

## Chapitre 6

### 6.3 Éléments finis en dimension $N \geq 2$

#### Plan du cours:

- ➔ 6.3.1 Éléments finis triangulaires (de Lagrange)
- ➔ 6.3.2 Convergence et estimation d'erreur
- ➔ 6.3.3 Éléments finis rectangulaires (de Lagrange)
- ➔ 13.1 Résolution numérique des systèmes linéaires

## Problème modèle

$\Omega$  ouvert borné de  $\mathbb{R}^N$  et  $f \in L^2(\Omega)$ .

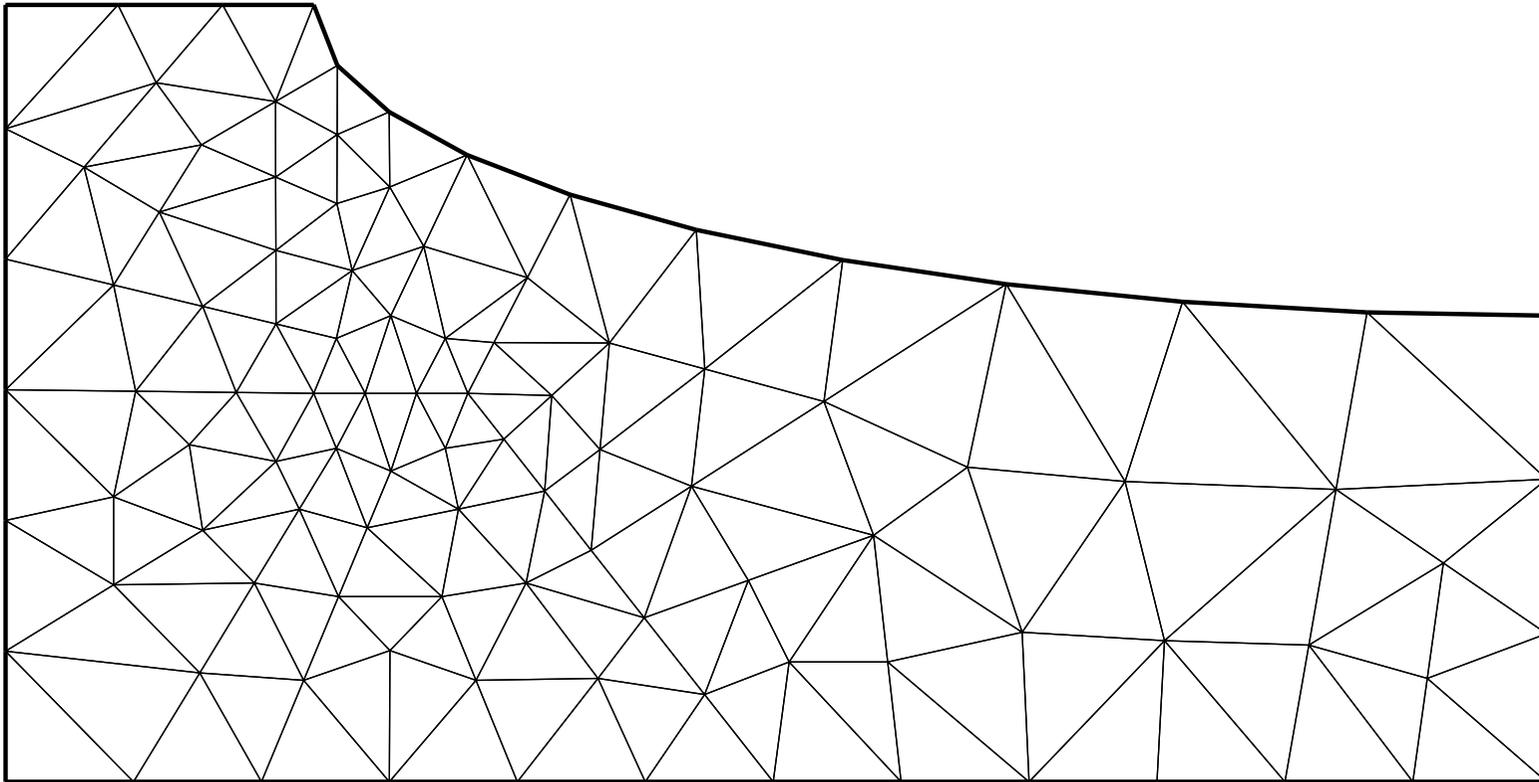
$$\begin{cases} -\Delta u = f & \text{dans } \Omega \\ u = 0 & \text{sur } \partial\Omega. \end{cases}$$

Il existe une solution unique dans  $H_0^1(\Omega)$ .

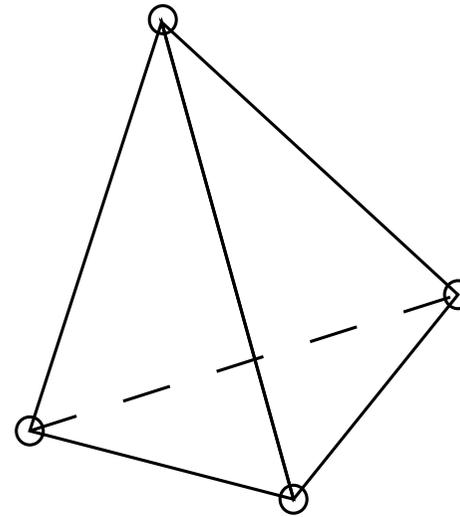
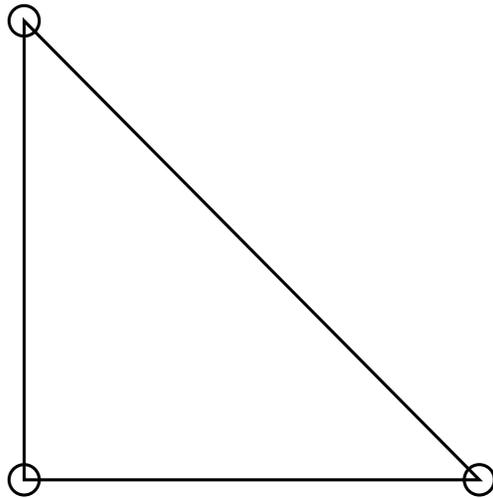
Dans tout ce qui suit nous supposerons que le domaine  $\Omega$  est **polyédrique** (polygonal si  $N = 2$ ), afin que nous puissions le mailler exactement.

### 6.3.1 Éléments finis triangulaires

Exemple de maillage en dimension  $N = 2$ :



## Mailles



Les mailles sont des  $N$ -simplexes (triangles en 2-D, tétraèdres en 3-D).

## N-simplexe

Un  $N$ -simplexe est l'enveloppe convexe de  $(N + 1)$  points  $(a_j)_{1 \leq j \leq N+1}$  de  $\mathbb{R}^N$ .

On note  $(a_{i,j})_{1 \leq i \leq N}$  les coordonnées du vecteur  $a_j$ . Le  $N$ -simplexe est **non dégénéré** si la matrice

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,N+1} \\ a_{2,1} & a_{2,2} & \dots & a_{2,N+1} \\ \vdots & \vdots & & \vdots \\ a_{N,1} & a_{N,2} & \dots & a_{N,N+1} \\ 1 & 1 & \dots & 1 \end{pmatrix}$$

est inversible (ce que l'on supposera toujours par la suite). Un  $N$ -simplexe a autant de faces que de sommets, qui sont elles-mêmes des  $(N - 1)$ -simplexes.

## Coordonnées barycentriques

Si  $K$  est un  $N$ -simplexe non dégénéré de sommets  $(a_j)_{1 \leq j \leq N+1}$ , les **coordonnées barycentriques**  $(\lambda_j)_{1 \leq j \leq N+1}$  de  $x \in \mathbb{R}^N$  sont définies par

$$\left\{ \begin{array}{l} \sum_{j=1}^{N+1} a_{i,j} \lambda_j = x_i \quad \text{pour } 1 \leq i \leq N \\ \sum_{j=1}^{N+1} \lambda_j = 1 \end{array} \right.$$

qui admet une solution unique car la matrice  $A$  est inversible. (Remarquons que  $x \rightarrow \lambda_j(x)$  est affine.) On vérifie que

$$K = \{x \in \mathbb{R}^N \text{ tel que } \lambda_j(x) \geq 0 \text{ pour } 1 \leq j \leq N + 1\},$$

et que les  $(N + 1)$  faces de  $K$  sont les intersections de  $K$  et des hyperplans  $\lambda_j(x) = 0$ ,  $1 \leq j \leq N + 1$ .

