

Méthodes de gradient multiniveau en optimisation de forme

François Beux

(Scuola Normale Superiore di Pisa)

CANUM 2006, 30 mai 2006

Mini-Symposium en Optimisation de Forme

Contexte

❖ Problèmes d'optimisation de forme en aérodynamique

$$\text{Trouver } \alpha^* \text{ tel que } j(\alpha^*) = \min_{\alpha} j(\alpha)$$

$$\text{avec } j(\alpha) = J(\alpha, W(\alpha)) \text{ et la contrainte } \Psi(\alpha, W(\alpha)) = 0$$

où α caractérise la frontière Γ de l'obstacle, W correspond aux variables physiques de l'écoulement et $\Psi=0$ correspond à un système d'équations aux dérivées partielles issu de la modélisation de l'écoulement (exemples: équations d'Euler, de Navier-Stokes, ...).

❖ Méthodes d'optimisation de type gradient

On considère un gradient discret, c'est à dire la différentiation du problème déjà discrétisé.

$$\text{Trouver } \alpha^* \in \mathbb{R}^p \text{ tel que } j_h(\alpha^*) = \min_{\alpha \in \mathbb{R}^p} j_h(\alpha) = \min_{\alpha \in \mathbb{R}^p} J_h(\alpha, W(\alpha))$$

$$\text{Algorithme itératif: } \begin{cases} \alpha^{(0)} & \text{configuration initiale} \\ \alpha^{(n+1)} & = \alpha^{(n)} - \rho_n \nabla j(\alpha^{(n)}) \end{cases}$$

Méthode de gradient multi-niveau (*Beux-Dervieux, Eng. Comp. 1994*)

- Soient E et F espaces de Hilbert, $P : F \rightarrow E$ linéaire, et j une fonctionnelle définie sur E à valeurs réelles. Au lieu de minimiser dans E , on considère le problème:

$$\text{Trouver } \bar{v} \in F \text{ tel que } j \circ P(\bar{v}) = \min_{v \in F} j \circ P(v)$$

- Utilisant une méthode de gradient pour résoudre le problème de minimisation, on obtient alors l'algorithme suivant:

$$u_0 \in E \text{ fixe, pour } k \geq 1 \quad u_{k+1} = u_k - \rho PP^* \Lambda_E j'(u_k)$$

où P^* est l'adjoint de P , Λ_E l'isomorphisme canonique entre E' et E .

Cet algorithme est une méthode de descente dans E .

Application: paramètres de contrôle discrets



Considérant m paramètres extraits de la paramétrisation complète, c'est à dire les coordonnées des points du maillage sur la frontière, et soit une application linéaire

$$\mathbb{R}^m \rightarrow \mathbb{R}^p, \text{ alors les résultats précédents peuvent être utilisés avec } E = \mathbb{R}^p \text{ et } F = \mathbb{R}^m$$

$$u_0 \in \mathbb{R}^p \text{ fixe, pour } k \geq 1 \quad u_{k+1} = u_k - \rho PP^* \text{grad}_E j(u_k)$$

Si P est associé à la matrice $M \in \mathbb{R}^{p \times m}$ alors P^* est associé à la transposé $M^T \in \mathbb{R}^{m \times p}$

Stratégie multiniveau

- On définit une famille de sous-paramétrisation
- ➔ Définition d'une stratégie pour décider quels paramètres utiliser à chaque itération 
- ➔ Choix de l'opérateur P 

- Minimiser sur un niveau grossier l correspond à remplacer le gradient par $d_l = P_l(P_l)^* \text{grad } j$
- Cela peut être interpréter comme un **préconditionnement** du gradient. En effet, à l'itération n :

$$\alpha^{n+1} = \alpha^n - \rho^n C_n \text{grad } j^n \quad C_n = M_l(M_l)^T \quad (\text{symétrique définite positive})$$

Intérêt

- **Effet de lissage**: réduit l'apparition de hautes fréquences sur la forme à optimiser typique quand on utilise une paramétrisation basée sur les noeuds du maillage.
- ▶ Manque de régularité de la dérivée d'une fonction par rapport au domaine [Courty–Dervieux]
 - **Accélération de la convergence** du processus d'optimisation (peu dépendant du nombre de paramètres de contrôle).
- ▶ Justification d'un point de vue théorique [Guillard, 1993]

Approche multiniveau et différentiation

- α : variables de contrôle caractérisant la frontière;
- La fonction objectif discrète dépend de α seulement à travers $\mathcal{L}(\alpha)$
 $\mathcal{L}(\alpha)$ étant les coordonnées des points du maillage sur la frontière discrétisée;

$$\forall \alpha \in \mathbb{R}^p \quad j(\alpha) = J(\mathcal{L}(\alpha), W(\mathcal{L}(\alpha))) = \mathcal{J}(\mathcal{L}(\alpha)) \quad \mathcal{J}(\gamma) = J(\gamma, W(\gamma))$$

Différentiation

$$\forall \delta \alpha \in \mathbb{R}^p \quad \left\langle \frac{dj}{d\alpha}, \delta \alpha \right\rangle = \left\langle \frac{\partial \mathcal{J}}{\partial \mathcal{L}}, \frac{d\mathcal{L}}{d\alpha} \delta \alpha \right\rangle = \left\langle \left(\frac{d\mathcal{L}}{d\alpha} \right)^* \frac{\partial \mathcal{J}}{\partial \mathcal{L}}, \delta \alpha \right\rangle$$

