}}^{(k)}, \quad k = 1, \dots, p$$

- Rayon spectral  $\mathbf{A}$

$$\rho(\mathbf{A}) = \max_k |\lambda^{(k)}|, \quad k = 1, \dots, p$$

# Systèmes d'équations linéaires

## Rappels

ainsi que les propriétés

- $\mathbf{A}$  est définie positive si

$$\mathbf{v}^T \mathbf{A} \mathbf{v} > 0, \quad \forall \mathbf{v} \in \mathbb{R}^p, \mathbf{v} \neq \mathbf{0}$$

- $\mathbf{A}$  est diagonale dominante stricte par ligne si

$$|A_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^p |A_{ij}|, \quad \forall i$$

- $\mathbf{A}$  est symétrique si

$$A_{ij} = A_{ji}, \quad \forall i, j$$

# Méthodes directes

## Systèmes triangulaires inférieurs

On considère le système triangulaire inférieur

$$\begin{pmatrix} L_{11} & 0 & 0 \\ L_{21} & L_{22} & 0 \\ L_{31} & L_{32} & L_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$$

Puisque le système est par hypothèse inversible, on a  $L_{ii} \neq 0$ . Ainsi, la solution s'écrit

$$\begin{aligned} x_1 &= \frac{b_1}{L_{11}} \\ x_2 &= \frac{1}{L_{22}}(b_2 - L_{21}x_1) \\ x_3 &= \frac{1}{L_{33}}(b_3 - L_{31}x_1 - L_{32}x_2) \end{aligned}$$

# Méthodes directes

## Systèmes triangulaires inférieurs

De manière générale, la solution du système triangulaire inférieur

$$\mathbf{L}\mathbf{x} = \mathbf{b}$$

est donnée par l'algorithme de substitution directe

$$x_1 = \frac{b_1}{L_{11}}$$

$$x_i = \frac{1}{L_{ii}} \left( b_i - \sum_{j=1}^{i-1} L_{ij} x_j \right), \quad i = 2, \dots, p$$

dont la complexité algorithmique est  $\propto p^2$  à cause de la somme allant de 1 jusqu'à  $j - 1$ .

# Méthodes directes

## Systèmes triangulaires supérieurs

De manière analogue, la solution du système triangulaire supérieur

$$\mathbf{U}\mathbf{x} = \mathbf{b}$$

est donnée par l'algorithme de substitution inverse

$$x_p = \frac{b_p}{U_{pp}}$$

$$x_i = \frac{1}{U_{ii}}(b_i - \sum_{j=i+1}^p U_{ij}x_j), \quad i = p-1, \dots, 1$$

dont la complexité algorithmique est  $\propto p^2$  à cause de la somme allant de  $i+1$  jusqu'à  $p$ .



# Méthodes directes

$$\mathbf{Ax} = \mathbf{b}$$

## Méthode d'élimination de Gauss

Pour une matrice  $\mathbf{A}$  non-triangulaire, on construit les suites de matrices et de seconds membres

$$\begin{aligned}\mathbb{A} &= \{\overbrace{\mathbf{A}^{(1)}}^{\equiv \mathbf{A}}, \mathbf{A}^{(2)}, \dots, \mathbf{A}^{(k)}, \dots, \overbrace{\mathbf{A}^{(p)}}^{\equiv \mathbf{U}}\} \\ \mathbb{B} &= \{\overbrace{\mathbf{b}^{(1)}}^{\equiv \mathbf{b}}, \mathbf{b}^{(2)}, \dots, \mathbf{b}^{(k)}, \dots, \overbrace{\mathbf{b}^{(p)}}^{\equiv \hat{\mathbf{b}}}\}\end{aligned}$$

telles que le système devienne triangulaire supérieur à l'étape  $p$

$$\mathbf{Ux} = \hat{\mathbf{b}}$$

et puisse être résolu par substitution inverse. La complexité totale est  $\propto p^3$ .