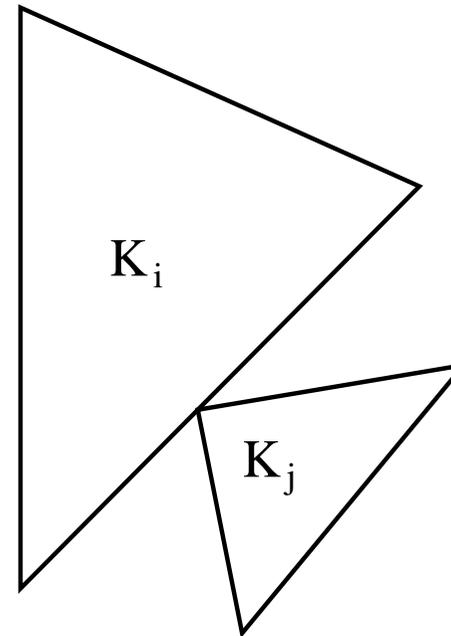
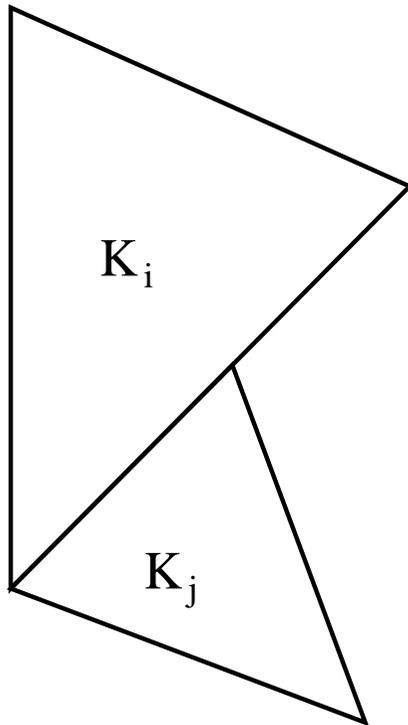
### Définition 6.3.1

Soit  $\Omega$  un ouvert connexe polyédrique de  $\mathbb{R}^N$ . Un maillage ou une triangulation de  $\bar{\Omega}$  est un ensemble  $\mathcal{T}_h$  de  $N$ -simplexes (non dégénérés)  $(K_i)_{1 \leq i \leq n}$  qui vérifient

1.  $K_i \subset \bar{\Omega}$  et  $\bar{\Omega} = \cup_{i=1}^n K_i$ ,
2. en dimension  $N = 2$ , l'intersection  $K_i \cap K_j$  de deux triangles distincts est soit vide, soit réduite à un sommet commun, soit à une arête commune **entière** (en dimension  $N = 3$ , l'intersection est soit vide, soit un sommet commun, soit une face commune entière, soit une arête commune entière).

Les **sommets** (ou **noeuds**) du maillage  $\mathcal{T}_h$  sont les sommets des  $N$ -simplexes  $K_i$  qui le composent. Par convention, le paramètre  $h$  désigne le maximum des diamètres des  $N$ -simplexes  $K_i$ .

## Situations interdites



## Treillis d'ordre $k$

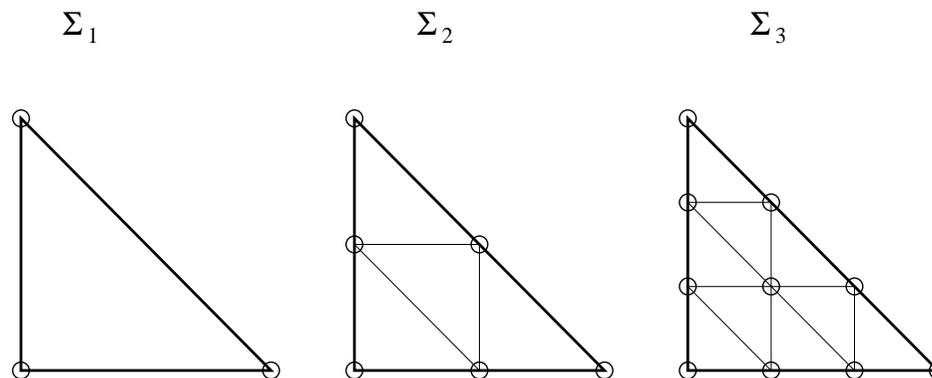
On appelle **treillis d'ordre  $k$**  l'ensemble (fini)

$$\Sigma_k = \left\{ x \in K \text{ tel que } \lambda_j(x) \in \left\{ 0, \frac{1}{k}, \dots, \frac{k-1}{k}, 1 \right\} \text{ pour } 1 \leq j \leq N \right\}$$

dont les points sont notés  $(\sigma_j)_{1 \leq j \leq n_k}$ .

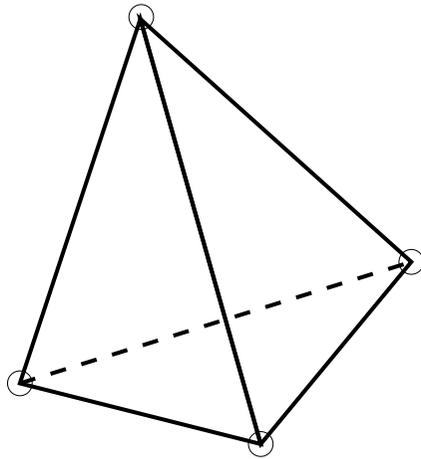
Pour  $k = 1$  il s'agit de l'ensemble des sommets de  $K$ , et pour  $k = 2$  des sommets et des points milieux des arêtes reliant deux sommets.

Dimension  $N = 2$

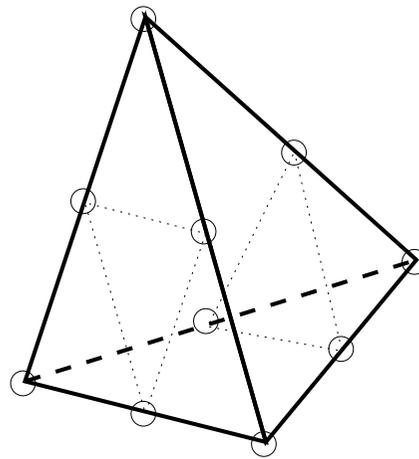


Dimension  $N = 3$

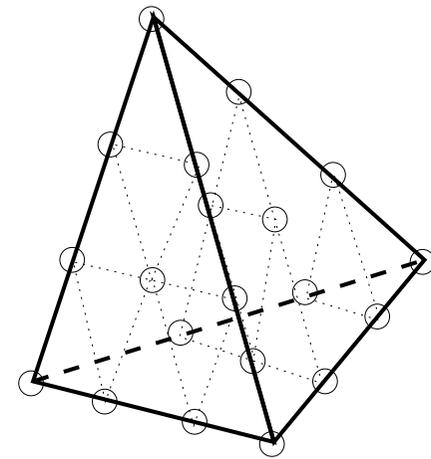
$\Sigma_1$



$\Sigma_2$



$\Sigma_3$



## Ensemble de polynômes $\mathbb{P}_k$

On définit l'ensemble  $\mathbb{P}_k$  des polynômes à coefficients réels de  $\mathbb{R}^N$  dans  $\mathbb{R}$  de **degré inférieur ou égal à  $k$** , c'est-à-dire que tout  $p \in \mathbb{P}_k$  s'écrit sous la forme

$$p(x) = \sum_{\substack{i_1, \dots, i_N \geq 0 \\ i_1 + \dots + i_N \leq k}} \alpha_{i_1, \dots, i_N} x_1^{i_1} \cdots x_N^{i_N} \quad \text{avec } x = (x_1, \dots, x_N).$$

Degré  $k = 1$ : polynômes affines

$$p(x) = \alpha_0 + \sum_{i=1}^N \alpha_i x_i \quad \text{avec } x = (x_1, \dots, x_N).$$

En général,  $k = 1$ , parfois  $k = 2$ , plus rarement au delà.

## Unisolvance de $\Sigma_k$ pour $\mathbb{P}_k$

**Lemme 6.3.3** Tout polynôme de  $\mathbb{P}_k$  est déterminé de **manière unique** par ses valeurs aux points  $(\sigma_j)_{1 \leq j \leq n_k}$  du **treillis**  $\Sigma_k$ . Autrement dit, il existe une base  $(\psi_j)_{1 \leq j \leq n_k}$  de  $\mathbb{P}_k$  telle que

$$\psi_j(\sigma_i) = \delta_{ij} \quad 1 \leq i, j \leq n_k.$$

**Preuve** ( $k = 1$ ). On vérifie que tout polynôme  $p \in \mathbb{P}_1$  se met sous la forme

$$p(x) = \sum_{j=1}^{N+1} p(a_j) \lambda_j(x)$$

car les coordonnées barycentriques  $\lambda_j(x)$  sont affines en  $x$ . Comme  $\lambda_j(a_k) = \delta_{jk}$ , la base recherchée est  $\psi_j = \lambda_j$ .

