Chapitre 2

Systèmes non linéaires

Dans le premier chapitre, on a étudié quelques méthodes de résolution de systèmes linéaires en dimension finie. L'objectif est maintenant de développer des méthodes de résolution de systèmes non linéaires, toujours en dimension finie. On se donne $g \in C(\mathbb{R}^n, \mathbb{R}^n)$ et on cherche x dans $\mathbb{R}^n$ solution de :

$$\begin{cases} x \in \mathbb{R}^n \\ g(x) = 0. \end{cases} \tag{2.1}$$

Au Chapitre I on a étudié des méthodes de résolution du système (2.1) dans le cas particulier $g(x) = Ax - b$, $A \in \mathcal{M}_n(\mathbb{R})$, $b \in \mathbb{R}^n$. On va maintenant étendre le champ d'étude au cas où g n'est pas forcément affine. On étudiera deux familles de méthodes pour la résolution approchée du système (2.1) :

— les méthodes de point fixe : point fixe de contraction et point fixe de monotonie

— les méthodes de type Newton [1].

2.1 Rappels et notations de calcul différentiel

Le premier chapitre faisait appel à vos connaissances en algèbre linéaire. Ce chapitre-ci, ainsi que le suivant (optimisation) s'appuieront sur vos connaissances en calcul différentiel, et nous allons donc réviser les quelques notions qui nous seront utiles.

Définition 2.1 (Application différentiable)**.** *Soient E et F des espaces vectoriels normés, f une application de E dans F et $x \in E$. On rappelle que f est différentiable en x s'il existe $T_x \in \mathcal{L}(E, F)$ (où $\mathcal{L}(E, F)$ est l'ensemble des applications linéaires continues de E dans F) telle que*

$$f(x + h) = f(x) + T_x(h) + \|h\|_E \varepsilon(h) \text{ avec } \varepsilon(h) \to 0 \text{ quand } h \to 0. \tag{2.2}$$

L'application T_x est alors unique [2] et on note $Df(x) = T_x \in \mathcal{L}(E, F)$ la différentielle de f au point x. Si f est différentiable en tout point de E, alors on appelle différentielle de f l'application $Df = E \to \mathcal{L}(E, F)$ qui à $x \in E$ associe l'application linéaire continue $Df(x)$ de E dans F.

Remarquons tout de suite que si f est une application linéaire continue de E dans F, alors f est différentiable, et $Df = f$. En effet, si f est linéaire, $f(x + h) - f(x) = f(h)$, et donc l'égalité (2.2) est vérifiée avec $T_x = f$ et $\varepsilon = 0$.
Voyons maintenant quelques cas particuliers d'espaces E et F :

1. Isaac Newton, 1643 - 1727, né d'une famille de fermiers, est un philosophe, mathématicien, physicien, alchimiste et astronome anglais. Figure emblématique des sciences, il est surtout reconnu pour sa théorie de la gravitation universelle et la création, en concurrence avec Leibniz, du calcul infinitésimal.

Cas où $E = \mathbb{R}$ et $F = \mathbb{R}$ Si f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$, dire que f est différentiable en x revient à dire que f est dérivable en x. En effet, dire que f est dérivable en x revient à dire que

$$\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \text{ existe, et } \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = f'(x),$$

ce qui s'écrit encore

$$\frac{f(x+h)-f(x)}{h} = f'(x) + \varepsilon(h), \text{ avec } \varepsilon(h) \to 0 \text{ lorsque } h \to 0,$$

c'est-à-dire

$$f(x+h) - f(x) = T_x(h) + h\varepsilon(h), \text{ avec } T_x(h) = f'(x)h,$$

ce qui revient à dire que f est différentiable en x, et que sa différentielle en x est l'application linéaire $T_x : \mathbb{R} \to \mathbb{R}$, qui à h associe $f'(x)h$. On a ainsi vérifié que pour une fonction de $\mathbb{R}$ dans $\mathbb{R}$, la notion de différentielle coïncide avec celle de dérivée.

Exemple 2.2. *Prenons $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = \sin x$. Alors f est dérivable en tout point et sa dérivée vaut $f'(x) = \cos x$. La fonction f est donc aussi différentiable en tout point. La différentielle de f au point x est l'application linéaire $Df(x)$ qui à $h \in \mathbb{R}$ associe $Df(x)(h) = \cos x\, h$. La différentielle de f est l'application de $\mathbb{R}$ dans $\mathcal{L}(\mathbb{R}, \mathbb{R})$, qui à x associe $Df(x)$ (qui est donc elle même une application linéaire).*

Cas où $E = \mathbb{R}^n$ et $F = \mathbb{R}^p$ Soit $f : \mathbb{R}^n \to \mathbb{R}^p$, $x \in \mathbb{R}^n$ et supposons que f est différentiable en x ; alors $Df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$; par caractérisation d'une application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}^n$, il existe une unique matrice $J_f(x) \in \mathcal{M}_{p,n}(\mathbb{R})$ telle que

$$\underbrace{Df(x)(y)}_{\in\mathbb{R}^p} = \underbrace{J_f(x)y}_{\in\mathbb{R}^p}, \ \forall y \in \mathbb{R}^n.$$

On confond alors souvent l'application linéaire $Df(x) \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)$ avec la matrice $J_f(x) \in \mathcal{M}_{p,n}(\mathbb{R})$ qui la représente, qu'on appelle **matrice jacobienne** de f au point x et qu'on note J_f. On écrit donc :

$$J_f(x) = Df(x) = (a_{i,j})_{1\le i\le p, 1\le j\le n} \text{ où } a_{i,j} = \partial_j f_i(x),$$

∂_j désignant la dérivée partielle par rapport à la j-ème variable.
Notons que si $n = p = 1$, la fonction f est de $\mathbb{R}$ dans $\mathbb{R}$ et la matrice jacobienne en x n'est autre que la dérivée en x : $J_f(x) = f'(x)$. On confond dans cette écriture la matrice $J_f(x)$ qui est de taille 1×1 avec le scalaire $f'(x)$.

Exemple 2.3. *Prenons $n = 3$ et $p = 2$; soit $f : \mathbb{R}^3 \to \mathbb{R}^2$ définie par :*

$$f(x) = \begin{pmatrix} x_1^2 + x_2^3 + x_3^4 \\ \\ 2x_1 - x_2 \end{pmatrix}, \ \forall x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Soit $h \in \mathbb{R}^3$ de composantes (h_1, h_2, h_3). Pour calculer la différentielle de f (en x appliquée à h), on peut calculer $f(x+h) - f(x)$:

$$\begin{aligned} f(x+h) - f(x) &= \begin{bmatrix} (x_1+h_1)^2 - x_1^2 + (x_2+h_2)^3 - x_2^3 + (x_3+h_3)^4 - x_3^4 \\ \\ 2(x_1+h_1) - 2x_1 - 2(x_2+h_2) + 2x_2 \end{bmatrix} \\ &= \begin{bmatrix} 2x_1h_1 + h_1^2 +' 3x_2^2h_2 + 3x_2h_2^2 + h_2^3 + -4x_3^3h_3 + -4x_3^2h_3^2 + h_3^4 \\ \\ 2h_1 - 2 + h_2 \end{bmatrix} \end{aligned}$$

et on peut ainsi vérifier l'égalité (2.2) avec :

$$Df(x)h = \begin{bmatrix} 2x_1h_1 + 3x_2^2h_2 + 4x_3^3h_3 \\ 2h_1 - h_2 \end{bmatrix}$$

et donc, avec les notations précédentes,

$$J_f(x) = \begin{bmatrix} 2x_1 & 3x_2^2 & 4x_3^3 \\ 2 & -1 & 0 \end{bmatrix}$$

Bien sûr, dans la pratique, on n'a pas besoin de calculer la différentielle en effectuant la différence $f(x+h)-f(x)$. On peut directement calculer les dériées partielles pour calculer la matrice jacobienne J_f.