# Méthodes directes

## Méthode d'élimination de Gauss

On a donc par définition

$$\mathbf{A}^{(1)} = \mathbf{A}, \quad \mathbf{b}^{(1)} = \mathbf{b}$$

Les seconds termes de la suite sont définis par

$$A_{ij}^{(2)} = A_{ij}^{(1)} - m_i^{(1)} A_{1j}^{(1)}, \quad i, j = 2, \dots, p$$

$$b_i^{(2)} = b_i^{(1)} - m_i^{(1)} b_1^{(1)}, \quad i = 2, \dots, p$$

où le multiplicateur

$$m_i^{(1)} = \frac{A_{i1}^{(1)}}{A_{11}^{(1)}}, \quad i = 2, \dots, p$$

# Méthodes directes

## Méthode d'élimination de Gauss

Sous forme matricielle, les seconds termes s'écrivent

$$\mathbf{A}^{(2)} = \begin{pmatrix} A_{11}^{(1)} & A_{12}^{(1)} & \dots & A_{1p}^{(1)} \\ 0 & A_{22}^{(2)} & \dots & A_{2p}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & A_{p2}^{(2)} & \dots & A_{pp}^{(2)} \end{pmatrix}, \quad \mathbf{b}^{(2)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_p^{(2)} \end{pmatrix}$$

On a donc éliminé la partie inférieure de la première colonne de  $\mathbf{A}$ .

# Méthodes directes

## Méthode d'élimination de Gauss

On définit ainsi la suite de systèmes

$$\mathbf{A}^{(k)} \mathbf{x} = \mathbf{b}^{(k)}, \quad 1 \leq k \leq p$$

qui s'obtient par récurrence avec

$$A_{ij}^{(k+1)} = A_{ij}^{(k)} - m_i^{(k)} A_{kj}^{(k)}, \quad i, j = k+1, \dots, p$$

$$b_i^{(k+1)} = b_i^{(k)} - m_i^{(k)} b_k^{(k)}, \quad i = k+1, \dots, p$$

où le multiplicateur

$$m_i^{(k)} = \frac{A_{ik}^{(k)}}{A_{kk}^{(k)}}, \quad i = k+1, \dots, p$$

# Méthodes directes

## Méthode d'élimination de Gauss

Sous forme matricielle, les  $k$ -ièmes termes s'écrivent

$$\mathbf{A}^{(k)} = \begin{pmatrix} A_{11}^{(1)} & A_{12}^{(1)} & \dots & \dots & \dots & A_{1p}^{(1)} \\ 0 & A_{22}^{(2)} & & & & A_{2p}^{(2)} \\ \vdots & & \ddots & & & \vdots \\ 0 & \dots & 0 & A_{kk}^{(k)} & \dots & A_{kp}^{(k)} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & A_{pk}^{(k)} & \dots & A_{pp}^{(k)} \end{pmatrix}, \quad \mathbf{b}^{(k)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_k^{(k)} \\ \vdots \\ b_p^{(k)} \end{pmatrix}$$

On a donc éliminé la partie inférieure des  $k - 1$  premières colonne de  $\mathbf{A}$ .

# Méthodes directes

$$\mathbf{Ax} = \mathbf{b}$$

$$\mathbf{Ux} = \hat{\mathbf{b}}$$

## Méthode d'élimination de Gauss

A l'étape  $p$ , on obtient donc le système triangulaire supérieur recherché

$$\mathbf{A}^{(p)} = \mathbf{U}, \quad \mathbf{b}^{(p)} = \hat{\mathbf{b}}$$

qui s'écrit explicitement sous la forme

$$\mathbf{U} = \begin{pmatrix} A_{11}^{(1)} & A_{12}^{(1)} & \cdots & \cdots & A_{1p}^{(1)} \\ 0 & A_{22}^{(2)} & & & A_{2p}^{(2)} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & A_{pp}^{(p)} \end{pmatrix}, \quad \hat{\mathbf{b}} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ \vdots \\ b_p^{(p)} \end{pmatrix}$$

et qui peut être résolu par substitution inverse.

# Méthodes directes

## Méthode d'élimination de Gauss

La méthode de Gauss est applicable ssi les pivots

$$A_{kk}^{(k)} \neq 0$$

Cette condition est satisfaite si

- $\mathbf{A}$  est diagonale dominante par ligne
- $\mathbf{A}$  est diagonale dominante par colonne
- $\mathbf{A}$  est symétrique et définie positive

Dans les autres cas, on peut utiliser la méthode de Gauss avec changement de pivot.