## Continuité à l'interface entre 2 mailles

**Lemme 6.3.4** Soit  $K$  et  $K'$  deux  $N$ -simplexes ayant une **face commune**  $\Gamma = \partial K \cap \partial K'$ . Alors, leur treillis d'ordre  $k \geq 1$ ,  $\Sigma_k$  et  $\Sigma'_k$ , **coïncident** sur cette face  $\Gamma$ . De plus, étant donné  $p_K$  et  $p_{K'}$  deux polynômes de  $\mathbb{P}_k$ , la fonction  $v$  définie par

$$v(x) = \begin{cases} p_K(x) & \text{si } x \in K \\ p_{K'}(x) & \text{si } x \in K' \end{cases}$$

est **continue** sur  $K \cup K'$ , si et seulement si  $p_K$  et  $p_{K'}$  ont des valeurs qui coïncident aux points du treillis sur la face commune  $\Gamma$ .

**Preuve.** Par construction les treillis  $\Sigma_k$  et  $\Sigma'_k$  coïncident sur leur face commune  $\Gamma$ . Si les polynômes  $p_K$  et  $p_{K'}$  coïncident aux points de  $\Sigma_k \cap \Gamma$ , alors par application du Lemme 6.3.3 ils sont égaux sur  $\Gamma$ , ce qui prouve la continuité de  $v$ .

## Eléments finis $\mathbb{P}_k$

**Définition 6.3.5.** Etant donné un maillage  $\mathcal{T}_h$  d'un ouvert  $\Omega$ , la méthode des éléments finis  $\mathbb{P}_k$ , ou **éléments finis triangulaires de Lagrange d'ordre  $k$** , associée à ce maillage, est définie par l'espace discret

$$V_h = \{v \in C(\overline{\Omega}) \text{ tel que } v|_{K_i} \in \mathbb{P}_k \text{ pour tout } K_i \in \mathcal{T}_h\}.$$

On appelle **noeuds des degrés de liberté** l'ensemble des points (distincts)  $(\hat{a}_i)_{1 \leq i \leq n_{dl}}$  des treillis d'ordre  $k$  de chacun des  $N$ -simplexes  $K_i \in \mathcal{T}_h$ .

On appelle **degrés de liberté** d'une fonction  $v \in V_h$  l'ensemble des valeurs de  $v$  en ces noeuds  $(\hat{a}_i)_{1 \leq i \leq n_{dl}}$ .

On définit aussi le sous-espace  $V_{0h}$  par

$$V_{0h} = \{v \in V_h \text{ tel que } v = 0 \text{ sur } \partial\Omega\}.$$

**Proposition 6.3.7** L'espace  $V_h$  est un sous-espace de  $H^1(\Omega)$  dont la dimension est le nombre de degrés de liberté, et il existe une base  $(\phi_i)_{1 \leq i \leq n_{dl}}$  de  $V_h$  définie par

$$\phi_i(\hat{a}_j) = \delta_{ij} \quad 1 \leq i, j \leq n_{dl},$$

telle que

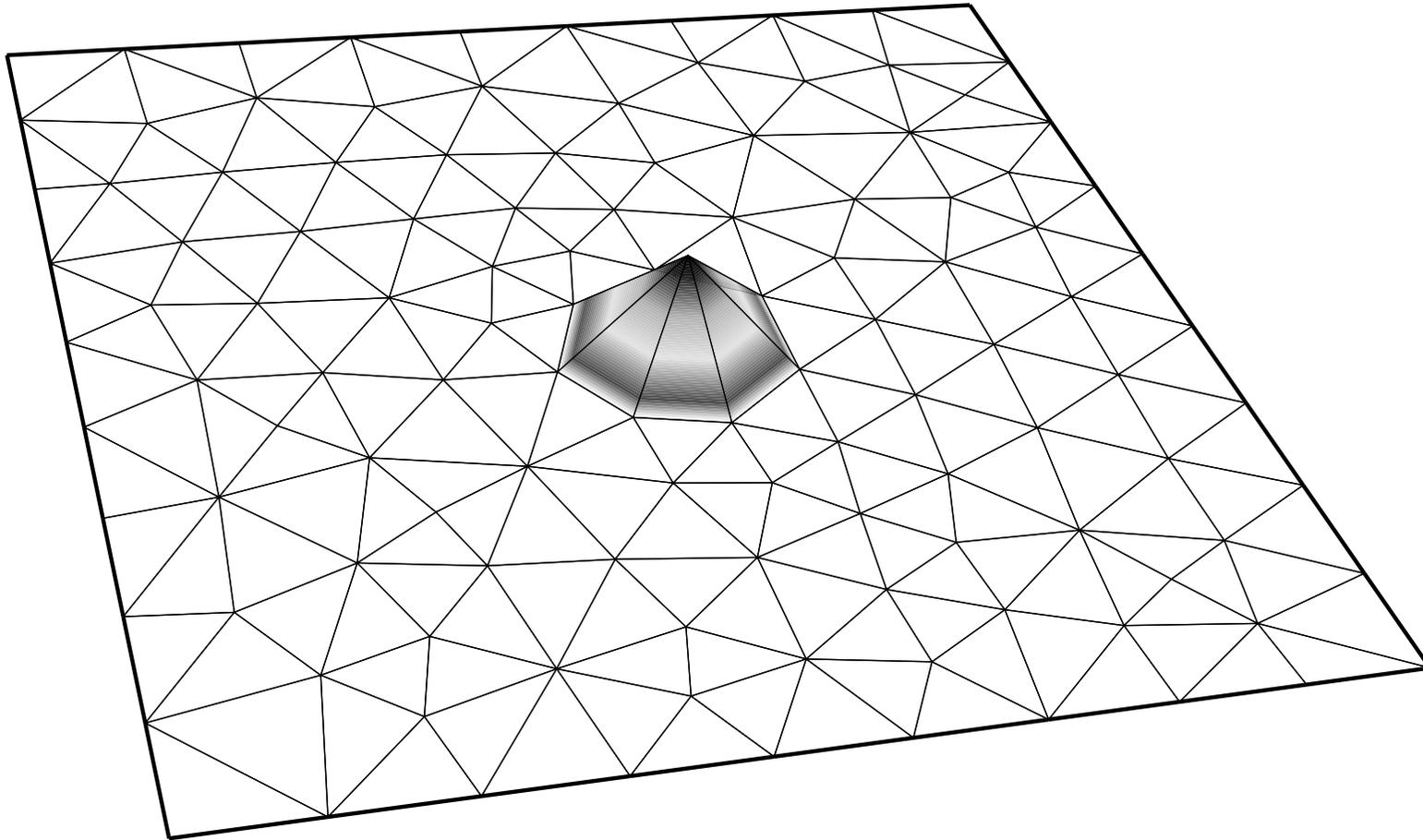
$$v(x) = \sum_{i=1}^{n_{dl}} v(\hat{a}_i) \phi_i(x).$$

**Preuve:** par simple combinaison des Lemmes 6.3.3 et 6.3.4.

**Remarque.** L'appellation “**éléments finis de Lagrange**” veut dire que toute fonction de l'espace  $V_h$  est caractérisée par ses **valeurs ponctuelles** (ses degrés de liberté) aux noeuds  $(\hat{a}_j)$ .

On parle **d'éléments finis de Hermite** si les degrés de liberté sont les valeurs de la fonction et de ses dérivées partielles d'ordre 1.

Fonction de base  $P_1$  en dimension  $N = 2$ .



## Résolution pratique

On résout le problème modèle par la méthode des éléments finis  $\mathbb{P}_k$ .

La formulation variationnelle de l'approximation interne est

$$\text{trouver } u_h \in V_{0h} \text{ tel que } \int_{\Omega} \nabla u_h \cdot \nabla v_h \, dx = \int_{\Omega} f v_h \, dx \quad \forall v_h \in V_{0h}.$$

On décompose  $u_h$  sur la base des  $(\phi_j)_{1 \leq j \leq n_{dl}}$  et on prend  $v_h = \phi_i$  ce qui donne

$$\sum_{j=1}^{n_{dl}} u_h(\hat{a}_j) \int_{\Omega} \nabla \phi_j \cdot \nabla \phi_i \, dx = \int_{\Omega} f \phi_i \, dx.$$

## Matrice de rigidité

Vecteur inconnu:  $U_h = (u_h(\hat{a}_j))_{1 \leq j \leq n_{dl}}$

Second membre:  $b_h = \left( \int_{\Omega} f \phi_i dx \right)_{1 \leq i \leq n_{dl}}$

Matrice de rigidité:  $\mathcal{K}_h = \left( \int_{\Omega} \nabla \phi_j \cdot \nabla \phi_i dx \right)_{1 \leq i, j \leq n_{dl}}$

La formulation variationnelle est équivalente au système linéaire

$$\mathcal{K}_h U_h = b_h.$$

En général, l'intersection des supports de  $\phi_j$  et  $\phi_i$  est vide et la plupart des coefficients de  $\mathcal{K}_h$  sont nuls. La matrice de rigidité  $\mathcal{K}_h$  est donc creuse.