Cas où $E = \mathbb{R}^n$**,** $F = \mathbb{R}$ C'est en fait un sous-cas du paragraphe précédent, puisqu'on est ici dans le cas $p = 1$. Soit $x \in \mathbb{R}^n$ et f une fonction de E dans F différentiable en x ; on a donc $J_f(x) \in \mathcal{M}_{1,n}(\mathbb{R})$: J_f est une matrice ligne. On définit le **gradient** de f en x comme le vecteur de $\mathbb{R}^n$ dont les composantes sont les coefficients de la matrice colonne $(J_f(x))^t$, ce qu'on écrit, avec un abus de notation, $\nabla f(x) = (J_f(x))^t \in \mathbb{R}^n$. (L'abus de notation est dû au fait qu'à gauche, il s'agit d'un vecteur de $\mathbb{R}^n$, et à droite, une matrice $n \times 1$, qui sont des objets mathématiques différents, mais qu'on identifie pour alléger les notations). Pour $(x, y) \in (\mathbb{R}^n)^2$, on a donc

$$Df(x)(y) = J_f(x)y = \sum_{j=i}^{n} \partial_j f(x) y_j = \nabla f(x) \cdot y \text{ où } \nabla f(x) = \begin{bmatrix} \partial_1 f(x) \\ \vdots \\ \partial_n f(x) \end{bmatrix} \in \mathbb{R}^n.$$

Attention, lorsque l'on écrit $J_f(x)y$ il s'agit d'un *produit matrice vecteur*, alors que lorsqu'on écrit $\nabla f(x) \cdot y$, il s'agit du *produit scalaire entre les vecteurs* $\nabla f(x)$ et y, qu'on peut aussi écrire $\nabla(f(x))^t y$.

Cas où E **est un espace de Hilbert et** $F = \mathbb{R}$**.** On généralise ici le cas présenté au paragraphe précédent. Soit $f : E \to \mathbb{R}$ différentiable en $x \in E$. Alors $Df(x) \in \mathcal{L}(E, \mathbb{R}) = E'$, où E' désigne le dual topologique de E, c.à.d. l'ensemble des formes linéaires continues sur E. Par le théorème de représentation de Riesz, il existe un unique $u \in E$ tel que $Df(x)(y) = (u|y)_E$ pour tout $y \in E$, où $(.|.)_E$ désigne le produit scalaire sur E. On appelle encore gradient de f en x ce vecteur u. On a donc $u = \nabla f(x) \in E$ et pour $y \in E$, $Df(x)(y) = (\nabla f(x)|y)_E$.

Différentielle d'ordre 2, matrice hessienne.

Revenons maintenant au cas général de deux espaces vectoriels normés E et F, et supposons maintenant que $f \in C^2(E, F)$. Le fait que $f \in C^2(E, F)$ signifie que $Df \in C^1(E, \mathcal{L}(E, F))$. Par définition, on a $D^2 f(x) \in \mathcal{L}(E, \mathcal{L}(E, F))$ et donc pour $y \in E$, $D^2 f(x)(y) \in \mathcal{L}(E, F)$, et pour $z \in E$, $D^2 f(x)(y)(z) \in F$.
Considérons maintenant le cas particulier $E = \mathbb{R}^n$ et $F = \mathbb{R}$. On a :

$$f \in C^2(\mathbb{R}^n, \mathbb{R}) \Leftrightarrow \left[f \in C^1(\mathbb{R}^n, \mathbb{R}) \text{ et } \nabla f \in C^1(\mathbb{R}^n, \mathbb{R}^n) \right].$$

et

$$D^2 f(x) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$$

Mais à toute application linéaire $\varphi \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$, on peut associer de manière unique une forme bilinéaire ϕ sur $\mathbb{R}^n$ de la manière suivante :

$$\phi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \tag{2.3}$$

$$(u, v) \mapsto \phi(u, v) = \underbrace{(\varphi(u))}_{\in \mathcal{L}(\mathbb{R}^n, \mathbb{R})} \underbrace{(v)}_{\in \mathbb{R}^n}. \tag{2.4}$$

On dit qu'il existe une isométrie canonique (un isomorphisme qui conserve la norme) entre l'espace vectoriel normé $\mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$ et l'espace des formes bilinéaires sur $\mathbb{R}^n$.
On appelle matrice hessienne de f et on note $H_f(x)$ la matrice de la forme bilinéaire ainsi associée à l'application linéaire $D^2 f(x) \in \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$.
On a donc $D^2 f(x)(y)(z) = y^t H_f(x) z$. La matrice hessienne $H_f(x)$ peut se calculer à l'aide des dérivées partielles : $H_f(x) = (b_{i,j})_{i,j=1\ldots N} \in \mathcal{M}_n(\mathbb{R})$ où $b_{i,j} = \partial^2_{i,j} f(x)$ et $\partial^2_{i,j}$ désigne la dérivée partielle par rapport

à la variable i de la dérivée partielle par rapport à la variable j. Notons que par définition (toujours avec l'abus de notation qui consiste à identifier les applications linéaires avec les matrices qui les représentent), $Dg(x)$ est la matrice jacobienne de $g = \nabla f$ en x.

Remarque 2.4 (Sur les différentielles, gradient et Hessienne)**.** *Pour définir la différentielle d'une fonction f d'un espace vectoriel de dimension finie E dans $\mathbb{R}$, on a besoin d'une norme sur E.*
Si f est différentiable en $x \in E$, pour définir le gradient de f en x, on a besoin d'un produit scalaire sur E pour pouvoir utiliser le théorème de representation de Riesz mentionné plus haut. Le gradient est défini de manière unique par le produit scalaire, mais ses composantes dépendent de la base choisie.
Enfin, si f est deux fois différentiable en $x \in E$, on a besoin d'une base de E pour définir la matrice hessienne en x, et cette matrice hessienne dépend de la base choisie.

2.2 Les méthodes de point fixe

2.2.1 Point fixe de contraction

Soit $g \in C(\mathbb{R}^n, \mathbb{R}^n)$, on définit la fonction $f \in C(\mathbb{R}^n, \mathbb{R}^n)$ par $f(x) = x - g(x)$. On peut alors remarquer que $g(x) = 0$ si et seulement si $f(x) = x$. Résoudre le système non linéaire (2.1) revient donc à trouver un point fixe de f. Encore faut-il qu'un tel point fixe existe... On rappelle le théorème de point fixe bien connu :

Théorème 2.5 (Point fixe)**.** *Soit E un espace métrique complet, d la distance sur E, et $f : E \to E$ une fonction strictement contractante, c'est–à–dire telle qu'il existe $\kappa \in]0,1[$ tel que $d(f(x), f(y)) \leq \kappa d(x,y)$ pour tout $x, y \in E$. Alors il existe un unique point fixe $\bar{x} \in E$ qui vérifie $f(\bar{x}) = \bar{x}$. De plus si $x^{(0)} \in E$, et $x^{(k+1)} = f(x^{(k)})$, $\forall k \geq 0$, alors $x^{(k)} \to \bar{x}$ quand $n + \infty$.*

DÉMONSTRATION – *Etape 1 : Existence de $\bar{x}$ et convergence de la suite*
Soit $x^{(0)} \in E$ et $(x^{(k)})_{k\in\mathbb{N}}$ la suite définie par $x^{(k+1)} = f(x^{(k)})$ pour $k \geq 0$. On va montrer que :

1. la suite $(x^{(k)})_{k\in\mathbb{N}}$ est de Cauchy (donc convergente car E est complet),
2. $\lim\limits_{n\to+\infty} x^{(k)} = \bar{x}$ est point fixe de f.