# Méthodes directes

## Factorisation LU

En factorisant la matrice **A** en partie triangulaire inférieure **L** et supérieure **U**

$$\mathbf{Ax} = \mathbf{L} \underbrace{\mathbf{Ux}}_{=\mathbf{y}} = \mathbf{b}$$

la résolution du système d'équations se ramène à la résolution successive des deux systèmes triangulaires

$$\mathbf{Ly} = \mathbf{b}$$

$$\mathbf{Ux} = \mathbf{y}$$

par substitution directe et inverse.



# Méthodes directes

## Factorisation LU

Les matrices triangulaires sont composées des multiplicateurs, par exemple

$$\mathbf{L} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ m_2^{(1)} & 1 & & & \vdots \\ \vdots & & m_3^{(2)} & \ddots & \vdots \\ \vdots & \vdots & & \ddots & 0 \\ m_p^{(1)} & m_p^{(2)} & \dots & m_p^{(p-1)} & 1 \end{pmatrix}$$

Cette méthode est ainsi équivalente à celle de Gauss en termes de complexité algorithmique.

# Méthodes directes

## Méthode de Thomas

Dans le cas où la matrice  $\mathbf{A}$  est tridiagonale, *i.e.*

$$\mathbf{A} = \begin{pmatrix} \alpha_1 & \gamma_1 & & & \\ \beta_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \beta_p & \gamma_{p-1} & \alpha_p \end{pmatrix}$$

la factorisation LU prend la forme

$$\mathbf{L} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ L_2 & & \ddots & & \\ & & \ddots & \ddots & \\ & & & L_p & 1 \end{pmatrix}, \quad \mathbf{U} = \begin{pmatrix} U_1 & \gamma_1 & & & \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ & & & \gamma_{p-1} & U_p \end{pmatrix}$$

# Méthodes directes

## Méthode de Thomas

Pour la matrice triangulaire supérieure, on a

$$U_1 = \alpha_1$$

$$U_i = \alpha_i - L_i \gamma_{i-1}, \quad i = 2, \dots, p$$

et pour la matrice triangulaire inférieure

$$L_i = \frac{\beta_i}{U_{i-1}}, \quad i = 2, \dots, p$$

# Méthodes directes

$$Ly = b$$

$$Ux = y$$

## Méthode de Thomas

Pour la résolution des systèmes triangulaires, on applique les relations de substitution directe simplifiées

$$y_1 = b_1, \quad y_i = b_i - L_i y_{i-1}, \quad i = 2, \dots, p$$

puis inverse

$$x_p = \frac{y_p}{U_p}, \quad x_i = \frac{1}{U_i} (y_i - \gamma_i x_{i+1}), \quad i = p-1, \dots, 1$$

dont la complexité algorithmique est  $\propto p$

# Méthodes itératives

## Historique

Pour résoudre l'équation de diffusion

$$\begin{cases} \nabla^2 u = f, & \mathbf{x} \in \Omega \subset \mathbb{R}^3 \\ u = 0 & \text{sur } \partial\Omega \end{cases}$$

avec des différences finies centrées du second ordre et  $p = p_i^3$  points de colocation, voici quelques algorithmes ainsi que leur complexité

Elimination de Gauss (1947)

$$\propto p_i^7$$

Gauss-Seidel sous-optimal (1954)

$$\propto 8p_i^5$$

Gauss-Seidel optimal (1960)

$$\propto 8p_i^4 \log(p_i)$$

Multigrille (1981)

$$\propto 30p_i^3 \log(p_i)$$

# Méthodes itératives

$$\mathbf{Ax} = \mathbf{b}$$

## Définition

A la place de construire une suite de systèmes équivalents (méthode directe), pour une méthode itérative, on construit la suite de vecteurs

$$\mathbb{X} = \{\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \dots, \mathbf{x}^{(k)}, \dots\}$$

Cette suite est dite convergente si

$$\lim_{k \rightarrow \infty} \mathbf{x}^{(k)} = \mathbf{x}, \quad \forall \mathbf{x}^{(0)}$$

# Méthodes itératives

$$\mathbf{Ax} = \mathbf{b}$$

## Définition

On définit l'erreur à l'itération  $k$

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$$

- L'erreur ne peut pas être calculée dans la pratique puisqu'on ne connaît pas *a priori* la solution du système.
- L'erreur ne doit pas être confondue avec le résidu

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{Ax}^{(k)}$$

qui peut être calculé à chaque itération.