## Traitement des conditions aux limites

- ➡ Condition aux limites de Neumann: **rien à faire !** (Elle est prise en compte par la formulation variationnelle: on dit qu'elle est implicite ou naturelle, voir la Remarque 5.2.11.)
- ➡ Condition aux limites de Fourier (i.e.  $\frac{\partial u}{\partial n} + \alpha u = 0$  sur  $\partial\Omega$ ): on rajoute un terme

$$\alpha \int_{\partial\Omega} uv \, ds$$

dans la formulation variationnelle.

- ➡ Condition aux limites de Dirichlet: l'inconnue comme la fonction test est nulle sur le bord. Deux possibilités: soit **on élimine** les degrés de liberté du bord, soit **on pénalise** les degrés de liberté du bord (i.e. Fourier avec  $\alpha \gg 1$ ).

## Calcul des intégrales

On peut utiliser la formule exacte en coordonnées barycentriques

$$\int_K \lambda_1(x)^{\alpha_1} \cdots \lambda_{N+1}(x)^{\alpha_{N+1}} dx = \text{Volume}(K) \frac{\alpha_1! \cdots \alpha_{N+1}! N!}{(\alpha_1 + \cdots + \alpha_{N+1} + N)!}.$$

On peut aussi utiliser des “**formules de quadrature**” approchées

formule du point milieu:  $\int_K \psi(x) dx \approx \text{Volume}(K) \psi(a_0),$

avec  $a_0 = (N + 1)^{-1} \sum_{i=1}^{N+1} a_i,$  le barycentre de  $K,$

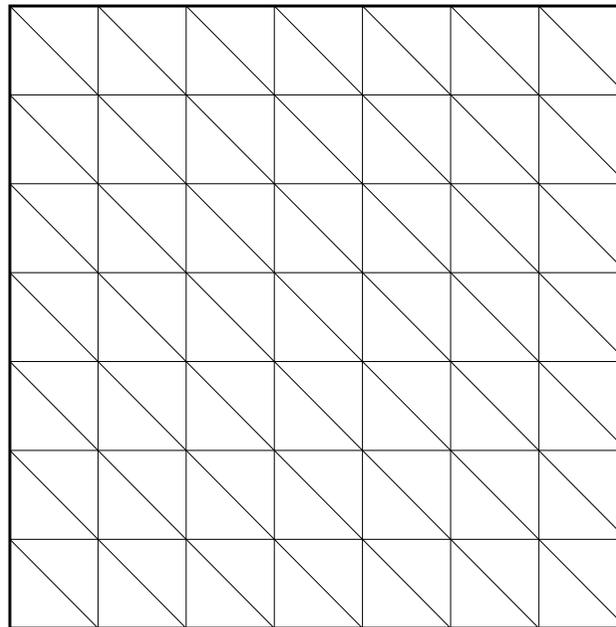
formule des trapèzes:  $\int_K \psi(x) dx \approx \frac{\text{Volume}(K)}{N + 1} \sum_{i=1}^{N+1} \psi(a_i).$

Ces formules sont exactes pour des fonctions affines et sont approchées à l'ordre 2 en  $h$  pour des fonctions régulières.

## Taille des matrices

La matrice de rigidité  $\mathcal{K}_h$  est creuse **mais elle est de grande taille!**

Exemple: maillage régulier  $n \times n$  en dimension  $N = 2$

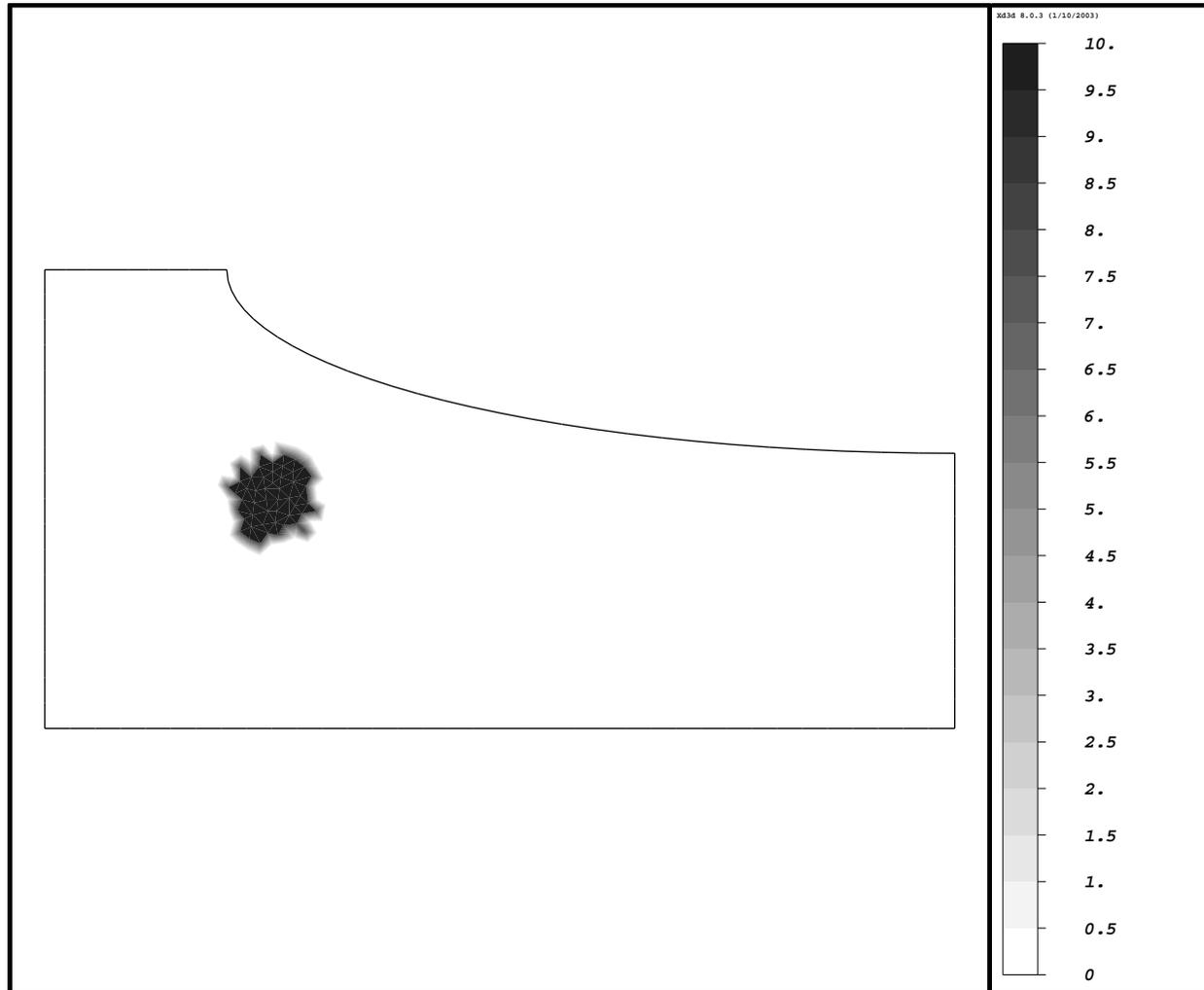


Matrice  $\mathcal{K}_h$  d'ordre  $n^2$  (ou bien  $n^3$  en dimension  $N = 3$ ).