Par hypothèse, on sait que pour tout $k \geq 1$,
$$d(x^{(k+1)}, x^{(k)}) = d(f(x^{(k)}), f(x^{(k-1)})) \leq \kappa d(x^{(k)}, x^{(k-1)}).$$
Par récurrence sur k, on obtient que
$$d(x^{(k+1)}, x^{(k)}) \leq \kappa^k d(x^{(1)}, x^{(0)}),\ \forall k \geq 0.$$
Soit $k \geq 0$ et $p \geq 1$, on a donc :
$$\begin{aligned}
d(x^{(k+p)}, x^{(k)}) &\leq d(x^{(k+p)}, x^{(k+p-1)}) + \cdots + d(x^{(k+1)}, x^{(k)}) \\
&\leq \sum_{q=1}^{p} d(x^{(k+q)}, x^{(k+q-1)}) \\
&\leq \sum_{q=1}^{p} \kappa^{k+q-1} d(x^{(1)}, x^{(0)}) \\
&\leq d(x^{(1)}, x^{(0)}) \kappa^k (1 + \kappa + \ldots + \kappa^{p-1}) \\
&\leq d(x^{(1)}, x^{(0)}) \frac{\kappa^k}{1-\kappa} \longrightarrow 0 \text{ quand } k \to +\infty \text{ car } \kappa < 1.
\end{aligned}$$
La suite $(x^{(k)})_{k\in\mathbb{N}}$ est donc de Cauchy, *i.e.* :
$$\forall \varepsilon > 0,\quad \exists k_\varepsilon \in \mathbb{N}\,;\quad \forall k \geq k_\varepsilon,\quad \forall\ \ p \geq 1\quad d(x^{(k+p)}, x^{(k)}) \leq \varepsilon.$$

Comme E est complet, on a donc $x^{(k)} \longrightarrow \bar{x}$ dans E quand $k \to +\infty$. Comme la fonction f est strictement contractante, elle est continue, donc on a aussi $f(x^{(k)}) \longrightarrow f(\bar{x})$ dans E quand $k \to +\infty$. En passant à la limite dans l'égalité $x^{(k+1)} = f(x^{(k)})$, on en déduit que $\bar{x} = f(\bar{x})$.

Etape 2 : Unicité

Soit $\bar{x}$ et $\bar{y}$ des points fixes de f, qui satisfont donc $\bar{x} = f(\bar{x})$ et $\bar{y} = f(\bar{y})$. Alors $d(f(\bar{x}), f(\bar{y})) = d(\bar{x}, \bar{y}) \leq \kappa d(\bar{x}, \bar{y})$; comme $\kappa < 1$, ceci est impossible sauf si $\bar{x} = \bar{y}$.

■

La méthode du point fixe s'appelle aussi méthode des itérations successives. Dans le cadre de ce cours, nous prendrons $E = \mathbb{R}^n$, et la distance associée à la norme euclidienne, que nous noterons $|\cdot|$.

$$\forall(\boldsymbol{x}, \boldsymbol{y}) \in \mathbb{R}^n \times \mathbb{R}^n \text{ avec } \boldsymbol{x} = (x_1, \ldots, x_n), \boldsymbol{y} = (y_1, \ldots, y_n), d(\boldsymbol{x}, \boldsymbol{y}) = |\boldsymbol{x} - \boldsymbol{y}| = \left(\sum_{i=1}^n (x_i - y_i)^2\right)^{\frac{1}{2}}.$$

A titre d'illustration, essayons de la mettre en oeuvre pour trouver les points fixes de la fonction $x \mapsto x^2$.

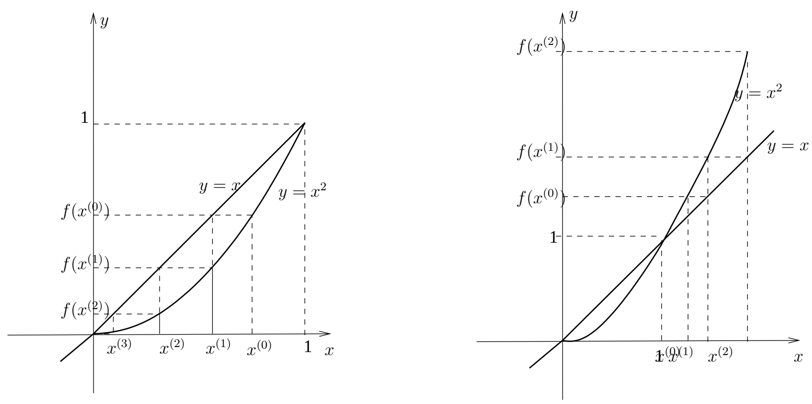

FIGURE 2.1: Comportement des itérés successifs du point fixe pour $x \mapsto x^2$– A gauche : $x^{(0)} < 1$, à droite : $x^{(0)} > 1$.

Pour la fonction $x \mapsto x^2$, on voit sur la figure 2.1, côté gauche, que si l'on part de $x = x^{(0)} < 1$, la méthode converge rapidement vers 0 ; or la fonction $x \mapsto x^2$ n'est strictement contractante que sur l'intervalle $]-\frac{1}{2}, \frac{1}{2}[$. Donc si $x = x^{(0)} \in]-\frac{1}{2}, \frac{1}{2}[$, on est dans les conditions d'application du théorème du point fixe. Mais en fait, la suite $(x^{(k)})_{k\in\mathbb{N}}$ définie par le point fixe converge pour tout $x^{(0)} \in]-1, 1[$; ceci est très facile à voir car $x^{(k)} = (x^{(k)})^2$ et on a donc convergence vers 0 si $|x| < 1$.
Par contre si l'on part de $x^{(0)} > 1$ (à droite sur la figure 2.1), on diverge rapidement : mais rien de surprenant à cela, puisque la fonction $x \mapsto x^2$ n'est pas contractante sur $[1, +\infty[$
Dans le cas de la fonction $x \mapsto \sqrt{x}$, on voit sur la figure 2.2 que les itérés convergent vers 1 que l'on parte à droite ou à gauche de $x = 1$; on peut même démontrer (exercice) que si $x^{(0)} > 0$, la suite $(x)_{k\in\mathbb{N}}$ converge vers 1 lorsque $k \to +\infty$. Pourtant la fonction $x \mapsto \sqrt{x}$ n'est contractante que pour $x > \frac{1}{4}$; mais on n'atteint jamais le point fixe 0, ce qui est moral, puisque la fonction n'est pas contractante en 0. On se rend compte encore sur cet exemple que le théorème du point fixe donne une condition suffisante de convergence, mais que cette condition n'est pas nécessaire.
Remarquons que l'hypothèse que f envoie E dans E est cruciale. Par exemple la fonction $f : x \mapsto \frac{1}{x}$ est lipschitzienne de rapport $k < 1$ sur $[1 + \varepsilon, +\infty[$ pour tout $\varepsilon > 0$ mais elle n'envoie pas $[1 + \varepsilon, +\infty[$ dans $[1 + \varepsilon, +\infty[$. La méthode du point fixe à partir du choix initial $x \neq 1$ donne la suite $x, \frac{1}{x}, x, \frac{1}{x}, \ldots, x, \frac{1}{x}$ qui ne converge pas.