# Méthodes itératives

$$\mathbf{Ax} = \mathbf{b}$$

## Définition

On considère la méthode itérative suivante

$$\begin{aligned} \mathbf{x}^{(0)} & \text{ donné} \\ \mathbf{x}^{(k+1)} &= \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}, \quad k > 0 \end{aligned}$$

avec  $\mathbf{B}$  la matrice d'itération. La méthode est consistante si

$$\mathbf{x} = \mathbf{B}\mathbf{x} + \mathbf{f}$$

ce qui peut s'écrire sous la forme

$$\mathbf{f} = (\mathbf{I} - \mathbf{B})\mathbf{x}$$



# Méthodes itératives

## Définition

$$\begin{aligned}\mathbf{e}^{(k)} &= \mathbf{x}^{(k)} - \mathbf{x} \\ \mathbf{x}^{(k+1)} &= \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} \\ \mathbf{f} &= (\mathbf{I} - \mathbf{B})\mathbf{x}\end{aligned}$$

En considérant une méthode consistante, l'erreur à l'itération  $k + 1$  s'écrit

$$\begin{aligned}\mathbf{e}^{(k+1)} &= \mathbf{x}^{(k+1)} - \mathbf{x} \\ &= \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} - \mathbf{x} \\ &= \mathbf{B}\mathbf{x}^{(k)} + (\mathbf{I} - \mathbf{B})\mathbf{x} - \mathbf{x} \\ &= \mathbf{B}\mathbf{e}^{(k)}\end{aligned}$$

En procédant par récurrence, on obtient facilement

$$\boxed{\mathbf{e}^{(k+1)} = \mathbf{B}^k \mathbf{e}^{(0)}}$$

# Méthodes itératives

## Définition

$$\begin{aligned}\mathbf{x}^{(k+1)} &= \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f} \\ \mathbf{e}^{(k+1)} &= \mathbf{B}^k \mathbf{e}^{(0)}\end{aligned}$$

Effectuons le changement de variable suivant

$$\underbrace{\mathbf{V}\boldsymbol{\epsilon}^{(k+1)}}_{= \mathbf{e}^{(k+1)}} = \mathbf{B} \underbrace{\mathbf{V}\boldsymbol{\epsilon}^{(k)}}_{= \mathbf{e}^{(k)}}$$

Si on choisit  $\mathbf{V}$  telle qu'elle diagonalise  $\mathbf{B}$ , on a le système découplé

$$\boldsymbol{\epsilon}^{(k+1)} = \mathbf{V}^{-1}\mathbf{B}\mathbf{V}\boldsymbol{\epsilon}^{(k)} = \boldsymbol{\Lambda}\boldsymbol{\epsilon}^{(k)}$$

où  $\boldsymbol{\Lambda}$  est la matrice des valeurs propres de  $\mathbf{B}$ . On a donc par récurrence et sous forme indicielle

$$\boxed{\epsilon_i^{(k+1)} = \lambda_i^k \epsilon_i^{(0)}}$$

# Méthodes itératives

## Définition

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$$

$$\mathbf{e}^{(k+1)} = \mathbf{B}^k \mathbf{e}^{(0)}$$

Ainsi, la condition sur le rayon spectral de la matrice d'itération

$$\rho(\mathbf{B}) < 1$$

implique la convergence

$$\lim_{k \rightarrow \infty} \mathbf{e}^{(k+1)} = \lim_{k \rightarrow \infty} \mathbf{B}^k \mathbf{e}^{(0)} = \mathbf{0}, \quad \forall \mathbf{x}^{(0)}$$

le taux de convergence asymptotique étant donné par

$$\kappa = -\log \rho(\mathbf{B})$$

# Méthodes itératives

## Définition

De manière générale, les méthodes itératives peuvent s'écrire sous la forme

$$\begin{aligned} \mathbf{x}^{(0)} &= \mathbf{g}^{(0)}(\mathbf{A}, \mathbf{b}) \\ \mathbf{x}^{(k+1)} &= \mathbf{g}^{(k)}(\mathbf{A}, \mathbf{b}, \mathbf{x}^{(k)}, \mathbf{x}^{(k-1)}, \dots, \mathbf{x}^{(k-s)}), \quad k \geq s \end{aligned}$$

On définit les propriétés suivantes

- $s$  est l'ordre de la méthode
- Si  $\mathbf{g}^{(k)}$  est indépendant de  $k$ , la méthode est dite stationnaire
- Si  $\mathbf{g}^{(k)}$  dépend linéairement des  $\mathbf{x}^{(i)}$ , la méthode est linéaire