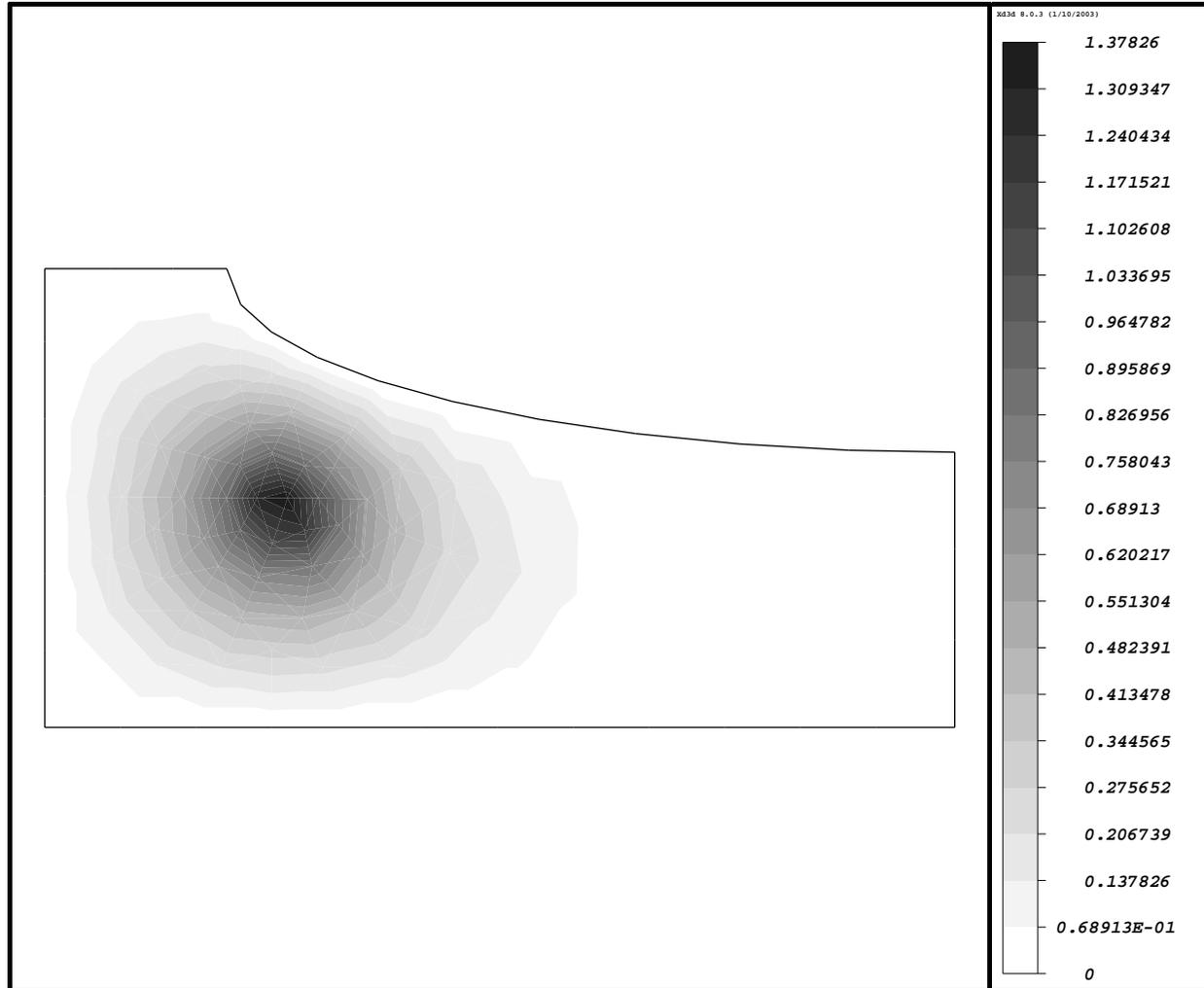
**Il faut optimiser la résolution du système linéaire !**

## Exemples numériques avec FreeFem++

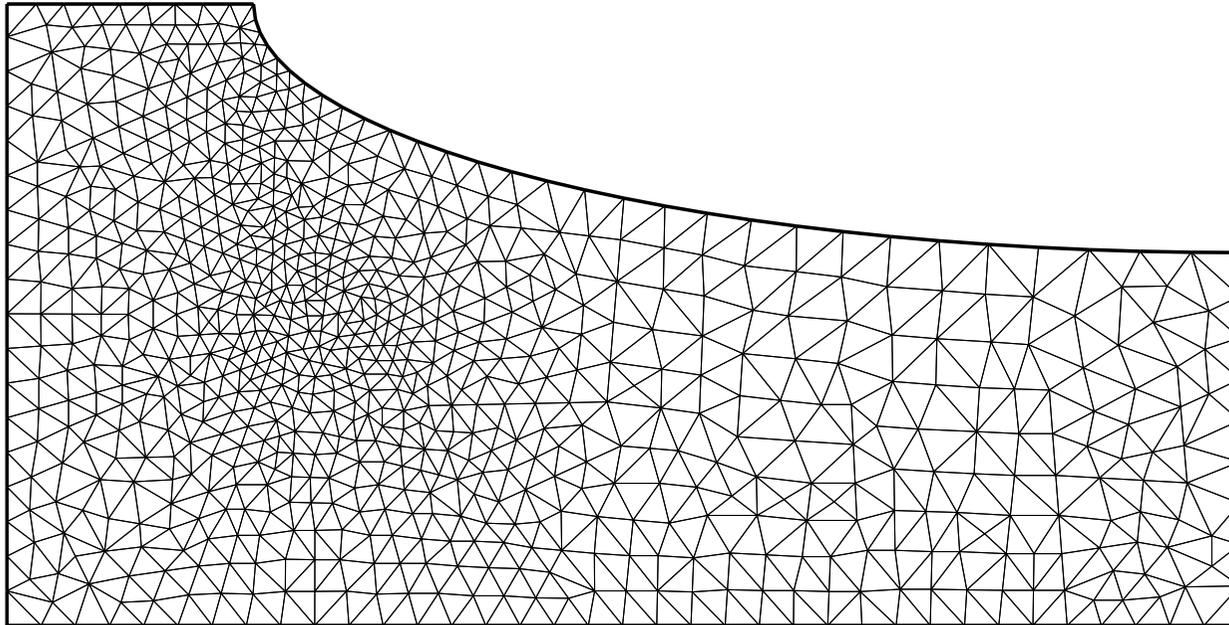
Terme source  $f$



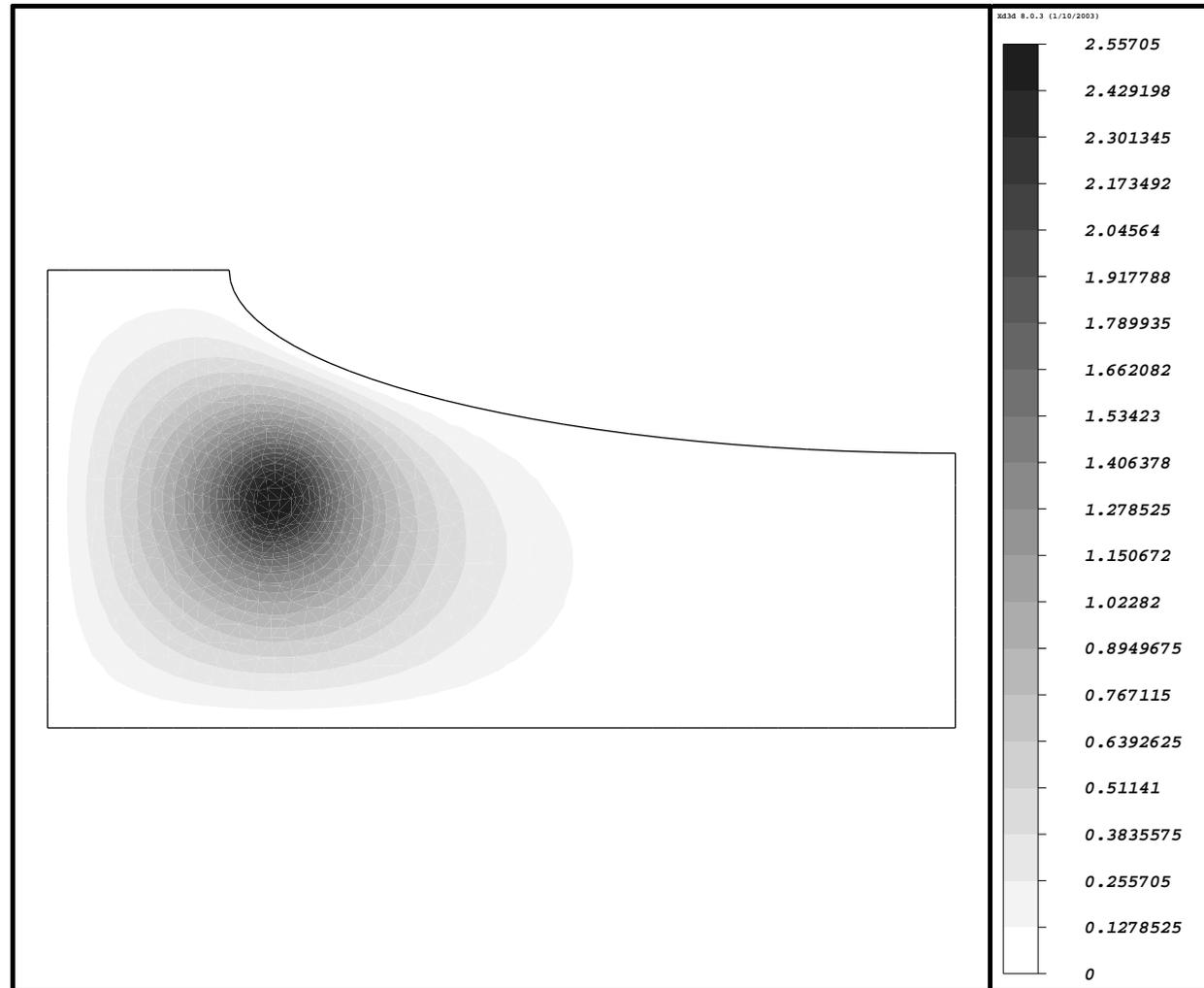
Solution approchée  $u_h$  pour le maillage “grossier”



Maillage triangulaire plus fin que le précédent

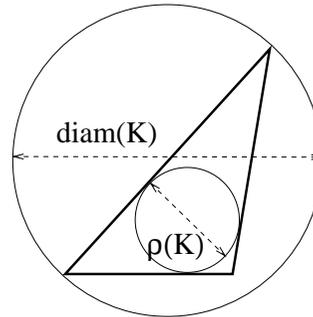


Solution approchée  $u_h$  pour le maillage "fin"



### 6.3.2 Convergence et estimation d'erreur

Diamètre  $h_K = \text{diam}(K)$  et rondeur  $\rho(K)$  d'un triangle  $K$



**Définition 6.3.11** Soit  $(\mathcal{T}_h)_{h>0}$  une suite de maillages de  $\Omega$ . On dit qu'il s'agit d'une **suite de maillages réguliers** si

1. la suite  $h = \max_{K_i \in \mathcal{T}_h} \text{diam}(K_i)$  tend vers 0,
2. il existe une constante  $C$  telle que, pour tout  $h > 0$  et tout  $K \in \mathcal{T}_h$ ,

$$1 \leq \frac{\text{diam}(K)}{\rho(K)} \leq C.$$

## Convergence et estimation d'erreur

**Théorème 6.3.13** Soit  $(\mathcal{T}_h)_{h>0}$  une suite de maillages réguliers de  $\Omega$ . Soit  $u \in H_0^1(\Omega)$ , la solution exacte, et  $u_h \in V_{0h}$ , la solution approchée par éléments finis  $\mathbb{P}_k$ .