Gradient

$$\text{grad } j = \left(\frac{d\mathcal{L}}{d\alpha} \right)^* \text{grad } \mathcal{J} \quad \text{si } \mathcal{L} \text{ est linéaire alors } \text{grad } j = \mathcal{L}^* \text{grad } \mathcal{J}$$

- ❖ Si α correspond aux coordonnées d'un sous-ensemble de noeuds du maillage sur la frontière:

$$\mathcal{L} \equiv P_l \quad \text{alors le gradient sur le niveau } l \text{ est } (P_l)^* \text{grad } j$$

- Retournant sur le niveau le plus fin, on retrouve: $d_n = P_l (P_l)^* \text{grad } j_n$

Développements successifs sur le gradient multiniveau

❖ **Généralisation à écoulements Euler 3D** (*Marco-Dervieux, 1997*)

→ Basée sur techniques d'agglomération et conditions de transpiration

❖ **Multiniveau additif** (*Courty-Dervieux*)

Principe: Au lieu de changer de niveau à chaque itération d'optimisation, tous les niveaux sont balayés à chaque itération.

→ Complexité *a priori* inférieure

→ Bon comportement associé avec gradient conjugué, méthode BFGS

Gradients approchés (écoulements RANS 2D)

❖ **“One-shot “** (*Held-Dervieux, 2002*)



❖ **Gradient incomplet/complété** (*de Michieli Vitturi-Beux*)



Généralisation à paramétrisation de type polynomiale

(Martinelli-Beux)



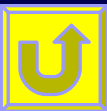
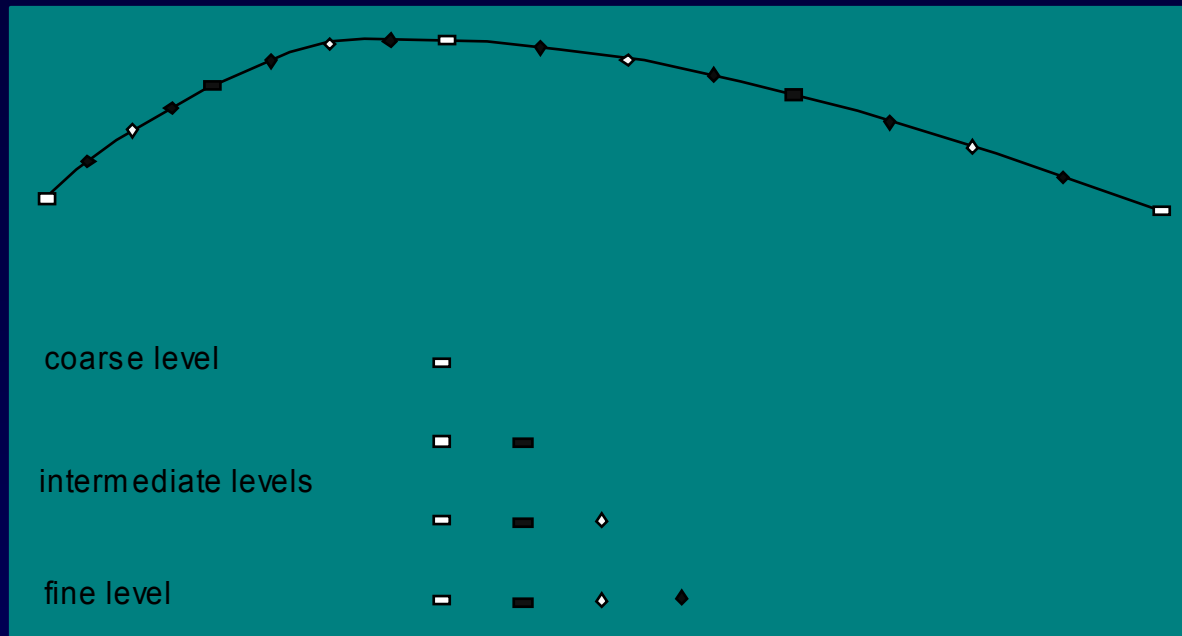
Paramétrisation hiérarchique

❖ Pour le passage d'un niveau au niveau suivant, on utilise une interpolation cubique.