# Méthodes itératives

## Méthodes linéaires

$$\mathbf{e}^{(k)} = \mathbf{x}^{(k)} - \mathbf{x}$$

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}$$

En décomposant la matrice du système sous la forme

$$\mathbf{A} = \mathbf{P} - \mathbf{N}$$

et en définissant la méthode itérative linéaire comme

$$\mathbf{x}^{(0)} \text{ donné}$$

$$\mathbf{P}\mathbf{x}^{(k+1)} = \mathbf{N}\mathbf{x}^{(k)} + \mathbf{b}, \quad k \geq 0$$

la consistance est garantie et on a

$$\mathbf{B} = \mathbf{P}^{-1}\mathbf{N} = \mathbf{I} - \mathbf{P}^{-1}\mathbf{A}$$

$$\mathbf{f} = \mathbf{P}^{-1}\mathbf{b}$$

# Méthodes itératives

## Méthodes linéaires

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ (\mathbf{I} - \mathbf{B})\mathbf{x} &= \mathbf{f} \\ \mathbf{B} &= \mathbf{I} - \mathbf{P}^{-1}\mathbf{A} \\ \mathbf{f} &= \mathbf{P}^{-1}\mathbf{b} \end{aligned}$$

Ainsi, la relation de consistance devient

$$(+\mathbf{I} \underbrace{-\mathbf{I} + \mathbf{P}^{-1}\mathbf{A}}_{= -\mathbf{B}})\mathbf{x} = \underbrace{\mathbf{P}^{-1}\mathbf{b}}_{= \mathbf{f}}$$

et se réduit sous la forme du système préconditionné

$$\boxed{\mathbf{P}^{-1}\mathbf{Ax} = \mathbf{P}^{-1}\mathbf{b}}$$

On résoud donc le système préconditionné dont la solution est identique à celle du système original pourvu que le préconditionneur à gauche  $\mathbf{P}$  soit inversible.

# Méthodes itératives

## Méthodes linéaires

$$\mathbf{P}\mathbf{x}^{(k+1)} = \mathbf{N}\mathbf{x}^{(k)} + \mathbf{b}$$

$$\mathbf{x}^{(k+1)} = \mathbf{B}\mathbf{x}^{(k)} + \mathbf{f}$$

$$\mathbf{A} = \mathbf{P} - \mathbf{N}$$

En utilisant le résidu  $\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$ , on peut aussi écrire la méthode sous la forme

$$\mathbf{P}\mathbf{x}^{(k+1)} = \underbrace{(\mathbf{A} + \mathbf{N})}_{=\mathbf{P}}\mathbf{x}^{(k)} + \mathbf{r}^{(k)}$$

Puis, en divisant par  $\mathbf{P}$ , on obtient

$$\boxed{\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{P}^{-1}\mathbf{r}^{(k)}}$$

On définit en outre la direction de descente

$$\mathbf{z}^{(k)} = \mathbf{P}^{-1}\mathbf{r}^{(k)}$$

# Méthodes itératives

## Méthodes linéaires

Par exemple, on peut choisir naïvement

$$\mathbf{P} = \mathbf{A}$$

Dans ce cas, l'itération s'écrit

$$\begin{aligned}\mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \mathbf{A}^{-1}\mathbf{r}^{(k)} \\ &= \mathbf{x}^{(k)} + \mathbf{A}^{-1}\mathbf{b} - \mathbf{x}^{(k)} = \mathbf{b}, \quad \forall \mathbf{x}^{(k)}\end{aligned}$$

- On a convergence en une seule itération mais le coût est celui du problème initial puisqu'on doit résoudre  $\mathbf{A}\mathbf{z}^{(k)} = \mathbf{r}^{(k)}$ .
- Il faut choisir un préconditionneur plus facile à inverser et qui permette quand même de converger en peu d'itérations.