La **méthode des éléments finis  $\mathbb{P}_k$  converge**, c'est-à-dire que

$$\lim_{h \rightarrow 0} \|u - u_h\|_{H^1(\Omega)} = 0.$$

De plus, si  $u \in H^{k+1}(\Omega)$  et si  $k + 1 > N/2$ , alors on a **l'estimation d'erreur**

$$\|u - u_h\|_{H^1(\Omega)} \leq Ch^k \|u\|_{H^{k+1}(\Omega)}$$

**Remarque.** Le Théorème 6.3.13 s'applique à toute méthode d'éléments finis de type Lagrange (aussi pour les éléments finis rectangulaires).

Pour  $N = 2$  ou  $N = 3$ , la condition  $k + 1 > N/2$  est **satisfaite** dès que  $k \geq 1$ .

**Démonstration du Théorème 6.3.13:** idée principale

Lemme 6.1.2 de Céa + interpolation ci-dessous.

**Définition d'un opérateur d'interpolation  $r_h$ .** Pour toute fonction continue  $v$  on définit son interpolée

$$r_h v(x) = \sum_{i=1}^{n_{dl}} v(\hat{a}_i) \phi_i(x)$$

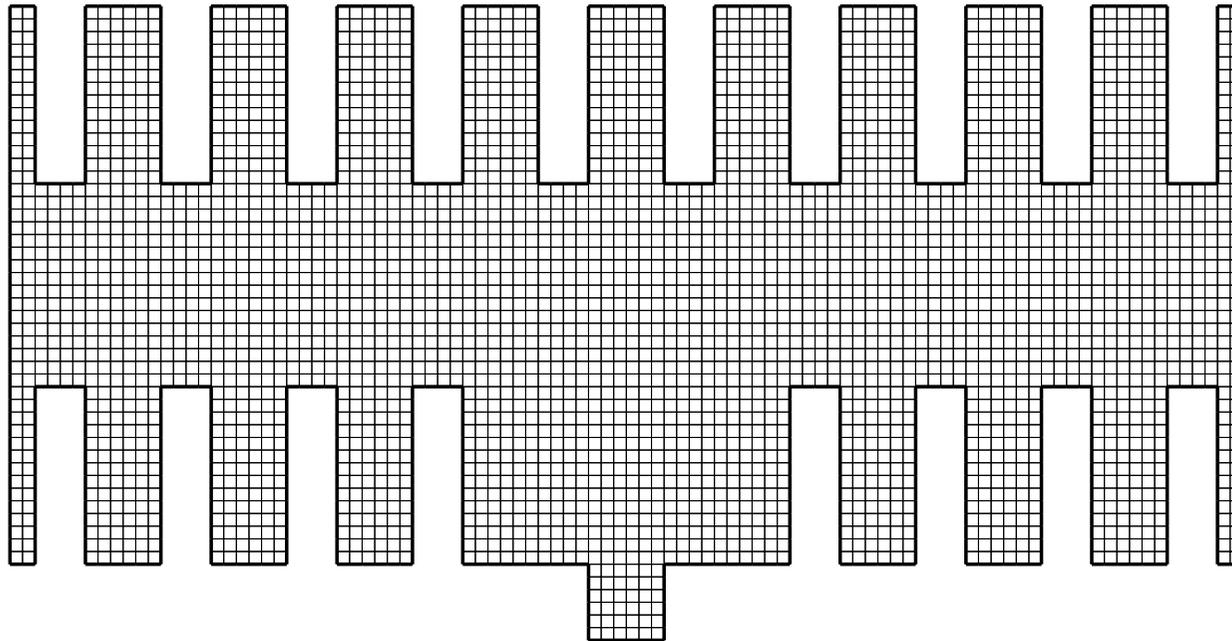
avec  $(\hat{a}_i)_{1 \leq i \leq n_{dl}}$  les noeuds des degrés de liberté et  $(\phi_i)_{1 \leq i \leq n_{dl}}$  la base de  $V_{0h}$  de la méthode des éléments finis  $\mathbb{P}_k$ .

**Proposition 6.3.16 (admise)** Soit  $(\mathcal{T}_h)_{h>0}$  une suite de maillages réguliers de  $\Omega$ . On suppose que  $k + 1 > N/2$ . Alors, pour tout  $v \in H^{k+1}(\Omega)$  l'interpolée  $r_h v$  est bien définie, et il existe une constante  $C$ , indépendante de  $h$  et de  $v$ , telle que

$$\|v - r_h v\|_{H^1(\Omega)} \leq C h^k \|v\|_{H^{k+1}(\Omega)}.$$

### 6.3.3 Éléments finis rectangulaires

Exemple de maillage rectangulaire en dimension  $N = 2$



Définition 6.3.21

Soit  $\Omega$  un ouvert connexe polyédrique de  $\mathbb{R}^N$ . Un **maillage rectangulaire** de  $\overline{\Omega}$  est un ensemble  $\mathcal{T}_h$  de  $N$ -rectangles (non dégénérés)  $(K_i)_{1 \leq i \leq n}$  qui vérifient

1.  $K_i \subset \overline{\Omega}$  et  $\overline{\Omega} = \cup_{i=1}^n K_i$ ,
2. en dimension  $N = 2$ , l'intersection  $K_i \cap K_j$  de deux rectangles distincts est soit vide, soit un sommet commun, soit une arête commune entière (en dimension  $N = 3$  il faut ajouter soit une face commune entière).

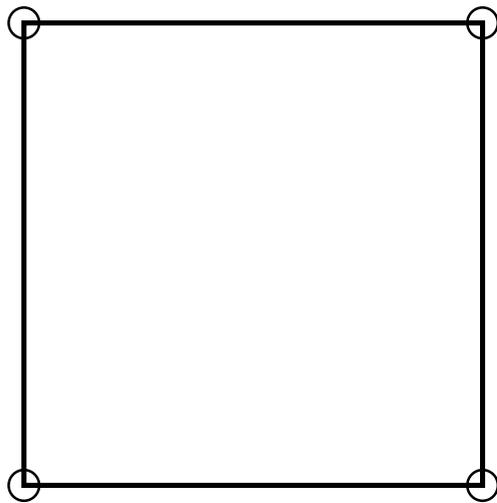
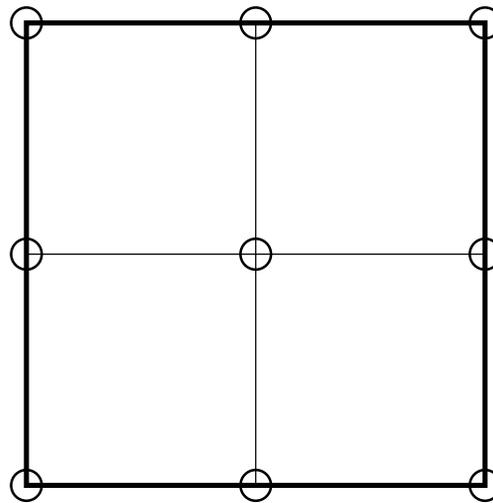
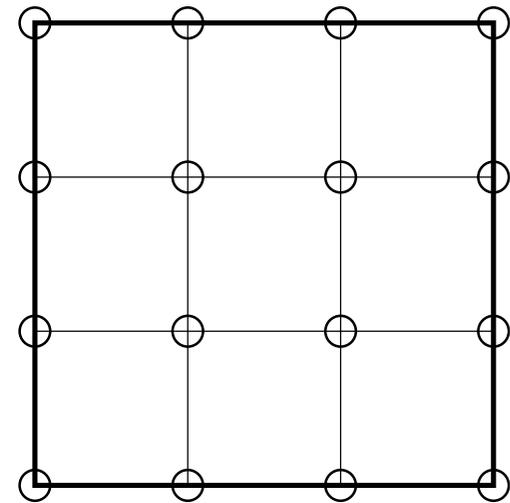
Les **sommets (ou noeuds)** du maillage  $\mathcal{T}_h$  sont les sommets des  $N$ -rectangles  $K_i$  qui le composent. Par convention, le paramètre  $h$  désigne le maximum des diamètres des  $N$ -rectangles  $K_i$ .