$$P_{l,l+1} : \mathbb{R}^{m_l} \rightarrow \mathbb{R}^{m_{l+1}}$$

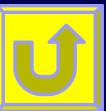
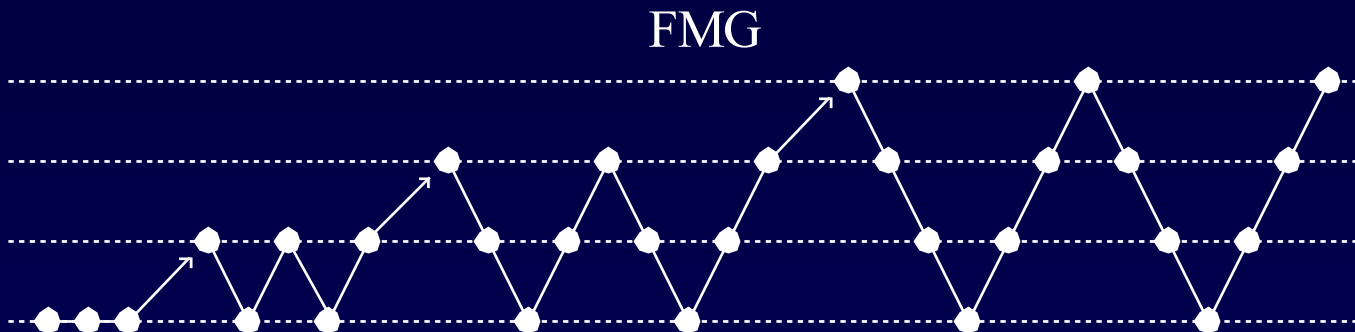
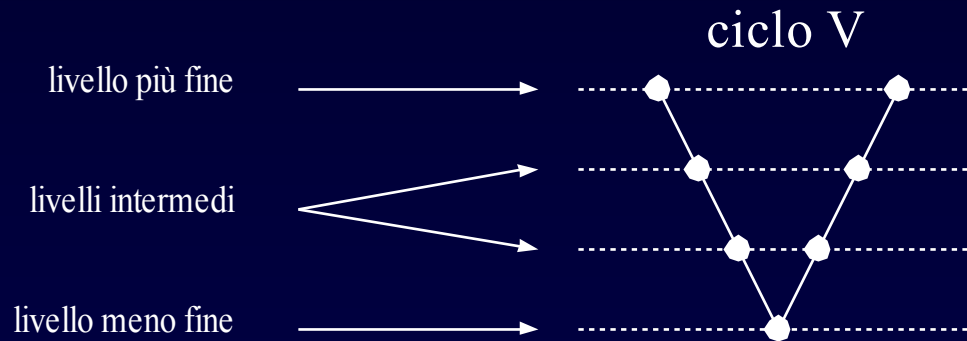
❖ Opérateur de prolongement du niveau l sur le niveau plus fin $P_l : \mathbb{R}^{m_l} \rightarrow \mathbb{R}^p$

❖ Pour obtenir P_l on itère la procédure d'interpolation $P_l = P_{L-1,L} \circ \dots \circ P_{l+1,l+2} \circ P_{l,l+1}$



Strategia multinivello

L'idea è di minimizzare alternativamente su differenti sotto-spazi di controllo utilizzando una strategia di cicli (come per le metodologie di tipo "multigrid").



Gradient approché et multiniveau

- ❖ On considère la dérivée de Fréchet de $j = \mathcal{J} \circ \mathcal{L}$ avec \mathcal{L} **linéaire**

$$j'(\alpha)(h) = \mathcal{J}'(\gamma)(\mathcal{L}'(\alpha)(h)) = \mathcal{J}'(\gamma)(\mathcal{L}(h)) = \lim_{\theta \rightarrow 0^+} \frac{j(\gamma + \theta \mathcal{L}(h)) - j(\gamma)}{\theta} \quad \text{où } \gamma = \mathcal{L}(\alpha)$$

- ❖ Approximation du gradient sur le niveau l ($\mathcal{L} \equiv P_l$)
$$\frac{j(\gamma^{(n)} + \varepsilon P_l(e_i)) - j(\gamma^{(n)})}{\varepsilon}$$

e_i : i -ième vecteur de la base canonique de \mathbb{R}^{m_l}

- ❖ Direction de descente sur le niveau fin:
$$(d_l^{(n)})_i = P_l \frac{j(\gamma^{(n)} + \varepsilon P_l(e_i)) - j(\gamma^{(n)})}{\varepsilon}$$

Perturbation de la forme seulement par rapport aux paramètres du niveau l

- ❖ Dans [**Held-Dervieux**]: couplage entre gradient approché multiniveau et approche “one-shot”

▶ On résoud simultanément les conditions d’optimalité et l’équation d’état

▶ En pratique, à chaque itération d’optimisation, le calcul de l’écoulement est effectué de façon très approximée (peu d’itérations pour la résolution des systèmes linéaires et peu d’itérations pseudo-temporelles)



Gradient incomplet

→ Si la fonctionnelle est basée sur des informations locales à proximité de la frontière (c'est en général le cas pour l'optimisation de forme en aérodynamique), le calcul du gradient est approximé considérant seulement la différentiation par rapport aux variables de type géométrique [Mohammadi, 1997]

introduisant $J : \mathbb{R}^l \times \mathbb{R}^m \times \mathbb{R}^n \rightarrow \mathbb{R}$ tel que:

$$\forall \alpha \in \mathbb{R}^p \quad j(\alpha) = J(\mathcal{L}(\alpha), q(\mathcal{L}(\alpha)), W(\mathcal{L}(\alpha)))$$

- $\mathcal{L}(\alpha)$: coordonnées des points du maillage sur la frontière;
- $q(\mathcal{L}(\alpha))$: coordonnées de tous les points du maillage;
- $W(\mathcal{L}(\alpha))$: variables physiques de l'écoulement.

Différentiation: $\left\langle \frac{dj}{d\alpha}, \delta\alpha \right\rangle = \underbrace{\left\langle \frac{\partial J}{\partial \mathcal{L}}, \frac{d\mathcal{L}}{d\alpha} \delta\alpha \right\rangle + \left\langle \frac{\partial J}{\partial q}, \frac{dq}{d\mathcal{L}} \frac{d\mathcal{L}}{d\alpha} \delta\alpha \right\rangle}_{\text{gradient incomplet}} + \left\langle \frac{\partial J}{\partial W}, \frac{dW}{d\mathcal{L}} \frac{d\mathcal{L}}{d\alpha} \delta\alpha \right\rangle$

Gradient complété [de Michieli Vitturi-Beux]

- L'idée est, partant d'une formulation gradient incomplet associée à une approche multiniveau, de chercher à obtenir une approximation du terme négligé mettant à profit les niveaux grossiers de l'approche multiniveau.

