

# Méthodes de Résolution Numérique des Équations

## Chapitre 1 : Introduction et Méthodes de Base

La résolution d'équations, notamment non linéaires, est une tâche fondamentale en sciences et ingénierie. Lorsque les solutions analytiques ne sont pas disponibles ou pratiques à obtenir, les méthodes numériques offrent des outils puissants pour approximer ces solutions. Ce chapitre introduit les concepts de base relatifs à la recherche de racines d'équations et présente plusieurs méthodes numériques fondamentales, telles que la dichotomie et la méthode de la fausse position. Nous aborderons également les notions d'existence, d'unicité des solutions, ainsi que la convergence de ces algorithmes.

## 1 Généralités et exemples

### 1.1 Rappels ((EO), $f(x)=0$ )

Une équation est une instruction mathématique qui stipule que deux expressions sont égales. La résolution d'une équation consiste à trouver les valeurs des inconnues pour lesquelles cette égalité est vérifiée. Dans ce chapitre, nous nous intéressons aux équations de la forme :

$$f(x) = 0$$

où  $f$  est une fonction définie sur un domaine  $D \subseteq \mathbb{R}$  à valeurs dans  $\mathbb{R}$ . On cherche  $x \in D$  tel que  $f(x) = 0$ . Une telle équation est parfois notée (EO).

**Definition 1.1** (Équation linéaire). Si la fonction  $f$  est affine, c'est-à-dire de la forme  $f(x) = ax + b$  avec  $a, b \in \mathbb{R}$ , l'équation  $f(x) = 0$  est appelée **équation linéaire**. Si  $a \neq 0$ , la solution unique est  $x = -b/a$ . Si  $a = 0$  et  $b \neq 0$ , il n'y a pas de solution. Si  $a = 0$  et  $b = 0$ , tout  $x \in \mathbb{R}$  est solution. Les équations qui ne sont pas linéaires sont dites **non linéaires**.

**Definition 1.2** (Solution ou Racine). Tout  $x^* \in D$  tel que  $f(x^*) = 0$  est appelé une **solution** de l'équation (EO). Une solution est également appelée une **racine** ou un **zéro** de la fonction  $f$ . L'ensemble de définition de  $f$  est noté  $D_f$ .

### 1.2 Définitions

Considérons une fonction  $f$  de classe  $\mathcal{C}^p(D, \mathbb{R})$  pour  $p \geq 1$ , où  $D$  est un intervalle de  $\mathbb{R}$ . Soit  $x^* \in D$  une racine de  $f$ , i.e.  $f(x^*) = 0$ .

**Definition 1.3** (Racine simple). La racine  $x^*$  est dite **simple** si  $f'(x^*) \neq 0$ .

**Definition 1.4** (Racine multiple). La racine  $x^*$  est dite de **multiplicité**  $p$  (où  $p \geq 1$  est un entier) si

$$f(x^*) = f'(x^*) = \dots = f^{(p-1)}(x^*) = 0 \quad \text{et} \quad f^{(p)}(x^*) \neq 0.$$

Une racine simple est donc une racine de multiplicité  $p = 1$ .

### 1.3 Exemples

Certains problèmes mènent à la résolution d'équations non linéaires. Par exemple, dans la résolution numérique d'équations différentielles ordinaires (EDO) de la forme  $x'(t) = f(t, x(t))$ , les schémas implicites nécessitent de résoudre une équation à chaque pas de temps.

**Exemple 1.5** (Euler implicite). Le schéma d'Euler implicite pour l'EDO  $x'(t) = f(t, x(t))$  est donné par :

$$x_{n+1} = x_n + hf(t_{n+1}, x_{n+1})$$

où  $x_n$  est une approximation de  $x(t_n)$ ,  $h$  est le pas de temps, et  $t_{n+1} = t_n + h$ . Pour déterminer  $x_{n+1}$  à l'itération  $n$ , il faut résoudre l'équation  $g_n(z) = 0$  pour  $z$ , où

$$g_n(z) = z - x_n - hf(t_{n+1}, z).$$

Cette équation est généralement non linéaire en  $z \equiv x_{n+1}$ .