## Treillis

Pour tout entier  $k \geq 1$  on définit le **treillis d'ordre  $k$**  du  $N$ -rectangle  $K$  comme l'ensemble (fini)

$$\Sigma_k = \left\{ x \in K \text{ tel que } \frac{x_j - l_j}{L_j - l_j} \in \left\{ 0, \frac{1}{k}, \dots, \frac{k-1}{k}, 1 \right\} \text{ pour } 1 \leq j \leq N \right\}.$$

Pour  $k = 1$  il s'agit de l'ensemble des sommets de  $K$ .

 $\Sigma_1$ 

 $\Sigma_2$ 

 $\Sigma_3$ 


## Ensemble de polynômes $\mathbb{Q}_k$

On définit  $\mathbb{Q}_k$  comme l'ensemble des polynômes à coefficients réels de  $\mathbb{R}^N$  dans  $\mathbb{R}$  de degré inférieur ou égal à  $k$  **par rapport à chaque variable**, c'est-à-dire que tout  $p \in \mathbb{Q}_k$  s'écrit sous la forme

$$p(x) = \sum_{0 \leq i_1 \leq k, \dots, 0 \leq i_N \leq k} \alpha_{i_1, \dots, i_N} x_1^{i_1} \cdots x_N^{i_N} \text{ avec } x = (x_1, \dots, x_N).$$

Remarquons que le degré total de  $p$  peut être supérieur à  $k$ , **ce qui différencie l'espace  $\mathbb{Q}_k$  de  $\mathbb{P}_k$** .

**Degré  $k = 1$  et dimension  $N = 2$ : polynômes bilinéaires**

$$p(x) = \alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_1 x_2.$$

En général,  $k = 1$ , parfois  $k = 2$ , plus rarement au delà.

Unisolvance de  $\Sigma_k$  pour  $\mathbb{Q}_k$

**Lemme 6.3.22** Soit  $K$  un  $N$ -rectangle. Soit un entier  $k \geq 1$ . Alors, tout polynôme de  $\mathbb{Q}_k$  est **déterminé de manière unique** par ses valeurs aux points du treillis d'ordre  $k$   $\Sigma_k$ .

**Lemme 6.3.23** Soit  $K$  et  $K'$  deux  $N$ -rectangles ayant une face commune  $\Gamma = \partial K \cap \partial K'$ . Alors, leur treillis d'ordre  $k \geq 1$ ,  $\Sigma_k$  et  $\Sigma'_k$ , coïncident sur cette face  $\Gamma$ . De plus, étant donné  $p_K$  et  $p_{K'}$  deux polynômes de  $\mathbb{Q}_k$ , la fonction  $v$  définie par

$$v(x) = \begin{cases} p_K(x) & \text{si } x \in K \\ p_{K'}(x) & \text{si } x \in K' \end{cases}$$

est continue sur  $K \cup K'$ , si et seulement si  $p_K$  et  $p_{K'}$  ont des valeurs qui coïncident aux points du treillis sur la face commune  $\Gamma$ .

## Eléments finis $\mathbb{Q}_k$

**Définition 6.3.25.** Etant donné un maillage rectangulaire  $\mathcal{T}_h$  d'un ouvert  $\Omega$ , la méthode des éléments finis  $\mathbb{Q}_k$  est définie par l'espace discret

$$V_h = \{v \in \mathcal{C}(\overline{\Omega}) \text{ tel que } v|_{K_i} \in \mathbb{Q}_k \text{ pour tout } K_i \in \mathcal{T}_h\}.$$

On appelle noeuds des **degrés de liberté** l'ensemble des points  $(\hat{a}_i)_{1 \leq i \leq n_{dl}}$  des treillis d'ordre  $k$  de chacun des  $N$ -rectangles  $K_i \in \mathcal{T}_h$ .

**Proposition 6.3.26** L'espace  $V_h$  est un sous-espace de  $H^1(\Omega)$  dont la dimension est le nombre de degrés de liberté  $n_{dl}$ . De plus, il existe une base de  $V_h$   $(\phi_i)_{1 \leq i \leq n_{dl}}$  définie par

$$\phi_i(\hat{a}_j) = \delta_{ij} \quad 1 \leq i, j \leq n_{dl},$$

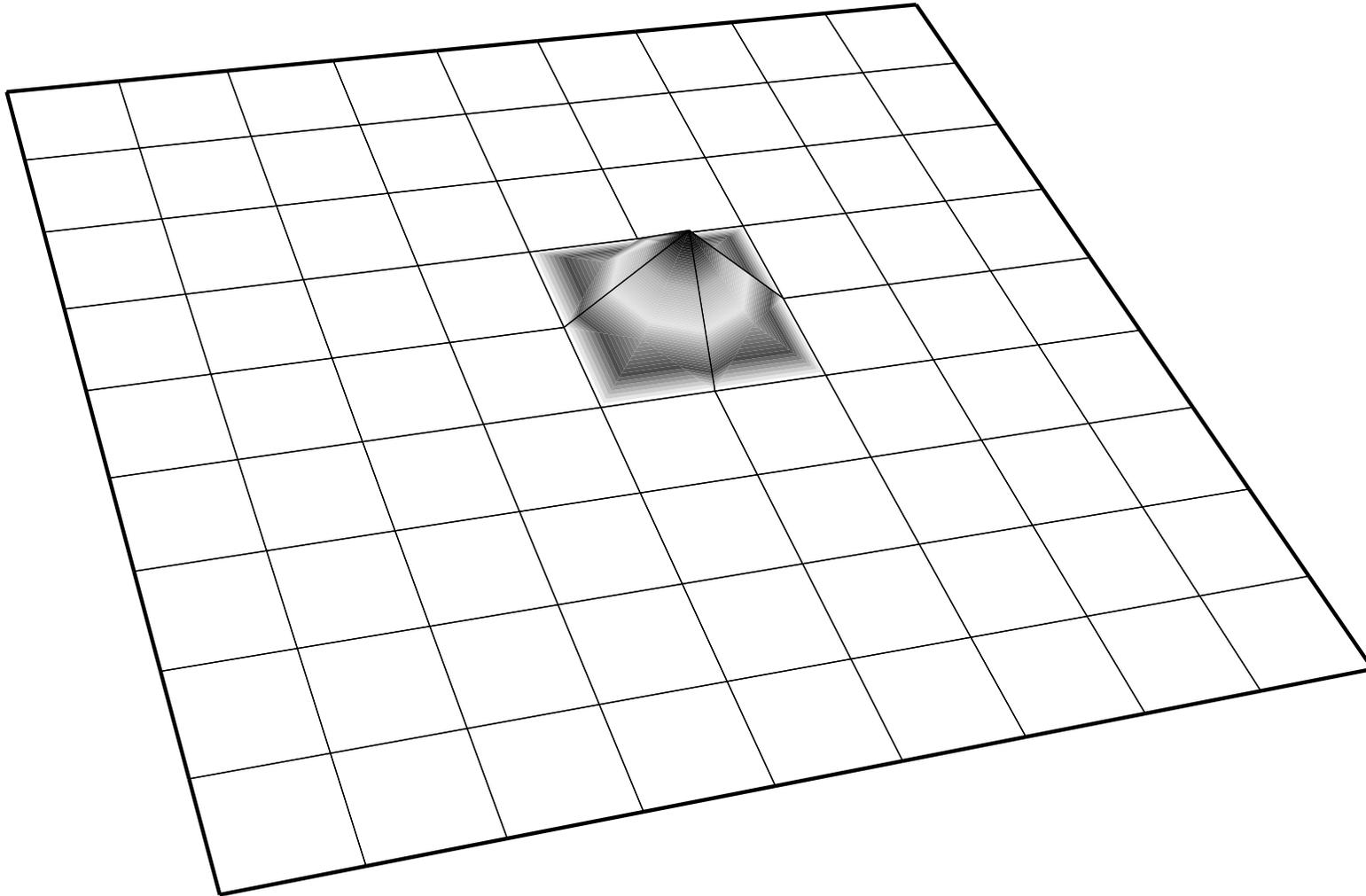
telle que

$$v(x) = \sum_{i=1}^{n_{dl}} v(\hat{a}_i) \phi_i(x).$$

**Remarque.** Il s'agit encore d'éléments finis de Lagrange.

Même résultat de convergence que pour les éléments finis triangulaires.

Fonction de base  $Q_1$  en dimension  $N = 2$ .



## 13.1 Résolution des systèmes linéaires

**Problème:** résoudre le système linéaire dans  $\mathbb{R}^n$

$$Ax = b \quad \text{avec } A \in \mathcal{M}_n(\mathbb{R}) \quad \text{et } n \text{ grand !}$$

On veut des algorithmes numériques **efficaces et stables !**

- ✗ Efficacité = faible temps de calcul, et faible occupation de mémoire.
- ✗ Stabilité = ne pas amplifier les erreurs d'arrondi.