Gradient multiniveau complété

Direction de descente à l'itération r : $d^{(r)} = P_{l_r} P_{l_r}^* d_I^{(r)} + P_{l_r} \delta_F^{(r)}$

avec $\mathcal{L} = P_{l_r}$ opérateur de prolongement du niveau l_r au niveau plus fin

$$[d_I]_i^{(r)} = \frac{\partial J^{(r)}}{\partial \mathcal{L}_i} + \sum_{j=1}^m \frac{\partial J^{(r)}}{\partial q_j} \frac{\partial q_j}{\partial \mathcal{L}_i},$$

$$[\delta_F]_i^{(r)} = \sum_{j=1}^N \frac{\partial J^{(r)}}{\partial W_j} \frac{\delta W_j^{(r)}}{\delta \alpha_i} \quad \text{avec} \quad \frac{\delta W_j^{(r)}}{\delta \alpha_i} = \frac{W_j(\gamma^{(r)} + \varepsilon P_{l_r}(e_i)) - W_j(\gamma^{(r)})}{\varepsilon}$$

➤ Pour limiter encore le coût du calcul, cette évaluation est faite seulement sur le niveau le plus grossier:

$$\gamma^{(n+1)} = \gamma^{(n)} - \rho^{(n)} \left[P_{l_n} P_{l_n}^* d_I^{(n)} + P_{l_n} \tilde{\delta}_F^{(r,n)} \right] \quad \text{avec} \quad \tilde{\delta}_F^{(r,n)} = \frac{\partial J^{(n)}}{\partial W} \left[P_{(r \rightarrow n)} \frac{\delta W^{(r)}}{\delta \alpha} \right]$$

$P_{(r \rightarrow n)}$ l'opérateur de prolongement du niveau l_r au niveau l_n



Approche multiniveau et courbes de Bézier

(M. Martinelli - F. Beux)

- Courbe de Bézier de degré n : représentation paramétrique

$$\begin{cases} x(t) = \sum_{q=0}^n B_n^q(t) x_q \\ y(t) = \sum_{q=0}^n B_n^q(t) y_q \end{cases} \quad \begin{array}{l} (x_q, y_q) : \text{ points de contrôle de Bézier} \\ B_n^q(t) : \text{ polynômes de Bernstein.} \end{array}$$

- variables de contrôle: ordonnées des points

$$\begin{cases} \alpha = (y_0, \dots, y_n)^T \\ \gamma = (y(t_0), \dots, y(t_m))^T \end{cases}$$

- fixant les paramètres $0 = t_0 \leq t_1 \leq \dots \leq t_{m-1} \leq t_m = 1$

on peut alors définir une application linéaire pour passer de α à γ

$$\begin{array}{l} P: F = \mathbb{R}^{n+1} \rightarrow E = \mathbb{R}^{m+1} \\ \alpha \rightarrow \gamma = M \alpha \end{array} \quad \text{où} \quad M_{ij} = B_n^j(t_i)$$

- ❖ On peut considérer une approche multiniveau prenant comme sous-niveau, au lieu d'un sous-ensemble de paramètres extraits de la paramétrisation complète, un ensemble de points de contrôle de Bézier.

Algorithme

1. Construction d'une paramétrisation adéquate au niveau l

Trouver $(t_0^{(l)}, \dots, t_m^{(l)})$ et $X^{(l)} = (x_0^{(l)}, \dots, x_{n_l}^{(l)})$ tels que $x_k^S = x(t_k^{(l)}) = \sum_{q=0}^{n_l} B_{n_l}^q(t_k^{(l)}) x_q^{(l)}$ $k = 0, \dots, m$

➤ On choisit $t_0^{(0)}, \dots, t_m^{(0)}$ e $X^{(0)}$ compatibles

➤ La distribution des paramètres t ne varie pas par **élévation de degré**



➤ $t_i^{(l)} = t_i^{(0)}$ pour $i = 0, \dots, m$

➤ $X^{(l)}$ est obtenu à partir de $X^{(0)}$ appliquant successivement l'élévation de degré

1. Direction de descente au niveau l

(i-ième composante)

$$\left(d_r^{(l)} \right)_i = \sum_{j=0}^m \underbrace{\sum_{q=0}^{n_l} B_{n_l}^q(t_i) B_{n_l}^q(t_j)}_{\left(M^{(l)} \left(M^{(l)} \right)^T \right)_{ij}} \left(g_r \right)_j$$

Réinterprétation de l'algorithme

• L'opérateur pour passer du niveau plus fin des points de contrôle de Bézier aux ordonnées des points du maillage sur la frontière

$$P^{(L)} : \mathbb{R}^{n_L} \rightarrow \mathbb{R}^{m+1}$$

$$\alpha \rightarrow M^{(L)} \alpha \quad \text{avec} \quad M_{ij}^{(L)} = B_{n_L}^j(t_i)$$

➤ Algorithme $\gamma_{r+1} = \gamma_r - \rho_r P^{(L)} p_r^{(l)}$ avec $p_r^{(l)} = P_l^L \left(P_l^L \right)^* \left(P^{(L)} \right)^* g_r$



Courbe de Bézier: élévation de degré



- Une courbe de Bézier de degré n associée aux $n+1$ points de contrôle $P_q = (x_q, y_q)^T$ peut être vue comme une courbe de Bézier de degré $n+1$ associée aux $n+2$ points de contrôle

$$P(t) = (x(t), y(t))^T = \sum_{q=0}^{n+1} B_{n+1}^q(t) P'_q$$

$$P'_q = (x'_q, y'_q)^T \quad \text{défini par} \quad P'_q = \frac{q}{n+1} P_{q-1} + \left(1 - \frac{q}{n+1}\right) P_q \quad q = 0, \dots, n+1$$

- ❖ On peut donc passer d'un ensemble de $n+1$ points de contrôle de Bézier à un ensemble de $n+2$ points de contrôle utilisant la matrice

$$D_n^{n+1} \quad \text{définie par} \quad (D_n^{n+1})_{ij} = \left(1 - \frac{i}{n+1}\right) \delta_{ij} + \frac{i}{n+1} \delta_{(i-1)j}$$

- ❖ **Remarque:** dans [Désidéri, 2003] on considère une approche multiniveau basée sur la **procédure d'élévation de degré** pour changer de niveau (les variables de contrôle correspondent à la variation des ordonnées des points de contrôle de Bézier).
- ❖ Cette approche a été conçue hors du contexte des méthodes de gradient.

