## Chapitre 12

# Méthode CESTAC sur un algorithme itératif

## 1 Introduction :

Soit un problème mathématique (P) avec une solution $x^*$ ; la résolution de ce problème par une méthode itérative est la construction d'une suite $x_n$ $n = 1$, 2, ... à partir d'un choix initial $x_0$ qui doit converger vers $x^*$ avec une relation de la forme $\mathcal{F}(x^*) = 0$.
Sur machine on a une approximation de cette solution, c'est-à-dire arrêter le calcul de la suite $x_n$ au bout d'un nombre fini $N$ d'itérations et de prendre $x_N$ comme solution approchée de $x^*$.
Les tests d'arrêt classiques des itérations sont :
Soient une valeur $\epsilon$ donnée a priori et une norme dans l'espace vectoriel considéré.

$$||x_n - x_{n-1}|| < \epsilon$$
$$||x_n - x_{n-1}|| < \epsilon||x_n||$$
$$||\mathcal{F}(x_n)|| < \epsilon$$

Ces tests sont très robustes pour une arithmétique exacte (précision infinie) mais pas sur machine car les $x_n$ et $\mathcal{F}(x_n)$ sont entachés d'erreurs de calcul (d'arrondi).
Soit $\mathbb{F}$ l'ensemble des flottants. Soit F une procédure de calcul de $\mathcal{F}$.l L'exécution d'un algorithme itératif donne pour image de $x_1, x_2, ..., x_n$, $X_1, X_2, ..., X_n$ associée à $\mathbb{F}$ et on n'a pas les équivalences entre les tests

$$||x_n - x_{n-1}|| < \epsilon \text{ et } ||X_n - X_{n-1}|| < \epsilon$$

$$||\mathcal{F}(x_n)|| < \epsilon \text{ et } ||\mathrm{F}(X_n)|| < \epsilon$$

Si on choisit une valeur de $\epsilon$ trop grande, les itération peuvent s'arrêter avant d'avoir obtenu la meilleure solution numérique possible, et si elle est trop petite on peut avoir :

- un grand nombre d'itérations avant l'arrêt de l'algorithme sans améliorer la solution.
- l'algorithme peut ne pas s'arrêter ou donner un itéré très loin de la solution cherchée par accumulation d'erreurs d'arrondi.

Il faut donc arrêter correctement les itérations, c'est-à-dire dès qu'une solution numérique est informatiquement satisfaisante pour le problème.

Définition :
Soit $R \in \mathbb{F}$ la valeur du résultat d'un calcul (qui peut être entaché d'erreurs) et $C$ le nombre de chiffres significatifs de $R$.
$R$ est un zéro informatique, noté $R = \underline{0}$ si :

$R = 0$ et $C \geq 1$ (0 significatif)

ou $R$ quelconque et $C < 1$ (non significatif)

La méthode CESTAC permet de calculer le nombre de chiffres significatifs $C$ pour un résultat numérique $R$ et donc de détecter le zéro informatique.

On peut construire à partir de cette définition un test d'arrêt optimal.
Ce test consiste à chaque itération $n$ à calculer la moyenne $\mathrm{F}(X_n)$ et le nombre de chiffres significatifs $C(\mathrm{F}(X_n))$ par la méthode CESTAC et de faire $\mathrm{F}(X_n) = \underline{0}$

# 2 Utilisation asynchrone de la méthode CESTAC

Elle consiste à faire exécuter trois fois de suite la même procédure itérative de calcul avec perturbations aléatoires. Le test d'arrêt optimal permet à chaque fois d'arrêter les itérations et les trois limites obtenues successivement doit permettre d'évaluer la précision sur la solution théorique.

Remarques :
Les trois suites peuvent être très voisines et produire une surestimation de la précision ou bien converger vers des limites différentes dans le cas d'un problème avec plusieurs solutions (calcul de racines de polynômes) et la comparaison de solutions différentes n'a plus de sens théoriquement ou bien si l'algorithme contient des tests avec les perturbations aléatoires les trois calculs peuvent se faire dans des branches différentes et l'estimation de la précision devient erronée.

# 3 Utilisation synchrone de la méthode CESTAC

On peut faire les trois calculs en parallèle de la suite donnée de manière synchrone :
Prenons l'itération $x_{k+1} = \varphi(x_k)$ et son image informatique $X_{k+1} = \phi(X_k)$.
On calcule simultanément les trois résultats et le nombre de chiffres significatifs issus de la méthode Cestac pour $X_1^{(i)} = \phi(X_0) \quad i = 1, 2, 3$ et $C(\overline{X}_1), \cdots, X_{k+1}^{(i)} = \phi(X_k^i) \quad i = 1, 2, 3$ et $C(\overline{X}_{k+1})$
où $C(\overline{X})$ est le nombre de chiffres significatifs calculé sur la moyenne $\overline{X}$ des trois représentants aléatoires $X^{(1)}$, $X^{(2)}$, $X^{(3)}$ d'une même valeur théorique, calculés avec les perturbations aléatoires de la méthode CESTAC.
Les itérations sont arrêtées de la manière suivante :

1. A une itération $k$ on a $C(\overline{X}_k) < 1$ (la valeur de $\overline{X}_k$ est non significative : les calculs de $\phi(X_{k-1}^{(i)})$ sont entachés d'une erreur supérieure à leur propre valeur et il est donc inutile de poursuivre le calcul).
2. A une itération $k$ on a pour au moins une valeur de $i$, $||X_k^{(i)} - \phi(X_{k-1}^{(i)})|| = \underline{0}$ (la différence entre un itéré et le suivant ne représente que des erreurs de calcul).
3. Si la suite $x_{k+1} = \varphi(x_k)$ converge mathématiquement vers la solution d'un problème tel qu'une fonction $\mathcal{F}$ s'annule à la limite et si pour une itération $k$ avec F image de $\mathcal{F}$ on a $\mathrm{F}(X_k) = \underline{0}$, les itérations doivent s'arrêter car la solution obtenue est informatiquement satisfaisante (cas du test d'arrêt optimal)

4. Si le nombre d'itérations dépasse une valeur $k_{max}$ fixée à l'avance, la suite informatique est considérée comme non convergente et on arrête les itérations.

Le cas 3 donne une solution informatique satisfaisante.

Remarques : (Comparaison avec l'utilisation asynchrone)

— Le temps de calcul est moindre.

— L'instabilité d'une suite est détectée car si les erreurs de calculs sont suffisantes pour que les trois suites calculées en parallèle convergent vers des valeurs distinctes sans chiffres significatifs en commun ; c'est le cas 1.

— Le calcul à chaque itération $k$ de trois valeurs distinctes $X_{k-1}^{(i)}$ évite la surestimation de la précision.