Remarque 2.6 (Vitesse de convergence)**.** *Sous les hypothèses du théorème 2.5, $d(x^{(k+1)}, \bar{x}) = d(f(x^{(k)}), f(\bar{x})) \leq kd(x^{(k)}, \bar{x})$; donc si $x^{(k)} \neq \bar{x}$ alors $\frac{d(x^{(k+1)}, \bar{x})}{d(x^{(k)}, \bar{x})} \leq k \ (< 1)$, voir à ce sujet la définition 2.14. La convergence est donc au moins linéaire (même si de fait, cette méthode converge en général assez lentement).*

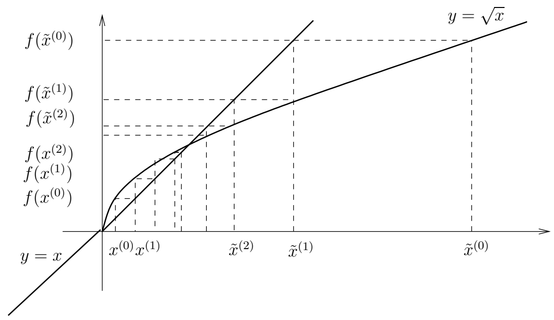

FIGURE 2.2: Comportement des itérés successifs du point fixe pour $x \mapsto \sqrt{x}$

Remarque 2.7 (Généralisation)**.** *Le théorème 2.5 se généralise en remplaçant l'hypothèse "f strictement contractante" par " il existe $k > 0$ tel que $f^{(k)} = \underbrace{f \circ f \circ \ldots \circ f}_{k \text{ fois}}$ est strictement contractante " (reprendre la démonstration du théorème pour le vérifier).*

La question qui vient alors naturellement est : que faire pour résoudre $g(x) = 0$ si la méthode du point fixe appliquée à la fonction $x \mapsto x - g(x)$ ne converge pas ? Dans ce cas, f n'est pas strictement contractante ; une idée possible est de pondérer la fonction g par un paramètre $\omega \neq 0$ et d'appliquer les itérations de point fixe à la fonction $f_\omega(x) = x - \omega g(x)$; on remarque là encore que x est encore solution du système (2.1) si et seulement si x est point fixe de $f_\omega(x)$. On aimerait dans ce cas trouver ω pour que f_ω soit strictement contractante, c.à.d. pour que

$$|f_\omega(x) - f_\omega(y)| = |x - y - \omega(g(x) - g(y))| \leq \kappa |x - y| \text{ pour } (x, y) \in \mathbb{R}^n \times \mathbb{R}^n, \text{ avec } \kappa < 1.$$

Or

$$\begin{aligned} |x - y - \omega(g(x) - g(y))|^2 &= \big(x - y - \omega(g(x) - g(y))\big) \cdot \big(x - y - \omega(g(x) - g(y))\big) \\ &= |x - y|^2 - 2(x - y) \cdot (\omega(g(x) - g(y))) + \omega^2 |g(x) - g(y)|^2. \end{aligned}$$

Supposons que g soit lipschitzienne, et soit $M > 0$ sa constante de Lipschitz :

$$|g(x) - g(y)| \leq M|x - y|, \ \forall x, y \in \mathbb{R}^n. \tag{2.5}$$

On a donc

$$|x - y - \omega(g(x) - g(y))|^2 \leq (1 + \omega^2 M^2)|x - y|^2 - 2(x - y) \cdot (\omega(g(x) - g(y)))$$

Or on veut $|x - y - \omega(g(x) - g(y))|^2 \leq \kappa |x - y|^2$, avec $\kappa < 1$. On a donc intérêt à ce que le terme $-2(x - y) \cdot (\omega(g(x) - g(y)))$ soit de la forme $-a|x - y|^2$ avec a strictement positif. Pour obtenir ceci, on va supposer de plus que :

$$\exists \alpha > 0 \text{ tel que } (g(x) - g(y)) \cdot (x - y) \geq \alpha |x - y|^2, \forall x, y \in \mathbb{R}^n, \tag{2.6}$$

On obtient alors :

$$|x - y - \omega(g(x) - g(y))|^2 \leq (1 + \omega^2 M^2 - 2\omega\alpha)|x - y|^2.$$

Et donc si $\omega \in]0, \frac{2\alpha}{M^2}[$, le polynôme $\omega^2 M^2 - 2\omega\alpha$ est strictement négatif : soit $-\mu$ (noter que $\mu \in]0, 1[$) et on obtient que

$$|x - y - \omega(g(x) - g(y))|^2 \leq (1 - \mu)|x - y|^2.$$

On peut donc énoncer le théorème suivant :

Théorème 2.8 (Point fixe de contraction avec relaxation)**.** *On désigne par $|\cdot|$ la norme euclidienne sur $\mathbb{R}^n$. Soit $g \in C(\mathbb{R}^n, \mathbb{R}^n)$ lipschitzienne de constante de Lipschitz $M > 0$, et telle que (2.6) est vérifiée : alors la fonction $f_\omega : x \mapsto x - \omega g(x)$ est strictement contractante si $0 < \omega < \dfrac{2\alpha}{M^2}$. Il existe donc un et un seul $\bar{x} \in \mathbb{R}^n$ tel que $g(\bar{x}) = 0$ et $x^{(k)} \to \bar{x}$ quand $n \to +\infty$ avec $x^{(k+1)} = f_\omega(x^{(k)}) = x^{(k)} - \omega g(x^{(k)})$.*

Remarque 2.9. *Le théorème 2.8 permet de montrer que sous les hypothèses* (2.6) *et* (2.5)*, et pour $\omega \in]0, \frac{2\alpha}{M^2}[$, on peut obtenir la solution de* (2.1) *en construisant la suite :*

$$\left\{ \begin{array}{l} x^{(k+1)} = x^{(k)} - \omega g(x^{(k)}) \;\; n \geq 0, \\ x^{(0)} \in \mathbb{R}^n. \end{array} \right. \tag{2.7}$$

Or on peut aussi écrire cette suite de la manière suivante (avec $f(x) = x - g(x)$) :

$$\left\{ \begin{array}{ll} \tilde{x}^{(k+1)} & = f(x^{(k)}), \;\; \forall n \geq 0 \\ x^{(k+1)} & = \omega \tilde{x}^{(k+1)} + (1-\omega) x^{(k)}, \quad x^{(0)} \in \mathbb{R}^n. \end{array} \right. \tag{2.8}$$

En effet si $x^{(k+1)}$ est donné par la suite (2.8)*, alors*

$$x^{(k+1)} = \omega \tilde{x}^{(k+1)} + (1-\omega) x^{(k)} = \omega f(x^{(k)}) + (1-\omega) x^{(k)} = -\omega g(x^{(k)}) + x^{(k)}.$$

Le procédé de construction de la suite (2.8) *est l'algorithme de relaxation sur f.*

La proposition suivante donne une condition suffisante pour qu'une fonction vérifie les hypothèses (2.6) et (2.5).