Courbe de Bézier: élévation de degré



❖ On peut passer d'un ensemble de $n + 1$ points de contrôle de Bézier à un ensemble de $n + 2$ points de contrôle utilisant la matrice

$$D_n^{n+1} \text{ définie par } (D_n^{n+1})_{ij} = \left(1 - \frac{i}{n+1}\right) \delta_{ij} + \frac{i}{n+1} \delta_{(i-1)j}$$

• **Stratégie**: considérons les ordonnées des points de contrôle de Bézier comme paramètres de contrôle et la **procédure d'élévation de degré** pour effectuer le changement de niveau dans l'approche multiniveau

➤ L'opérateur de prolongement d'un niveau p à un niveau q

$$P_p^q : F = \mathbb{R}^{n_p} \rightarrow E = \mathbb{R}^{n_q}$$

$$\alpha \rightarrow D_{n_p}^{n_q} \alpha \quad \text{ou } D_n^{n+s} = D_{n+s-1}^{n+s} \cdots D_{n+1}^{n+2} D_n^{n+1}$$

• Prolongement du niveau l au niveau plus fin L : $P_l^L = P_{L-1}^L \circ P_{L-2}^{L-1} \circ \cdots \circ P_{l+1}^{l+2} \circ P_l^{l+1}$

Algorithme $\alpha_{r+1} = \alpha_r - \rho_r C_r \mathcal{G}_r$ avec $C_r = D_{n_l}^{n_L} (D_{n_l}^{n_L})^T$

❖ $\mathcal{G}_r \neq g_r$ \mathcal{G}_r correspond au gradient par rapport aux points de contrôle de Bézier

Réinterprétation de la seconde approche multiniveau



- La fonction objectif discrète dépend de α seulement à travers $\mathcal{L}(\alpha)$
 $\mathcal{L}(\alpha)$ étant les coordonnées des points du maillage sur la frontière discrétisée;

$$\forall \alpha \in \mathbb{R}^p \quad j(\alpha) = J(\mathcal{L}(\alpha), W(\mathcal{L}(\alpha))) = \mathcal{J}(\mathcal{L}(\alpha)) \quad \mathcal{J}(\gamma) = J(\gamma, W(\gamma))$$

Différentiation

$$\forall \delta \alpha \in \mathbb{R}^p \quad \left\langle \frac{dj}{d\alpha}, \delta \alpha \right\rangle = \left\langle \frac{\partial \mathcal{J}}{\partial \mathcal{L}}, \frac{d\mathcal{L}}{d\alpha} \delta \alpha \right\rangle = \left\langle \left(\frac{d\mathcal{L}}{d\alpha} \right)^* \frac{\partial \mathcal{J}}{\partial \mathcal{L}}, \delta \alpha \right\rangle$$

Gradient

$$\text{grad } j = \left(\frac{d\mathcal{L}}{d\alpha} \right)^* \text{grad } \mathcal{J} \quad \text{si } \mathcal{L} \text{ est linéaire alors } \text{grad } j = \mathcal{L}^* \text{grad } \mathcal{J}$$