Deux types de méthodes:

- ⇒ Méthodes directes (solution exacte en un nombre fini d'opérations).
- ⇒ Méthodes itératives (suite de solutions approchées).

## Stabilité et conditionnement

**Définition 13.1.1** norme matricielle **subordonnée**  $\|A\| = \max_{x \in \mathbb{C}^n} \frac{\|Ax\|}{\|x\|}$ .

**Définition 13.1.9** On appelle **conditionnement** d'une matrice  $A \in \mathcal{M}_n(\mathbb{C})$ , relatif à une norme matricielle subordonnée, la valeur définie par

$$\text{cond}(A) = \|A\| \cdot \|A^{-1}\|$$

**Proposition 13.1.10** Soit  $A$  une matrice inversible et  $b \in \mathbb{R}^n, b \neq 0$ .

1. Si  $Ax = b$  et  $A(x + \delta x) = b + \delta b$ , alors on a

$$\frac{\|\delta x\|}{\|x\|} \leq \text{cond}(A) \frac{\|\delta b\|}{\|b\|}.$$

2. Si  $Ax = b$  et  $(A + \delta A)(x + \delta x) = b$ , alors on a

$$\frac{\|\delta x\|}{\|x + \delta x\|} \leq \text{cond}(A) \frac{\|\delta A\|}{\|A\|}.$$

**Démonstration.**

1.  $A\delta x = \delta b$  donc  $\|\delta x\| \leq \|A^{-1}\| \|\delta b\|$ .
2.  $A\delta x = -\delta A(x + \delta x)$ , donc  $\|\delta x\| \leq \|A^{-1}\| \|\delta A\| \|x + \delta x\|$ .

**Exercice.**

Si  $A$  est symétrique réelle définie positive, on trouve

$$\text{cond}_2(A) = \frac{\lambda_n(A)}{\lambda_1(A)},$$

où  $0 < \lambda_1(A) \leq \dots \leq \lambda_n(A)$  sont les valeurs propres de  $A$ .

### Exemple

Pour les éléments finis  $\mathbb{P}_1$  appliqués au Laplacien, la matrice de rigidité est

$$\mathcal{K}_h = h^{-1} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix},$$

dont le conditionnement est  $\text{cond}_2(\mathcal{K}_h) \approx \frac{4}{\pi^2 h^2}$  pour  $h \approx 0$ .

La matrice de rigidité  $\mathcal{K}_h$  est mal conditionnée.

Il faut faire attention à la **stabilité** dans la résolution des systèmes linéaires issus de la méthode des éléments finis.

### 13.1.3 Méthodes directes

Matrice réelle inversible  $A$  d'ordre  $n$ .

- ✗ Elimination de Gauss.
- ✗ Factorisation LU.
- ✗ Factorisation de Cholesky pour les matrices symétriques.

Caractéristiques:

- ⇒ Mémoire requise: de l'ordre de  $n^2$  réels.
- ⇒ Temps nécessaire: de l'ordre de  $n^3$  opérations arithmétiques.

Avantage: simple, robuste, précis.

Inconvénient: trop chères, voire impossibles, si  $n$  est grand (ce qui est systématique en 3-D).

## Factorisation LU

Il s'agit de la méthode d'élimination de Gauss sans pivot.

**Proposition 13.1.15** Soit une matrice  $A = (a_{ij})_{1 \leq i, j \leq n}$  d'ordre  $n$ . Sous une hypothèse technique (vérifiée si  $A$  est définie positive), il existe un unique couple de matrices triangulaires  $(L, U)$  tel que  $A = LU$  avec

$$L = \begin{pmatrix} 1 & 0 & \dots & 0 \\ l_{2,1} & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ l_{n,1} & \dots & l_{n,n-1} & 1 \end{pmatrix}, \quad U = \begin{pmatrix} u_{1,1} & \dots & \dots & u_{1,n} \\ 0 & u_{2,2} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & u_{n,n} \end{pmatrix}.$$

Intérêt: il est facile de résoudre des systèmes triangulaires.

## Calcul pratique de la factorisation LU

$$A = LU \Rightarrow a_{i,j} = \sum_{k=1}^n l_{i,k} u_{k,j} = \sum_{k=1}^{\min(i,j)} l_{i,k} u_{k,j}.$$

Au fur et à mesure que l'on lit les colonnes de  $A$  on en déduit les coefficients des colonnes de  $L$  et de  $U$ .

**Colonne  $j$  de  $A$ :** on calcule la  $j$ -ème colonne de  $L$  et de  $U$

$$\begin{array}{ll}
 a_{1,j} = l_{1,1}u_{1,j} & \Rightarrow u_{1,j} = a_{1,j} \\
 \vdots & \vdots \\
 a_{j,j} = l_{j,1}u_{1,j} + \cdots + l_{j,j}u_{j,j} & \Rightarrow u_{j,j} = a_{j,j} - \sum_{k=1}^{j-1} l_{j,k}u_{k,j} \\
 a_{j+1,j} = l_{j+1,1}u_{1,j} + \cdots + l_{j+1,j}u_{j,j} & \Rightarrow l_{j+1,j} = \frac{a_{j+1,j} - \sum_{k=1}^{j-1} l_{j+1,k}u_{k,j}}{u_{j,j}} \\
 \vdots & \vdots \\
 a_{n,j} = l_{n,1}u_{1,j} + \cdots + l_{n,j}u_{j,j} & \Rightarrow l_{n,j} = \frac{a_{n,j} - \sum_{k=1}^{j-1} l_{n,k}u_{k,j}}{u_{j,j}}
 \end{array}$$

## Compte d'opérations

Pour  $n$  grand on ne compte que les multiplications ou divisions.

- factorisation LU : le nombre d'opérations  $N_{op}$  est

$$N_{op} = \sum_{j=1}^{n-1} \sum_{i=j+1}^n (1 + \sum_{k=j+1}^n 1),$$

qui, au premier ordre, donne  $N_{op} \approx n^3/3$ .

- substitution (ou remontée-descente sur les deux systèmes triangulaires) : le nombre d'opérations  $N_{op}$  est

$$N_{op} = 2 \sum_{j=1}^n j,$$

qui, au premier ordre, donne  $N_{op} \approx n^2$ .

## Factorisation de Cholesky

**Proposition 13.1.19** Soit  $A$  une matrice **symétrique réelle, définie positive**. Il existe une unique matrice réelle  $B$  triangulaire inférieure, telle que tous ses éléments diagonaux soient positifs, et qui vérifie

$$A = BB^*.$$

$$A = \begin{pmatrix} b_{1,1} & 0 & \dots & 0 \\ b_{2,1} & b_{2,2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ b_{n,1} & \dots & b_{n,n-1} & b_{n,n} \end{pmatrix} \begin{pmatrix} b_{1,1} & \dots & \dots & b_{n,1} \\ 0 & b_{2,2} & & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & b_{n,n} \end{pmatrix}.$$



### 13.1.4 Méthodes itératives

**Définition 13.1.24** Soit  $A$  une matrice inversible. Soit une décomposition régulière  $(M, N)$  de  $A$  (avec  $M$  inversible) telle que

$$A = M - N.$$

La méthode itérative basée sur le splitting  $(M, N)$  est définie par

$$\begin{cases} x_0 \text{ donné dans } \mathbb{R}^n, \\ Mx_{k+1} = Nx_k + b \quad \forall k \geq 1. \end{cases}$$

**Avantages:** faible stockage mémoire et faible temps CPU si convergence rapide (en peu d'itérations). **Les seules possibles si  $n$  est grand !**

**Jacobi:**  $M = \text{diag}(A)$ .

**Gauss-Seidel:**  $M =$  partie triangulaire inférieure de  $A$ .

**Gradient:**  $M = \alpha^{-1} \text{Id}$  avec  $\alpha > 0$ .

## Méthode du gradient conjugué

**Proposition 13.1.39** Soit  $A$  une matrice symétrique définie positive, et  $x_0 \in \mathbb{R}^n$ . Soit  $(x_k, r_k, p_k)$  trois suites définies par les relations de récurrence

$$p_0 = r_0 = b - Ax_0, \text{ et pour } 0 \leq k \left\{ \begin{array}{l} x_{k+1} = x_k + \alpha_k p_k \\ r_{k+1} = r_k - \alpha_k A p_k \\ p_{k+1} = r_{k+1} + \beta_k p_k \end{array} \right.$$

avec

$$\alpha_k = \frac{\|r_k\|^2}{A p_k \cdot p_k} \text{ et } \beta_k = \frac{\|r_{k+1}\|^2}{\|r_k\|^2}.$$

Alors, la suite  $(x_k)_{k \geq 0}$  converge en moins de  $n$  itérations vers la solution exacte de  $Ax = b$

Méthode la plus efficace (avec un préconditionnement pour converger plus vite).