D'autres schémas numériques sont des exemples de suites récurrentes :

**Exemple 1.6** (Schéma de Crank-Nicolson type (CN)). Soit  $x_0 = x_0^*$  (valeur initiale). La suite est définie par :

$$x_{n+1} = x_n + \frac{h}{2}[f(t_n, x_n) + f(t_{n+1}, x_{n+1}^*)] \quad \text{pour } n = 0, \dots, N-1.$$

La valeur  $x_{n+1}^*$  doit être définie. Si  $x_{n+1}^* = x_{n+1}$ , le schéma est implicite (Crank-Nicolson standard) et nécessite la résolution d'une équation pour  $x_{n+1}$ . Si  $x_{n+1}^*$  est une prédiction explicite de  $x_{n+1}$ , le schéma peut devenir explicite.

**Exemple 1.7** (Schéma de Heun). Ce schéma est un exemple de méthode de prédicteur-correcteur :

$$\begin{aligned} x_0 &= x_0^* \quad (\text{valeur initiale}) \\ x_{n+1}^* &= x_n + hf(t_n, x_n) \quad (\text{prédicteur}) \\ x_{n+1} &= x_n + \frac{h}{2}[f(t_n, x_n) + f(t_{n+1}, x_{n+1}^*)] \quad (\text{correcteur}) \end{aligned}$$

Ce schéma, tel qu'écrit, est explicite et ne requiert pas la résolution d'une équation pour  $x_{n+1}$  à chaque étape, car  $x_{n+1}^*$  est calculé en premier.

## 2 Partition Correcte du Problème (EO)

Lorsqu'on aborde la résolution de l'équation  $f(x) = 0$ , les premières questions qui se posent sont :

- L'équation (EO) admet-elle une solution ? (Existence)
- Si oui, cette solution est-elle unique ? (Unicité)

De plus, la dépendance continue de la solution par rapport aux données du problème (stabilité) est une propriété souhaitable. Pour simplifier, nous nous plaçons dans le cadre scalaire où  $f : \mathbb{R} \rightarrow \mathbb{R}$ .

### 2.1 Existence et unicité de la solution

**Proposition 2.1** (Cas  $f(x) = 0$ ). Soit  $f$  une fonction continue sur un intervalle fermé  $I = [a, b]$ .

- Si  $f(a)f(b) < 0$  (c'est-à-dire si  $f(a)$  et  $f(b)$  sont de signes opposés), alors il existe au moins une racine  $x^* \in ]a, b[$  telle que  $f(x^*) = 0$ . (Ceci découle du Théorème des Valeurs Intermédiaires).
- Si de plus  $f$  est strictement monotone sur  $I$ , alors cette racine  $x^*$  est unique dans  $I$ .

*Preuve (esquisse) :* L'existence vient du Théorème des Valeurs Intermédiaires. Pour l'unicité, si  $f$  est strictement monotone, elle est injective sur  $I$ . Donc, si  $x_1 \neq x_2$  dans  $I$ , alors  $f(x_1) \neq f(x_2)$ . Ainsi,

$f(x) = 0$  ne peut avoir plus d'une solution.