Proposition 2.10. *Soit $h \in C^2(\mathbb{R}^n, \mathbb{R})$, et $(\lambda_i)_{i=1,n}$ les valeurs propres de la matrice hessienne de h. On suppose qu'il existe des réels strictement positifs α et M tels que*

$$\alpha \leq \lambda_i(x) \leq M, \; \forall i \in \{1 \ldots n\}, \; \forall x \in \mathbb{R}^n.$$

(Notons que cette hypothèse est plausible puisque les valeurs propres de la matrice hessienne sont réelles). Alors la fonction $g = \nabla h$ (gradient de h) vérifie les hypothèses (2.6) *et* (2.5) *du théorème 2.8.*

DÉMONSTRATION – Montrons d'abord que l'hypothèse (2.6) est vérifiée. Soit $(x, y) \in (\mathbb{R}^n)^2$, on veut montrer que $(g(x) - g(y)) \cdot (x - y) \geq \alpha |x - y|^2$. On introduit pour cela la fonction $\varphi \in C^1(\mathbb{R}, \mathbb{R}^n)$ définie par :

$$\varphi(t) = g(x + t(y - x)).$$

On a donc

$$\varphi(1) - \varphi(0) = g(y) - g(x) = \int_0^1 \varphi'(t) dt.$$

Or $\varphi'(t) = Dg(x + t(y - x))(y - x)$. Donc

$$g(y) - g(x) = \int_0^1 Dg(x + t(y - x))(y - x) dt.$$

On en déduit que :

$$(g(y) - g(x)) \cdot (y - x) = \int_0^1 \left(Dg(x + t(y - x))(y - x) \cdot (y - x) \right) dt.$$

Comme $\lambda_i(x) \in [\alpha, M]$ $\forall i \in \{1, \ldots, n\}$, on a

$$\alpha |w|^2 \leq Dg(z) w \cdot w \leq M |w|^2 \text{ pour tout } w, z \in \mathbb{R}^n$$

On a donc :

$$(g(y) - g(x)) \cdot (y - x) \geq \int_0^1 \alpha |y - x|^2 dt = \alpha |y - x|^2$$

ce qui montre que l'hypothèse (2.6) est bien vérifiée.

Montrons maintenant que l'hypothèse (2.5) est vérifiée. On veut montrer que $|g(y) - g(x)| \leq M|y - x|$. Comme

$$g(y) - g(x) = \int_0^1 Dg(x + t(y - x))(y - x)dt,$$

on a

$$\begin{aligned}|g(y) - g(x)| &\leq \int_0^1 |Dg(x + t(y - x))(y - x)|dt \\ &\leq \int_0^1 |Dg(x + t(y - x))||y - x|dt,\end{aligned}$$

où $|.|$ est la norme sur $\mathcal{M}_n(\mathbb{R})$ induite par la norme euclidienne sur $\mathbb{R}^n$.
Or, comme $\lambda_i(x) \in [\alpha, M]$ pour tout $i = 1, \ldots, n$, la matrice $Dg(x + t(y - x))$ est symétrique définie positive et donc, d'après la proposition 1.30 page 65, son rayon spectral est égal à sa norme, pour la norme induite par la norme euclidienne. On a donc :

$$|Dg(x + t(y - x)| = \rho(Dg(x + t(y - x)) \leq M.$$

On a donc ainsi montré que : $|g(y) - g(x)| \leq M|y - x|$, ce qui termine la démonstration. ■

2.2.2 Point fixe de monotonie

Dans de nombreux cas issus de la discrétisation d'équations aux dérivées partielles, le problème de résolution d'un problème non linéaire apparaît sous la forme $Ax = R(x)$ où A est une matrice carrée d'ordre n inversible, et $R \in C(\mathbb{R}^n, \mathbb{R}^n)$. On peut le réécrire sous la forme $x = A^{-1}R(x)$ et appliquer l'algorithme de point fixe sur la fonction $f : x \mapsto A^{-1}Rx$, ce qui donne comme itération : $x^{(k+1)} = A^{-1}R(x^{(k)})$. Si on pratique un point fixe avec relaxation, dont le paramètre de relaxation $\omega > 0$, alors l'itération s'écrit :

$$\tilde{x}^{(k+1)} = A^{-1}R(x^{(k)}), \qquad x^{(k+1)} = \omega\tilde{x}^{(k+1)} + (1 - \omega)x^{(k)}.$$

Si la matrice A possède une propriété dite "de monotonie", on peut montrer la convergence de l'algorithme du point fixe ; c'est l'objet du théorème suivant.

Théorème 2.11 (Point fixe de monotonie)**.**
Soient $A \in \mathcal{M}_n(\mathbb{R})$ et $R \in C(\mathbb{R}^n, \mathbb{R}^n)$. On suppose que :

1. *La matrice A est une matrice d'inverse positive, ou IP-matrice (voir exercice 10), c'est-à-dire que A est inversible et tous les coefficients de A^{-1} sont positifs ou nuls, ce qui est équivalent à dire que :*

$$Ax \geq 0 \Rightarrow x \geq 0,$$

au sens composante par composante, c'est-à-dire

$$\big((Ax)_i \geq 0,\ \forall i = 1, \ldots, n\big) \Rightarrow \big(x_i \geq 0,\ \forall i = 1, \ldots, n\big).$$

2. *R est monotone, c'est-à-dire que si $x \geq y$ (composante par composante) alors $R(x) \geq R(y)$ (composante par composante).*
3. *0 est une sous-solution du problème, c'est-à-dire que $0 \leq R(0)$ et il existe $\tilde{x} \in \mathbb{R}^n$; $\tilde{x} \geq 0$ tel que $\tilde{x}$ est une sur-solution du problème, c'est-à-dire que $A\tilde{x} \geq R(\tilde{x})$.*

On pose $x^{(0)} = 0$ et $Ax^{(k+1)} = R(x^{(k)})$. On a alors :

1. $0 \leq x^{(k)} \leq \tilde{x},\ \ \forall k \in \mathbb{N}$,
2. $x^{(k+1)} \geq x^{(k)},\ \ \forall k \in \mathbb{N}$,
3. $x^{(k)} \to \bar{x}$ *quand* $k \to +\infty$ *et* $A\bar{x} = R(\bar{x})$.

DÉMONSTRATION – Comme A est inversible la suite $(x^{(k)})_{n\in\mathbb{N}}$ vérifiant

$$\begin{cases} x^{(0)} = 0, \\ Ax^{(k+1)} = R(x^{(k)}), \;\; k \geq 0 \end{cases}$$

est bien définie. On va montrer par récurrence sur k que $0 \leq x^{(k)} \leq \tilde{x}$ pour tout $k \geq 0$ et que $x^{(k)} \leq x^{(k+1)}$ pour tout $k \geq 0$.

1. Pour $k = 0$, on a $x^{(0)} = 0$ et donc $0 \leq x^{(0)} \leq \tilde{x}$ et $Ax^{(1)} = R(0) \geq 0$. On en déduit que $x^{(1)} \geq 0$ grâce aux hypothèses 1 et 3 et donc $x^{(1)} \geq x^{(0)} = 0$.
2. On suppose maintenant (hypothèse de récurrence) que $0 \leq x^{(p)} \leq \tilde{x}$ et $x^{(p)} \leq x^{(p+1)}$ pour tout $p \in \{0, \ldots, n-1\}$.
   On veut montrer que $0 \leq x^{(k)} \leq \tilde{x}$ et que $x^{(k)} \leq x^{(k+1)}$. Par hypothèse de récurrence pour $p = k-1$, on sait que $x^{(k)} \geq x^{(k-1)}$ et que $x^{(k-1)} \geq 0$. On a donc $x^{(k)} \geq 0$. Par hypothèse de récurrence, on a également que $x^{(k-1)} \leq \tilde{x}$ et grâce à l'hypothèse 2, on a donc $R(x^{(k-1)}) \leq R(\tilde{x})$. Par définition de la suite $(x^{(k)})_{k\in\mathbb{N}}$, on a $Ax^{(k)} = R(x^{(k-1)})$ et grâce à l'hypothèse 3, on sait que $A\tilde{x} \geq R(\tilde{x})$. On a donc : $A(\tilde{x} - x^{(k)}) \geq R(\tilde{x}) - R(x^{(k-1)}) \geq 0$. On en déduit alors (grâce à l'hypothèse 1) que $x^{(k)} \leq \tilde{x}$.
   De plus, comme $Ax^{(k)} = R(x^{(k-1)})$ et $Ax^{(k+1)} = R(x^{(k)})$, on a $A(x^{(k+1)} - x^{(k)}) = R(x^{(k)}) - R(x^{(k-1)}) \geq 0$ par l'hypothèse 2, et donc grâce à l'hypothèse 1, $x^{(k+1)} \geq x^{(k)}$.