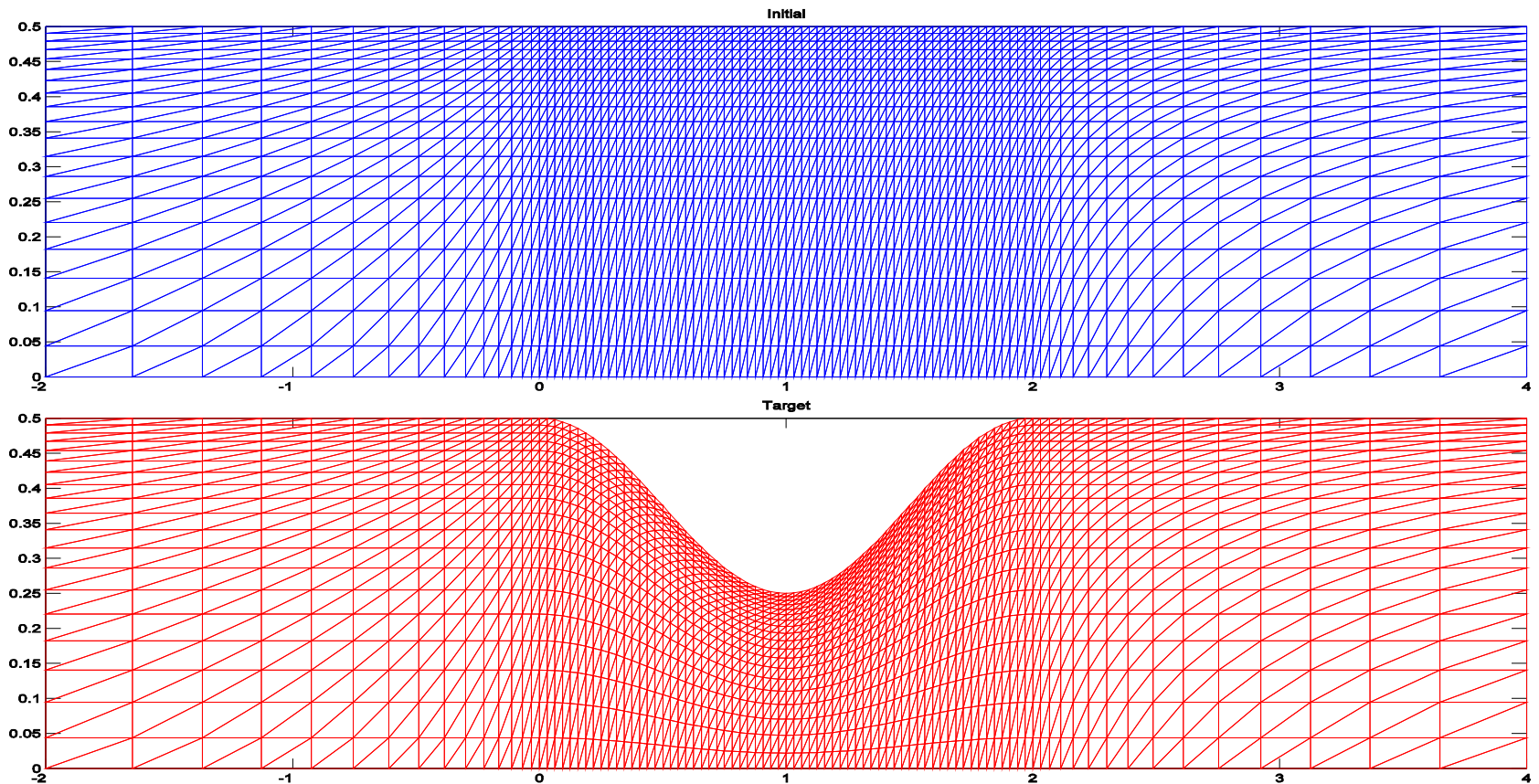
➤ Dans notre cas: $\mathcal{L} \equiv P^{(L)}$ donc $G_r = \left(P^{(L)} \right)^* g_r$

➤ Algorithme multiniveau: $\alpha_{r+1} = \alpha_r - \rho_r \underbrace{P_l^L \left(P_l^L \right)^*}_{P_r^{(l)}} G_r$

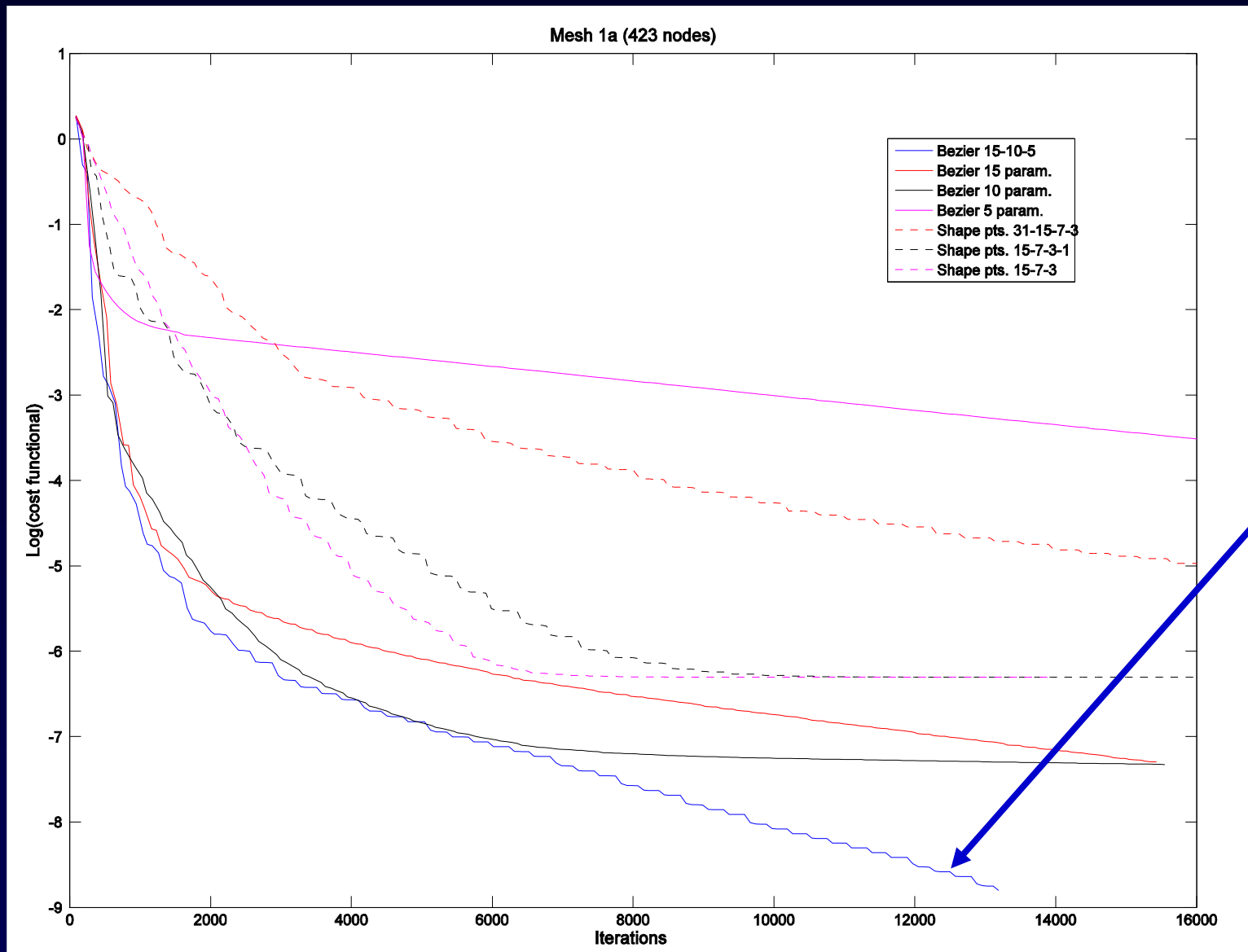
➤ Les ordonnées des points du maillage sur la frontière sont alors ajournées $\gamma_{r+1} = P^{(L)} \alpha_{r+1} = P^{(L)} \left(\alpha_r - \rho_r P_r^{(l)} \right) = \underbrace{P^{(L)} \alpha_r}_{\gamma_r} - \rho_r P^{(L)} P_r^{(l)}$