# Méthodes itératives

## Méthodes linéaire de Jacobi

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{P}^{-1} \mathbf{r}^{(k)}$$

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A} \mathbf{x}^{(k)}$$

Pour la méthode de Jacobi, le préconditionneur est choisi comme la diagonale de  $\mathbf{A}$

$$\mathbf{P} \equiv \omega^{-1} \mathbf{D}, \quad D_{ij} = A_{ij} \delta_{ij}$$

où  $\omega$  est un paramètre de sous- ou sur-relaxation qui permet de modifier les propriétés de convergence. La méthode de Jacobi est donc définie par l'itération

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \underbrace{(\omega^{-1} \mathbf{D})^{-1}}_{\equiv \mathbf{P}} \mathbf{r}^{(k)}$$

Le coût de l'itération de Jacobi est donc  $\propto p^1$ .

# Méthodes itératives

## Méthodes linéaire de Gauss-Seidel

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{P}^{-1} \mathbf{r}^{(k)}$$

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A} \mathbf{x}^{(k)}$$

Pour la méthode de Gauss-Seidel, le préconditionneur est choisi comme la partie triangulaire inférieure de  $\mathbf{A}$

$$\mathbf{P} \equiv \omega^{-1} \mathbf{D} - \mathbf{E}, \quad E_{ij} = \begin{cases} -A_{ij} & i > j \\ 0 & i \leq j \end{cases}$$

La méthode de Gauss-Seidel est donc définie par l'itération

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \underbrace{(\omega^{-1} \mathbf{D} - \mathbf{E})^{-1}}_{\equiv \mathbf{P}} \mathbf{r}^{(k)}$$

Le coût de l'itération de Gauss-Seidel est donc  $\propto p^2$ . On espère donc faire moins d'itérations qu'avec la méthode de Jacobi.

# Méthodes itératives

## Méthodes linéaires - Résultats de convergence

- Pour  $\mathbf{A}$  diagonale dominante stricte et  $\omega = 1$ , les méthodes de Jacobi et de Gauss-Seidel sont convergentes
- Pour  $\mathbf{A}$  symétrique et définie positive, la méthode de Jacobi converge si

$$0 < \omega < \frac{2}{\rho(\mathbf{D}^{-1}\mathbf{A})}$$

- Pour  $\mathbf{A}$  tridiagonale, symétrique et définie positive, on a

$$\rho(\mathbf{B}_{\text{GS}}) = \rho^2(\mathbf{B}_{\text{J}})$$

c'est-à-dire que la méthode de Gauss-Seidel converge plus rapidement que la méthode de Jacobi

# Méthodes itératives

## Méthodes linéaires - Résultats de convergence

- Sans hypothèses sur  $\mathbf{A}$ , les méthodes de Jacobi et de Gauss-Seidel ne peuvent être convergentes que pour

$$0 < \omega < 2$$

- Pour  $\mathbf{A}$  diagonale dominante stricte, la convergence de la méthode de Gauss-Seidel est assurée si

$$0 < \omega < 1$$

# Méthodes itératives

## Critères d'arrêt

$$\mathbf{r}^{(k)} = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)}$$

$$\mathbf{z}^{(k)} = \mathbf{P}^{-1}\mathbf{r}^{(k)}$$

- Nombre d'itérations

$$k \geq M$$

où  $M$  est un nombre d'itérations maximal fixé

- Convergence

$$\frac{\|\mathbf{P}^{-1}\mathbf{r}^{(k)}\|}{\|\mathbf{P}^{-1}\mathbf{b}\|} \leq \tau$$

où  $\tau$  est une tolérance fixée

# Méthodes itératives

## Méthodes de Richardson

$$\mathbf{B} = \mathbf{I} - \mathbf{P}^{-1}\mathbf{A}$$

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \mathbf{P}^{-1}\mathbf{r}^{(k)}$$

- Méthode de Richardson stationnaire

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha \mathbf{P}^{-1} \mathbf{r}^{(k)}, \quad k \geq 0$$

- Méthode de Richardson instationnaire

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{P}^{-1} \mathbf{r}^{(k)}, \quad k \geq 0$$

- Matrice d'itération de Richardson

$$\mathbf{R} = \mathbf{I} - \alpha^{(k)} \mathbf{P}^{-1} \mathbf{A}$$

# Méthodes itératives

## Méthodes de Richardson

### 1. Résoudre le système linéaire

$$\mathbf{P}\mathbf{z}^{(k)} = \mathbf{r}^{(k)}$$

### 2. Calculer le paramètre d'accélération

$$\alpha^{(k)} = f(\mathbf{z}^{(k)})$$

### 3. Mettre à jour la solution

$$\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{z}^{(k)}$$

### 4. Mettre à jour le résidu

$$\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha^{(k)}\mathbf{A}\mathbf{z}^{(k)}$$