On a donc ainsi montré (par récurrence) que

$$\begin{aligned} 0 \leq x^{(k)} \leq \tilde{x}, \quad \forall k \geq 0 \\ x^{(k)} \leq x^{(k+1)}, \quad \forall k \geq 0. \end{aligned}$$

Ces inégalités s'entendent composante par composante, c.à.d. que si $x^{(k)} = (x_1^{(k)} \ldots x_n^{(k)})^t \in \mathbb{R}^n$ et $\tilde{x} = (\tilde{x}_1 \ldots \tilde{x}_n)^t \in \mathbb{R}^n$, alors $0 \leq x_i^{(k)} \leq \tilde{x}_i$ et $x_i^{(k)} \leq x_i^{(k+1)}$, $\forall i \in \{1, \ldots, n\}$, et $\forall k \geq 0$.

Soit $i \in \{1, \ldots, n\}$; la suite $(x_i^{(k)})_{n\in\mathbb{N}} \subset \mathbb{R}$ est croissante et majorée par $\tilde{x}_i$ donc il existe $\bar{x}_i \in \mathbb{R}$ tel que $\bar{x}_i = \lim\limits_{k\to+\infty} x_i^{(k)}$. Si on pose $\bar{x} = (\bar{x}_1 \ldots \bar{x}_n)^t \in \mathbb{R}^n$, on a donc $x^{(k)} \longrightarrow \bar{x}$ quand $k \to +\infty$.

Enfin, comme $Ax^{(k+1)} = R(x^{(k)})$ et comme R est continue, on obtient par passage à la limite lorsque $k \to +\infty$ que $A\bar{x} = R(\bar{x})$ et que $0 \leq \bar{x} \leq \tilde{x}$. ■

L'hypothèse 1 du théorème 2.11 est vérifiée par exemple par les matrices A qu'on a obtenues par discrétisation par différences finies des opérateurs $-u''$ sur l'intervalle $]0,1[$ (voir page 11 et l'exercice 51) et Δu sur $]0,1[\times]0,1[$ (voir page 14).

Théorème 2.12 (Généralisation du précédent)**.**
Soit $A \in \mathcal{M}_n(\mathbb{R})$, $R \in C^1(\mathbb{R}^n, \mathbb{R}^n)$, $R = (R_1, \ldots, R_n)^t$ tels que

1. *Pour tout $\beta \geq 0$ et pour tout $x \in \mathbb{R}^n$, $Ax + \beta x \geq 0 \Rightarrow x \geq 0$*
2. *$\dfrac{\partial R_i}{\partial x_j} \geq 0$, $\forall i, j$ t.q. $i \neq j$ (R_i est monotone croissante par rapport à la variable x_j si $j \neq i$) et $\exists \gamma > 0$, $-\gamma \leq \dfrac{\partial R_i}{\partial x_i} \leq 0$, $\forall x \in \mathbb{R}^n$, $\forall i \in \{1, \ldots, n\}$ (R_i est monotone décroissante par rapport à la variable x_i).*
3. *$0 \leq R(0)$ (0 est sous-solution) et il existe $\tilde{x} \geq 0$ tel que $A(\tilde{x}) \geq R(\tilde{x})$ ($\tilde{x}$ est sur-solution).*

Soient $x^{(0)} = 0$, $\beta \geq \gamma$, et $(x^{(k)})_{n\in\mathbb{N}}$ la suite définie par $Ax^{(k+1)} + \beta x^{(k+1)} = R(x^{(k)}) + \beta x^{(k)}$. Cette suite converge vers $\bar{x} \in \mathbb{R}^n$ et $A\bar{x} = R(\bar{x})$. De plus, $0 \leq x^{(k)} \leq \tilde{x}$ $\forall n \in \mathbb{N}$ et $x^{(k)} \leq x^{(k+1)}$, $\forall n \in \mathbb{N}$.

DÉMONSTRATION – On se ramène au théorème précédent avec $A + \beta Id$ au lieu de A et $R + \beta$ au lieu de R. ■

Remarque 2.13 (Point fixe de Brouwer)**.** *On s'est intéressé ici uniquement à des théorèmes de point fixe "constructifs", i.e. qui donnent un algorithme pour le déterminer. Il existe aussi un théorème de point fixe dans $\mathbb{R}^n$ avec des* *hypothèses beaucoup plus générales (mais le théorème est non constructif), c'est le théorème de Brouwer*[3] *: si f est une fonction continue de la boule unité de $\mathbb{R}^n$ dans la boule unité, alors elle admet un point fixe dans la boule unité.*

2.2.3 Vitesse de convergence

Définition 2.14 (Vitesse de convergence)**.** *Soit $(x^{(k)})_{n\in\mathbb{N}} \in \mathbb{R}^n$ et $\bar{x} \in \mathbb{R}^n$. On suppose que $x^{(k)} \to \bar{x}$ lorsque $k \to +\infty$, que la suite est non stationnaire, c.à.d. que $x^{(k)} \neq \bar{x}$ pour tout $k \in \mathbb{N}$, et que*

$$\lim_{k\to+\infty} \frac{\|x^{(k+1)} - \bar{x}\|}{\|x^{(k)} - \bar{x}\|} = \beta \in [0, 1]. \tag{2.9}$$

On s'intéresse à la "vitesse de convergence" de la suite $(x^{(k)})_{k\in\mathbb{N}}$. On dit que :

1. *La convergence est* **sous–linéaire** *si $\beta = 1$.*
2. *La convergence est* **au moins linéaire** *si $\beta \in [0, 1[$.*
3. *La convergence* **est linéaire** *si $\beta \in]0, 1[$.*
4. *La convergence est* **super linéaire** *si $\beta = 0$. Dans ce cas, on dit également que :*
   (a) *La convergence est* **au moins quadratique** *s'il existe $\gamma \in \mathbb{R}_+$ et il existe $n_0 \in \mathbb{N}$ tels que si $k \geq n_0$ alors $\|x^{(k+1)} - \bar{x}\| \leq \gamma \|x^{(k)} - \bar{x}\|^2$.*
   (b) *La convergence est* **quadratique** *si*

$$\lim_{k\to+\infty} \frac{\|x^{(k+1)} - \bar{x}\|}{\|x^{(k)} - \bar{x}\|^2} = \gamma > 0.$$

Plus généralement, on dit que :

(a) *La convergence est* **au moins d'ordre** *p s'il existe $\gamma \in \mathbb{R}_+$ et il existe $k_0 \in \mathbb{N}$ tels que si $k \geq k_0$ alors $\|x^{(k+1)} - \bar{x}\| \leq \gamma \|x^{(k)} - \bar{x}\|^p$.*

(b) *La convergence est* **d'ordre** *p si*

$$\lim_{k\to+\infty} \frac{\|x^{(k+1)} - \bar{x}\|}{\|x^{(k)} - \bar{x}\|^p} = \gamma > 0.$$

Remarque 2.15 (Sur la vitesse de convergence des suites)**.**

— Remarquons d'abord que si une suite $(x^{(k)})_{k\in\mathbb{N}}$ de $\mathbb{R}^n$ converge vers $\bar{x}$ lorsque k tend vers l'infini, et qu'il existe β vérifiant (2.9), alors on a forcément $\beta \leq 1$. En effet, si la suite vérifie (2.9) avec $\beta > 1$, alors il existe $k_0 \in \mathbb{N}$ tel que si $k \geq k_0$, $|x_n - \bar{x}| \geq |x_{k_0} - \bar{x}|$ pour tout $k \geq k_0$, ce qui contredit la convergence.