Cas test: problème inverse (tuyère)

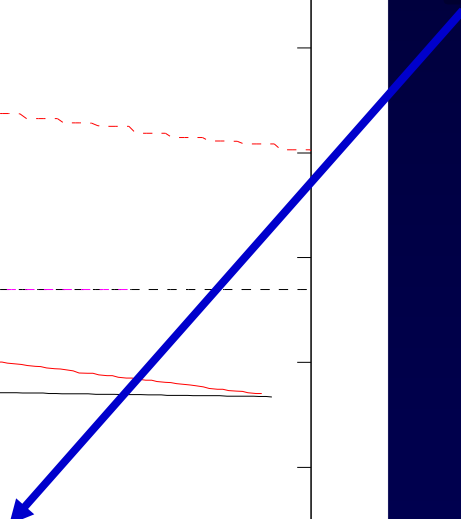
- ❖ écoulement compressible subsonique ($M_\infty = 0.3$) non visqueux
- ❖ Stratégie multiniveau: 3 niveaux (5-10-15 paramètres)
- ❖ Calcul du gradient: exact et approché (avec “one-shot”)
- ❖ maillages: 423 et 1900 éléments



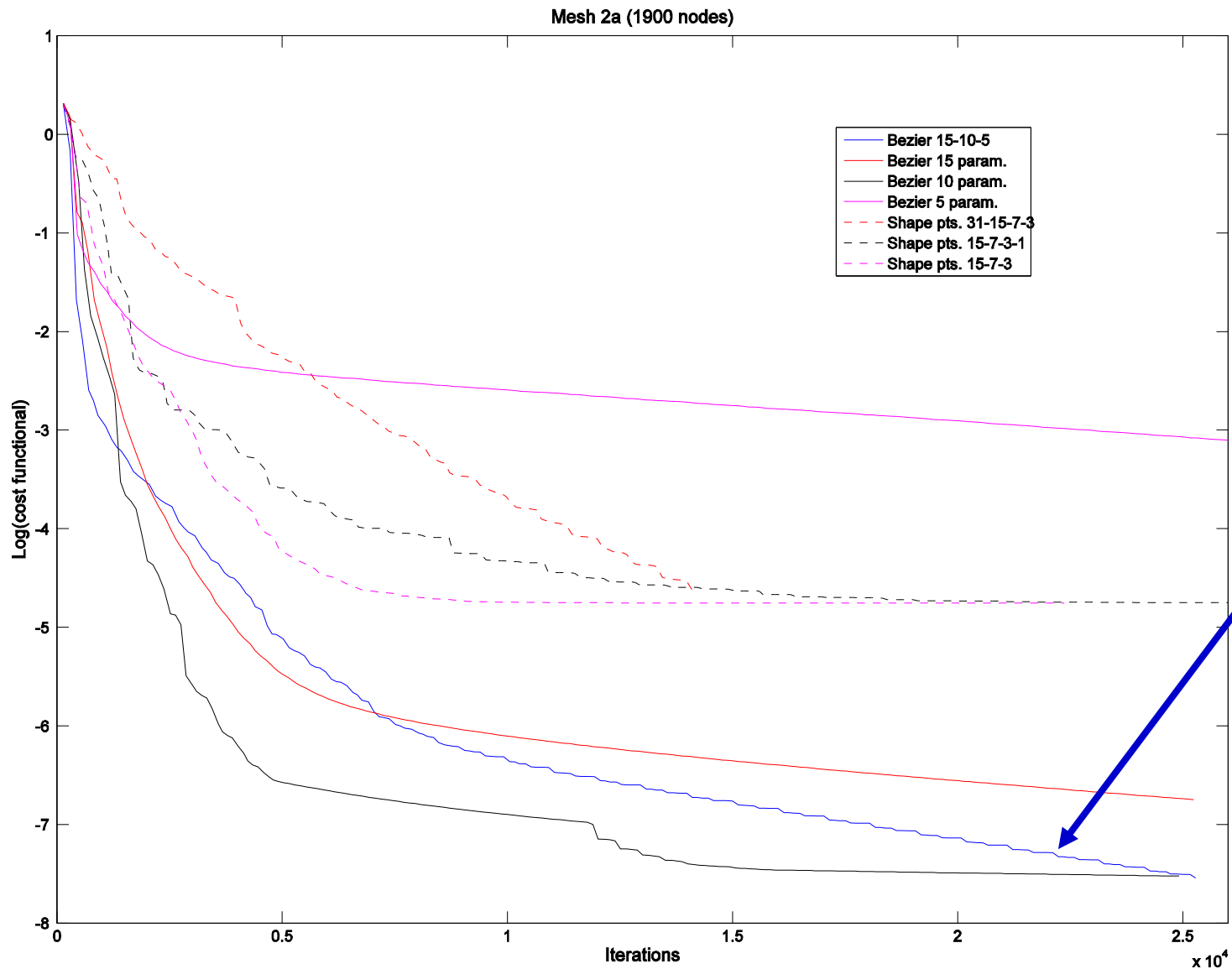
Courbes de convergence (maillage grossier)