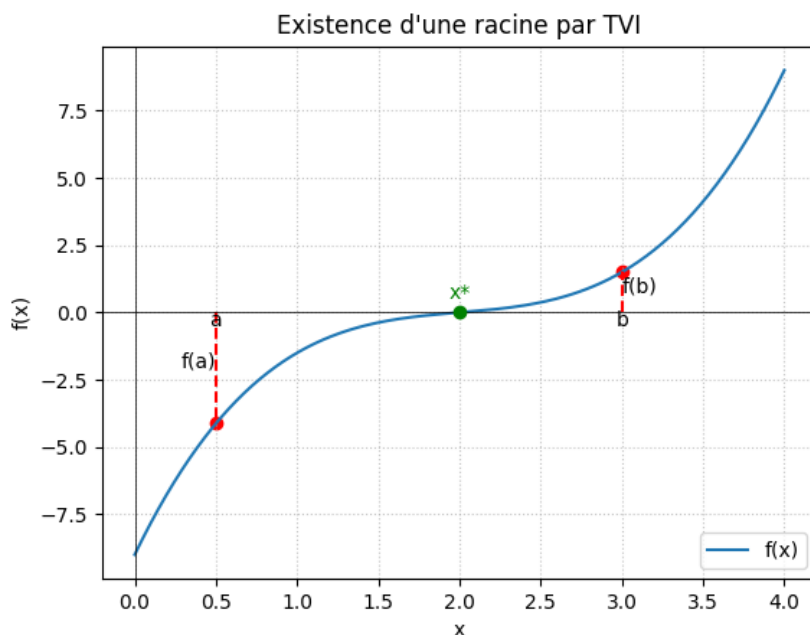


Figure 1: Illustration du Théorème des Valeurs Intermédiaires : si  $f(a)f(b) < 0$  et  $f$  est continue, il existe  $x^* \in ]a, b[$  tel que  $f(x^*) = 0$ .

Un autre type courant d'équation est l'équation de point fixe  $x = g(x)$ .

**Proposition 2.2** (Cas  $x = g(x)$ ). Soit  $g$  une fonction continue et  $I$  un intervalle fermé de  $\mathbb{R}$  tel que  $g(I) \subseteq I$  (c'est-à-dire que  $g$  envoie  $I$  dans lui-même).

- Alors il existe au moins un point fixe  $x^* \in I$  tel que  $g(x^*) = x^*$ .
- Si de plus  $g$  est contractante sur  $I$ , c'est-à-dire s'il existe  $K \in [0, 1)$  tel que pour tous  $x, y \in I$ ,  $|g(x) - g(y)| \leq K|x - y|$ , alors le point fixe  $x^*$  est unique dans  $I$ .
- Si  $g$  est dérivable sur  $I$  et s'il existe  $K \in [0, 1)$  tel que  $|g'(x)| \leq K$  pour tout  $x \in I$ , alors  $g$  est contractante sur  $I$  (par le théorème des accroissements finis).

## 2.2 Propriétés utiles

**Remark 2.3** (Multiplicité des racines). Lorsque la racine cherchée  $x^*$  est de multiplicité  $p > 1$  (racine multiple), des difficultés numériques peuvent apparaître. Par exemple, si  $f'(x^*) = 0$ , la pente de la fonction est nulle à la racine, ce qui peut ralentir la convergence de certaines méthodes ou les rendre instables. Numériquement, il sera souvent plus difficile de localiser précisément une racine multiple qu'une racine simple.

### 3 Construction de schémas pour (EO)

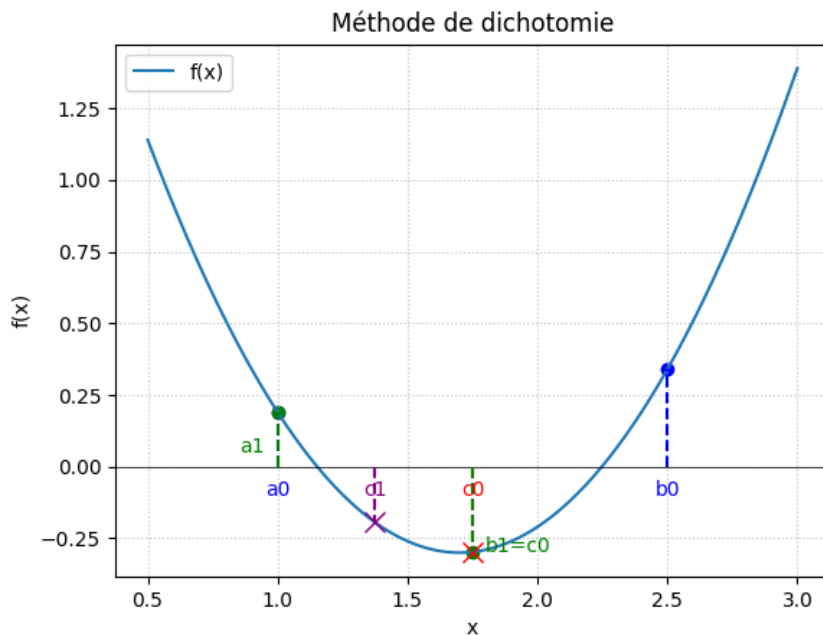
#### 3.1 Méthode de dichotomie (cas $f(x) = 0$ )

La méthode de dichotomie (ou méthode de bisection) est l'une des méthodes les plus simples et robustes pour trouver une racine d'une fonction continue.