— *Quelques exemples de suites qui convergent sous-linéairement* : $x_k = \frac{1}{\sqrt{k}}$, $x_k = \frac{1}{k}$, mais aussi, de manière moins intuitive : $x_k = \frac{1}{k^2}$. Toutes ces suites vérifient l'égalité (2.9) avec $\beta = 1$.

— Attention donc, contrairement à ce que pourrait suggérer son nom, la convergence linéaire (au sens donné ci-dessus), est déjà une convergence très rapide. Les suites géométriques définies par $x_k = \beta^k$ avec $\beta \in]0, 1[$ sont des suites qui convergent linéairement (vers 0), car elles verifient évidemment bien (2.9) avec $\beta \in]0, 1[$.

— la convergence quadratique est encore plus rapide ! Par exemple la suite définie par $x_{k+1} = x_k^2$ converge de manière quadratique pour un choix initial $x_0 \in] - 1, 1[$. Mais si par malheur le choix initial est en dehors

3. Luitzen Egbertus Jan Brouwer (1881-1966), mathématicien néerlandais.

de cet intervalle, la suite diverge alors très vite... de manière exponentielle, en fait (pour $x_0 > 1$, on a $x_k = e^{2k \ln x_0}$).
C'est le cas de la méthode de Newton, que nous allons introduire maintenant. Lorsqu'elle converge, elle converge très vite (nous démontrerons que la vitesse de convergence est quadratique). Mais lorsqu'elle diverge, elle diverge aussi très vite...

Pour construire des méthodes itératives qui convergent "super vite", nous allons donc essayer d'obtenir des vitesses de convergence super linéaires. C'est dans cet esprit que nous étudions dans la proposition suivante des conditions suffisantes de convergence de vitesse quadratique pour une méthode de type point fixe, dans le cas d'une fonction f de $\mathbb{R}$ dans $\mathbb{R}$.

Proposition 2.16 (Vitesse de convergence d'une méthode de point fixe)**.** *Soit $f \in C^1(\mathbb{R},\mathbb{R})$; on suppose qu'il existe $\bar{x} \in \mathbb{R}$ tel que $f(\bar{x}) = \bar{x}$. On construit la suite*

$$x^{(0)} \in \mathbb{R}$$
$$x^{(k+1)} = f(x^{(k)}).$$

1. *Si on suppose que $f'(\bar{x}) \neq 0$ et $|f'(\bar{x})| < 1$, alors il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha = [\bar{x}-\alpha, \bar{x}+\alpha]$ on a $x^{(k)} \to \bar{x}$ lorsque $k \to +\infty$. De plus si $x^{(k)} \neq \bar{x}$ pour tout $k \in \mathbb{N}$, alors*

$$\frac{|x^{(k+1)} - \bar{x}|}{|x^{(k)} - \bar{x}|} \to |f'(\bar{x})| = \beta \text{ avec } \beta \in]0,1[.$$

La convergence est donc linéaire.

2. *Si on suppose maintenant que $f'(\bar{x}) = 0$ et $f \in C^2(\mathbb{R},\mathbb{R})$, alors il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha = [\bar{x}-\alpha, \bar{x}+\alpha]$, alors $x^{(k)} \to \bar{x}$ quand $k \to +\infty$, et si $x^{(k)} \neq \bar{x}$, $\forall n \in \mathbb{N}$ alors*

$$\frac{|x^{(k+1)} - \bar{x}|}{|x^{(k)} - \bar{x}|^2} \to \beta = \frac{1}{2}|f''(\bar{x})|.$$

Dans ce cas, la convergence est donc au moins quadratique.

DÉMONSTRATION –

1. Supposons que $|f'(\bar{x})| < 1$, et montrons qu'il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha$ alors $x^{(k)} \to \bar{x}$. Comme $f \in C^1(\mathbb{R},\mathbb{R})$ il existe $\alpha > 0$ tel que $\gamma = \max_{x \in I_\alpha} |f'(x)| < 1$ (par continuité de f').
On va maintenant montrer que $f : I_\alpha \to I_\alpha$ est strictement contractante, on pourra alors appliquer le théorème du point fixe à $f_{|I_\alpha}$, (I_α étant fermé), pour obtenir que $x^{(k)} \to \bar{x}$ où $\bar{x}$ est l'unique point fixe de $f_{|I\alpha}$.
Soit $x \in I_\alpha$; montrons d'abord que $f(x) \in I_\alpha$: comme $f \in C^1(\mathbb{R},\mathbb{R})$, il existe $\xi \in]x, \bar{x}[$ tel que $|f(x) - \bar{x}| = |f(x) - f(\bar{x})| = |f'(\xi)||x - \bar{x}| \leq \gamma|x - \bar{x}| < \alpha$, ce qui prouve que $f(x) \in I_\alpha$. On vérifie alors que $f_{|I_\alpha}$ est strictement contractante en remarquant que pour tous $x, y \in I_\alpha$, $x < y$, il existe $\xi \in]x, y[(\subset I_\alpha)$ tel que $|f(x) - f(y)| = |f'(\xi)||x - y| \leq \gamma|x - y|$ avec $\gamma < 1$. On a ainsi montré que $x^{(k)} \to \bar{x}$ si $x^{(0)} \in I_\alpha$.

   Cherchons maintenant la vitesse de convergence de la suite. Supposons que $f'(\bar{x}) \neq 0$ et $x^{(k)} \neq \bar{x}$ pour tout $n \in \mathbb{N}$. Comme $x^{(k+1)} = f(x^{(k)})$ et $\bar{x} = f(\bar{x})$, on a $|x^{(k+1)} - \bar{x}| = |f(x^{(k)}) - f(\bar{x})|$. Comme $f \in C^1(\mathbb{R},\mathbb{R})$, il existe $\xi_k \in]x^{(k)}, \bar{x}[$ ou $]\bar{x}, x^{(k)}[$, tel que $f(x^{(k)}) - f(\bar{x}) = f'(\xi_k)(x^{(k)} - \bar{x})$. On a donc
$$\frac{|x^{(k+1)} - \bar{x}|}{|x^{(k)} - \bar{x}|} = |f'(\xi_k)| \longrightarrow |f'(\bar{x})| \text{ car } x^{(k)} \to \bar{x} \text{ et } f' \text{ est continue.}$$
On a donc une convergence linéaire.