15-10-5



Courbes de convergence (maillage fin)

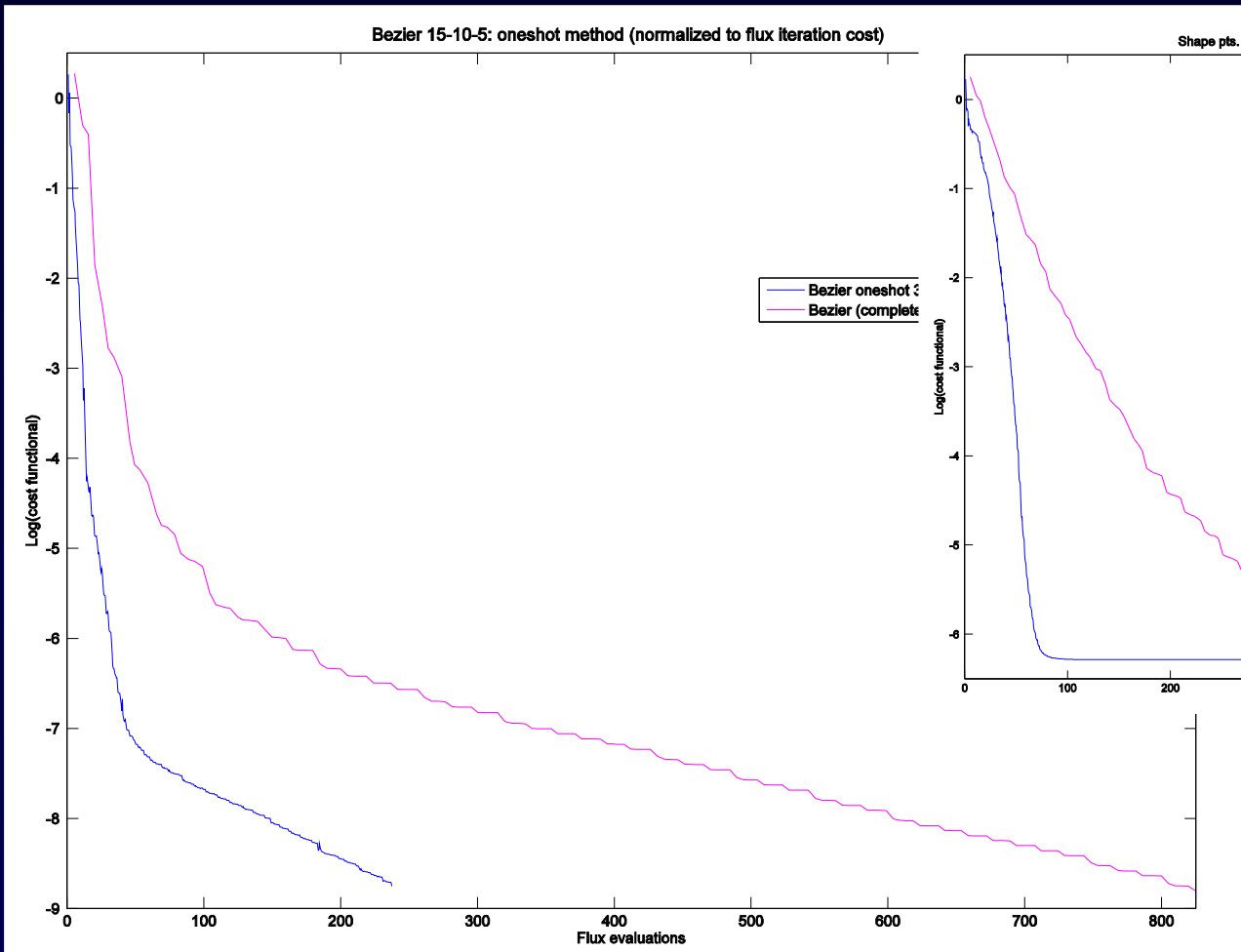


15-10-5

Comparaison gradient exact – “one-shot”

Bézier sur 3 niveaux: 15-10-5 param.

Points maillage sur 3 niveaux: 15-7-3 param.



Gradient exact

Gradient approché + one-shot.

(1 itér. Pseudo-temporelle + 3 itér. Jacobi)

Développements possibles

- ❖ Compléter la validation.
- ❖ Couplage avec une formulation gradient “incomplet-complété”
- ❖ Généralisation à d’autres types de paramétrisation:
 - ❖ B-splines
 - propriété d’élévation de niveau
 - matrices de passage entre B-splines et courbes de Bézier
 - ❖ (*Désidéri, 2003*) approche multiniveau basée sur l’élévation de degré, mais hors du contexte des méthodes de gradient (simplex, algorithmes génétiques)
 - possibilité de définir une approche hybride ?