# Méthodes itératives

$$\mathbf{Ax} = \mathbf{b}$$

## Méthodes de Richardson - Méthode du gradient

On considère la forme quadratique

$$\Phi(\mathbf{y}) = \frac{1}{2}\mathbf{y}^T \mathbf{A} \mathbf{y} - \mathbf{y}^T \mathbf{b}$$

Pour des matrices symétriques, le gradient est donné par

$$\nabla \Phi(\mathbf{y}) = \frac{1}{2}(\mathbf{A}^T + \mathbf{A})\mathbf{y} - \mathbf{b} = \mathbf{A}\mathbf{y} - \mathbf{b}$$

Ainsi

$$\nabla \Phi(\mathbf{y}) = 0 \quad \rightarrow \quad \mathbf{y} = \mathbf{x}$$

On doit donc résoudre un problème de minimisation.



# Méthodes itératives

## Méthodes de Richardson - Méthode du gradient

$$\begin{aligned}\nabla \Phi(\mathbf{y}) &= \mathbf{A}\mathbf{y} - \mathbf{b} \\ \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{P}^{-1} \mathbf{r}^{(k)} \\ \mathbf{z}^{(k)} &= \mathbf{P}^{-1} \mathbf{r}^{(k)}\end{aligned}$$

- Direction de descente

Méthode de la plus forte pente (gradient)

$$\mathbf{z}^{(k)} = -\nabla \Phi(\mathbf{x}^{(k)}) = \mathbf{b} - \mathbf{A}\mathbf{x}^{(k)} = \mathbf{r}^{(k)}$$

$$\mathbf{P} = \mathbf{I}$$

- Paramètre d'accélération

$$\frac{\partial}{\partial \alpha^{(k)}} \Phi(\mathbf{x}^{(k+1)}) = 0$$

→

$$\alpha^{(k)} = \frac{\mathbf{r}^{(k)\top} \mathbf{r}^{(k)}}{\mathbf{r}^{(k)\top} \mathbf{A} \mathbf{r}^{(k)}}$$

# Méthodes itératives

## Méthode Richardson (gradient)

1. Résoudre le système linéaire  $\mathbf{P}\mathbf{z}^{(k)} = \mathbf{r}^{(k)}$
2. Calculer le paramètre d'accélération  $\alpha^{(k)} = f(\mathbf{z}^{(k)})$
3. Mettre à jour la solution  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{z}^{(k)}$
4. Mettre à jour le résidu  $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha^{(k)}\mathbf{A}\mathbf{z}^{(k)}$

La méthode du gradient est ainsi donnée par l'algorithme

1.  $\mathbf{P} = \mathbf{I} \rightarrow \mathbf{z}^{(k)} = \mathbf{r}^{(k)}$
2.  $\alpha^{(k)} = \frac{\mathbf{r}^{(k)\top} \mathbf{r}^{(k)}}{\mathbf{r}^{(k)\top} \mathbf{A} \mathbf{r}^{(k)}}$
3.  $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)} \mathbf{z}^{(k)}$
4.  $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha^{(k)} \mathbf{A} \mathbf{z}^{(k)}$

qui est un cas particulier des méthodes de Richardson.

# Méthodes itératives

## Méthodes de Richardson

- Méthode du gradient

$$\alpha^{(k)} = \frac{\mathbf{r}^{(k)\top} \mathbf{r}^{(k)}}{\mathbf{r}^{(k)\top} \mathbf{A} \mathbf{r}^{(k)}}, \quad \mathbf{P} = \mathbf{I}$$

- Méthode linéaire de Jacobi

$$\alpha^{(k)} = 1, \quad \mathbf{P} = \omega^{-1} \mathbf{D}$$

- Méthode linéaire de Gauss-Seidel

$$\alpha^{(k)} = 1, \quad \mathbf{P} = \omega^{-1} \mathbf{D} - \mathbf{E}$$

## Références

- *Méthodes numériques pour le calcul scientifique*, A. Quarteroni, R. Sacco, F. Saleri, Springer, 2000
- *Iterative methods for sparse linear systems*, Y. Saad, PWS, 1996
- *High-order methods for incompressible fluid flow*, M.O. Deville, P.F. Fischer, E.H. Mund, Cambridge University Press, 2002