**Principe :** On part d'un intervalle initial  $I_0 = [a_0, b_0]$  tel que  $f(a_0)f(b_0) < 0$ . On est donc assuré par le TVI qu'il existe au moins une racine dans  $]a_0, b_0[$ . À chaque étape  $n$ , on considère l'intervalle  $[a_n, b_n]$ :

1. On calcule le milieu de l'intervalle :  $c_n = \frac{a_n + b_n}{2}$ .
2. On évalue  $f(c_n)$ .
3. Si  $f(c_n) = 0$  (ou est numériquement proche de zéro), alors  $c_n$  est la racine (ou une bonne approximation).
4. Sinon :
  - Si  $f(a_n)f(c_n) < 0$ , la racine se trouve dans  $[a_n, c_n]$ . On pose alors  $[a_{n+1}, b_{n+1}] = [a_n, c_n]$ .
  - Si  $f(c_n)f(b_n) < 0$ , la racine se trouve dans  $[c_n, b_n]$ . On pose alors  $[a_{n+1}, b_{n+1}] = [c_n, b_n]$ . (Note: si  $f(a_n)f(c_n) > 0$  et  $f(c_n) \neq 0$ , alors  $f(c_n)f(b_n)$  doit être  $< 0$  car  $f(a_n)$  et  $f(b_n)$  sont de signes opposés).

On répète ce processus, divisant la taille de l'intervalle de recherche par deux à chaque étape.



- Initialiser le compteur d'itérations  $k = 0$ .

2. **Itérations** : Tant que  $(b - a) > \varepsilon$  (ou  $(b - a)/2 > \varepsilon$ ) et  $k < k_{\max}$  :

- Calculer  $c = (a + b)/2$ .
- Si  $f(c) = 0$  (ou  $|f(c)| < \varepsilon_{\text{res}}$ , une tolérance sur le résidu), alors  $x^* = c$ , arrêter.
- Si  $f(a)f(c) < 0$ , alors  $b = c$ .
- Sinon (si  $f(c)f(b) < 0$ ), alors  $a = c$ .
- $k = k + 1$ .

3. **Résultat** : La racine est approximée par  $c = (a + b)/2$ . L'erreur absolue  $|x^* - c|$  est inférieure à  $(b - a)/2$ .

Les critères d'arrêt peuvent varier, par exemple : la largeur de l'intervalle  $(b_N - a_N) \leq \varepsilon$ , la valeur de la fonction  $|f(c_N)| \leq \varepsilon_f$ , ou un nombre maximal d'itérations atteint.

#### Propriétés :

- **Convergence** : La méthode de dichotomie converge toujours si  $f$  est continue et  $f(a_0)f(b_0) < 0$ . L'erreur à l'étape  $n$  (en prenant  $c_n$  comme approximation de  $x^*$ ) est bornée :

$$|x^* - c_n| \leq \frac{b_n - a_n}{2} = \frac{b_0 - a_0}{2^{n+1}}$$

Comme  $(b_0 - a_0)/2^{n+1} \rightarrow 0$  quand  $n \rightarrow \infty$ , la suite  $c_n$  converge vers  $x^*$ . La convergence est **linéaire** (ou d'ordre 1), avec un taux de convergence  $\mu = 1/2$ . C'est-à-dire  $\lim_{n \rightarrow \infty} \frac{|x^* - c_{n+1}|}{|x^* - c_n|} = \frac{1}{2}$  si  $c_n$  est pris comme le milieu de l'intervalle qui devient le point final de l'intervalle suivant.

- **Nombre d'itérations** : Pour atteindre une précision  $\varepsilon$  (i.e.,  $b_N - a_N \leq \varepsilon$ ), le nombre d'itérations  $N$  nécessaire est  $N \approx \log_2 \left( \frac{b_0 - a_0}{\varepsilon} \right)$ .
- **Robustesse** : La méthode est très robuste car elle est garantie de converger.
- **Inconvénients** : La convergence peut être lente comparée à d'autres méthodes, surtout si la précision requise est élevée. Elle nécessite de nombreux calculs de  $f(x)$  (un par itération, plus les deux initiaux).

### 3.2 Méthode de la fausse position (Regula Falsi)

La méthode de la fausse position (ou *Regula Falsi*) est une autre méthode de bracketing pour trouver une racine. Elle est similaire à la dichotomie, mais au lieu de diviser l'intervalle en deux, elle utilise une interpolation linéaire.