2. Supposons maintenant que $f \in C^2(\mathbb{R},\mathbb{R})$ et $f'(\bar{x}) = 0$. On sait déjà par ce qui précède qu'il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha$ alors $x^{(k)} \to \bar{x}$ lorsque $k \to +\infty$. On veut estimer la vitesse de convergence ; on suppose pour cela que $x^{(k)} \neq \bar{x}$ pour tout $k \in \mathbb{N}$. Comme $f \in C^2(\mathbb{R},\mathbb{R})$, il existe $\xi_k \in]x^{(k)}, \bar{x}[$ tel que
$$f(x^{(k)}) - f(\bar{x}) = f'(\bar{x})(x^{(k)} - \bar{x}) + \frac{1}{2}f''(\xi_k)(x^{(k)} - \bar{x})^2.$$

On a donc : $x^{(k+1)} - \bar{x} = \frac{1}{2} f''(\xi_k)(x^{(k)} - \bar{x})^2$ ce qui entraîne, par continuité de f'', que

$$\frac{|x^{(k+1)} - \bar{x}|}{|x^{(k)} - \bar{x}|^2} = \frac{1}{2}|f''(\xi_k)| \longrightarrow \frac{1}{2}|f''(\bar{x})| \text{ quand } k \to +\infty.$$

La convergence est donc au moins quadratique.

■

2.2.4 Méthode de Newton dans $\mathbb{R}$

On va étudier dans le paragraphe suivant la méthode de Newton pour la résolution d'un système non linéaire. (En fait, il semble que l'idée de cette méthode revienne plutôt à Simpson[4] Donnons l'idée de la méthode de Newton dans le cas $n = 1$ à partir des résultats de la proposition précédente. Soit $g \in C^3(\mathbb{R}, \mathbb{R})$ et $\bar{x} \in \mathbb{R}$ tel que $g(\bar{x}) = 0$. On cherche une méthode de construction d'une suite $(x^{(k)})_{k\in\mathbb{N}} \subset \mathbb{R}^n$ qui converge vers $\bar{x}$ de manière quadratique. On pose

$$f(x) = x - h(x)g(x) \text{ avec } h \in C^2(\mathbb{R}, \mathbb{R}) \text{ tel que } h(x) \neq 0 \ \ \forall x \in \mathbb{R},$$

et on a donc

$$f(x) = x \Leftrightarrow g(x) = 0.$$

Si par miracle $f'(\bar{x}) = 0$, la méthode de point fixe sur f va donner (pour $x^{(0)} \in I_\alpha$ donné par la proposition 2.16) $(x^{(k)})_{n\in\mathbb{N}}$ tel que $x^{(k)} \to \bar{x}$ de manière au moins quadratique. Or on a $f'(x) = 1 - h'(x)g(x) - g'(x)h(x)$ et donc $f'(\bar{x}) = 1 - g'(\bar{x})h(\bar{x})$. Il suffit donc de prendre h tel que $h\,(\bar{x}) = \dfrac{1}{g'(\bar{x})}$. Ceci est possible si $g'(\bar{x}) \neq 0$.

En résumé, si $g \in C^3(\mathbb{R}, \mathbb{R})$ est telle que $g'(\overline{x}) \neq 0 \ \forall x \in \mathbb{R}$ et $g(\bar{x}) = 0$, on peut construire, pour x assez proche de $\bar{x}$, la fonction $f \in C^2(\mathbb{R}, \mathbb{R})$ définie par

$$f(x) = x - \frac{g(x)}{g'(x)}.$$

Grâce à la proposition 2.16, il existe $\alpha > 0$ tel que si $x^{(0)} \in I_\alpha$ alors la suite définie par

$$x^{(k+1)} = f(x^{(k)}) = x^{(k)} - \frac{g(x^{(k)})}{g'(x^{(k)})}$$

converge vers $\bar{x}$ de manière au moins quadratique.

Remarquons que dans le cas $n = 1$, la suite de Newton peut s'obtenir naturellement en remplaçant l'équation $g(\overline{x}) = 0$ par $g(x^{(k+1)}) = 0$, et $g(x^{(k+1)})$ par le développement limité en x^k :

$$g(x^{(k+1)}) = g(x^{(k)}) + g'(x^{(k)})(x^{(k+1)} - x^{(k)}) + |x^{(k+1)} - x^{(k)}|\epsilon(x^{(k+1)} - x^{(k)}).$$

C'est le plus sûr moyen mnémotechnique pour retrouver l'itération de Newton :

$$g(x^{(k)}) + g'(x^{(k)})(x^{(k+1)} - x^{(k)}) = 0 \text{ ou encore } g'(x^{(k)})(x^{(k+1)} - x^{(k)}) = -g(x^{(k)}). \tag{2.10}$$

Comparons sur un exemple les méthodes de point fixe et de Newton. On cherche le zéro de la fonction $g : x \mapsto x^2 - 3$ sur $\mathbb{R}_+$. Notons en passant que la construction de la suite $x^{(k)}$ par point fixe ou Newton permet l'approximation effective de $\sqrt{3}$. Si on applique le point fixe standard, la suite $x^{(k)}$ s'écrit

$$x^{(0)} \text{ donné },$$
$$x^{(k+1)} = x^{(k)} - (x^{(k)})^2 + 3.$$

4. Voir Nick Kollerstrom (1992). *Thomas Simpson and "Newton's method of approximation" : an enduring myth*, The British Journal for the History of Science, 25, pp 347-354 doi :10.1017/S0007087400029150 – Thomas Simpson est un mathématicien anglais du 18-ème siècle à qui on attribue généralement la méthode du même nom pour le calcul approché des intégrales, probablement à tort car celle-ci apparaît déjà dans les travaux de Kepler deux siècles plus tôt !

Si on applique le point fixe avec paramètre de relaxation ω, la suite $x^{(k)}$ s'écrit

$$x^{(0)} \text{ donné },$$
$$x^{(k+1)} = -x^{(k)} + \omega(-x^{(k)})^2 + 3)$$

Si maintenant on applique la méthode de Newton, la suite $x^{(k)}$ s'écrit

$$x^{(0)} \text{ donné },$$
$$x^{(k+1)} = -\frac{(x^{(k)})^2 - 3}{2x^{(k)}}.$$

Comparons les suites produites par scilab à partir de $x^{(0)} = 1$ par le point fixe standard, le point fixe avec relaxation ($\omega = .1$) et la méthode de Newton.

— **point fixe standard :** 1. 3. - 3 - 9 - 87 - 7653 - 58576059 - 3.431D+15 - 1.177D+31

— **point fixe avec relaxation :**
1. 1.2 1.356 1.4721264 1.5554108 1.6134805 1.6531486 1.6798586 1.6976661
1.7094591 1.717234 1.7223448 1.7256976 1.7278944 1.7293325 1.7302734 1.7308888
1.7312912 1.7315543 1.7317263 1.7318387 1.7319122 1.7319602 1.7319916 1.7320121
1.73204 1.7320437 1.7320462 1.7320478 1.7320488 1.7320495 1.73205 1.7320503
1.7320504 1.7320506 1.7320507 1.7320507 1.7320507 1.7320508

— **Newton :**
1. 2. 1.75 1.7321429 1.7320508 1.7320508

Remarque 2.17 (Attention à l'utilisation du théorème des accroissements finis...)**.** *On a fait grand usage du théorème des accroissements finis dans ce qui précède. Rappelons que sous la forme qu'on a utilisée, ce théorème n'est valide que pour les fonctions de* $\mathbb{R}$ *dans* $\mathbb{R}$. *On pourra s'en convaincre en considérant la fonction de* $\mathbb{R}$ *dans* $\mathbb{R}^2$ *définie par :*

$$\varphi(x) = \begin{bmatrix} \sin x \\ \cos x \end{bmatrix}.$$

On peut vérifier facilement qu'il n'existe pas de $\xi \in \mathbb{R}$ *tel que* $\varphi(2\pi) - \varphi(0) = 2\pi\varphi'(\xi)$.