**Principe** : Comme pour la dichotomie, on part d'un intervalle  $[a_n, b_n]$  tel que  $f(a_n)f(b_n) < 0$ . Au lieu de prendre le milieu  $c_n = (a_n + b_n)/2$ , on approxime  $f$  par la droite passant par les points  $(a_n, f(a_n))$  et  $(b_n, f(b_n))$ . Le point  $c_n$  est alors l'intersection de cette droite avec l'axe des abscisses. L'équation de la droite est  $y - f(a_n) = \frac{f(b_n) - f(a_n)}{b_n - a_n}(x - a_n)$ . En posant  $y = 0$  et  $x = c_n$ , on obtient :

$$c_n = a_n - f(a_n) \frac{b_n - a_n}{f(b_n) - f(a_n)} = \frac{a_n f(b_n) - b_n f(a_n)}{f(b_n) - f(a_n)}$$

Ensuite, comme pour la dichotomie, on compare les signes de  $f(a_n)f(c_n)$  et  $f(c_n)f(b_n)$  pour déterminer le nouvel intervalle  $[a_{n+1}, b_{n+1}]$  qui contient la racine.

**Algorithme de la fausse position** : L'algorithme est similaire à celui de la dichotomie, la seule différence étant le calcul de  $c_n$  :

1. **Initialisation** : Idem dichotomie.
2. **Itérations** : Tant que  $(b - a) > \varepsilon$  (ou une autre condition d'arrêt) et  $k < k_{\max}$  :

- (a) Calculer  $c = \frac{af(b)-bf(a)}{f(b)-f(a)}$ .
- (b) Si  $f(c) = 0$  (ou  $|f(c)| < \varepsilon_{\text{res}}$ ), alors  $x^* = c$ , arrêter.
- (c) Si  $f(a)f(c) < 0$ , alors  $b = c$ .
- (d) Sinon (si  $f(c)f(b) < 0$ ), alors  $a = c$ .
- (e)  $k = k + 1$ .

3. **Résultat :** La racine est approximée par  $c$ .

Le coût par itération est d'une évaluation de  $f$ .

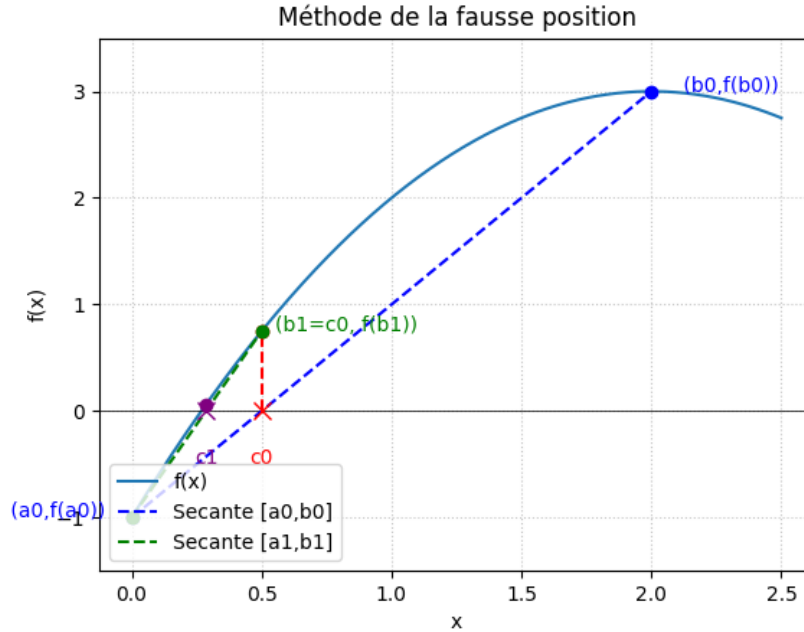


Figure 3: Illustration de la méthode de la fausse position. Le point  $c_n$  est l'intersection de la sécante avec l'axe des  $x$ . Un des points d'extrémité de l'intervalle ( $a_n$  ici) peut rester fixe pendant plusieurs itérations si la fonction a une convexité/concavité marquée.

#### Convergence de la fausse position :

- En général, la fausse position converge plus vite que la dichotomie. Cependant, sa convergence reste linéaire.
- Une particularité est que si la fonction  $f$  est convexe ou concave sur  $[a, b]$  (c'est-à-dire si  $f''$  ne change pas de signe sur  $[a, b]$ ), l'un des points d'extrémité de l'intervalle ( $a_n$  ou  $b_n$ ) restera fixe après quelques itérations (ou dès la première). L'autre extrémité convergera vers la racine  $x^*$ .

**Proposition 3.1.** Si  $f \in \mathcal{C}^2([a, b])$ ,  $f(a)f(b) < 0$ , et si  $f''$  n'a aucune racine dans  $]a, b[$  (i.e.,  $f$  est strictement convexe ou strictement concave sur  $[a, b]$ ), alors une des suites ( $a_n$ ) ou ( $b_n$ ) demeure constante à partir d'un certain rang. La suite ( $c_n$ ) générée par la méthode de la fausse position converge linéairement vers la racine  $x^*$  de  $f$ .

- Si, par exemple,  $b_n$  est constant (égal à  $b$ ) à partir d'un certain rang, alors la suite  $a_n$  converge vers  $x^*$ , et l'erreur  $e_n = x^* - a_n$  satisfait :

$$\lim_{n \rightarrow \infty} \frac{x^* - a_{n+1}}{x^* - a_n} = 1 - \frac{f'(x^*)(x^* - b)}{f(x^*) - f(b)} = 1 + \frac{f'(x^*)(b - x^*)}{f(b)}$$

(car  $f(x^*) = 0$ ). Un résultat similaire s'applique si  $a_n$  est constant et  $b_n \rightarrow x^*$ .

Bien que souvent plus rapide que la dichotomie en termes de nombre d'itérations, la convergence peut être lente si le point fixe est très éloigné de la racine par rapport à l'autre point. Des variantes existent pour améliorer ce comportement (ex: méthode de l'Illinois).

## 4 Algorithmes : Méthodes Itératives

Les méthodes de dichotomie et de fausse position sont des exemples de méthodes itératives. Une méthode itérative est un procédé mathématique qui génère une suite de solutions approchées  $(x_n)_{n \in \mathbb{N}}$  dans le but de converger vers la solution exacte  $x^*$ .

### 4.1 Convergence

Pour une méthode itérative générant une suite  $(x_n)$  qui converge vers une racine  $x^*$ , on s'intéresse à la vitesse de convergence.

**Definition 4.1** (Ordre de convergence). Soit  $(x_n)_{n \in \mathbb{N}}$  une suite qui converge vers  $x^*$ . On dit que la convergence est d'ordre  $p \geq 1$  s'il existe une constante  $\mu > 0$  (appelée taux de convergence asymptotique) telle que :

$$\lim_{n \rightarrow \infty} \frac{|x_{n+1} - x^*|}{|x_n - x^*|^p} = \mu$$

- Si  $p = 1$  et  $0 < \mu < 1$ , la convergence est dite **linéaire**.
- Si  $p = 1$  et  $\mu = 0$ , la convergence est dite **superlinéaire**.
- Si  $p = 2$ , la convergence est dite **quadratique**.
- Si  $p = 3$ , la convergence est dite **cubique**.

Plus l'ordre  $p$  est élevé, plus la convergence est rapide (en général). Si  $p = 1$ , un  $\mu$  petit indique une convergence plus rapide.

**Example 4.2** (Convergence de la dichotomie). Comme mentionné précédemment, la méthode de dichotomie a une convergence linéaire ( $p = 1$ ) avec un taux  $\mu = 1/2$ . En notant  $e_n = |x^* - c_n|$ , on a  $e_{n+1} \approx \frac{1}{2}e_n$ .

**Example 4.3** (Convergence de la fausse position). La méthode de la fausse position a également une convergence linéaire ( $p = 1$ ). Le taux de convergence  $\mu$  dépend de la fonction  $f$  et de l'intervalle initial (voir la formule donnée précédemment). Dans de nombreux cas,  $\mu$  est plus petit que  $1/2$ , ce qui rend la fausse position plus rapide que la dichotomie, mais elle peut être très proche de 1 dans des cas défavorables (si le point fixe est "mauvais").