Agrégation interne de mathématiques

CONTINUITÉ, DÉRIVABILITÉ ET CONVEXITÉ

Jean-Etienne Rombaldi

1-2046-TE-PA-02-16

Table des matières

1

Continuité, dérivabilité des fonctions d'une variable réelle

Pour ce chapitre I désigne un intervalle réel non réduit à un point et f une fonction définie sur I et à valeurs réelles ou complexes.

1.1 Continuité en un point, continuité sur I

Définition 1.1 *On dit que la fonction f est continue au point $a \in I$ si :*

$$\forall \varepsilon > 0,\ \exists \eta > 0 \mid \forall x \in I \cap \left]a - \eta, a + \eta\right[,\ \left|f(x) - f(a)\right| < \varepsilon \tag{1.1}$$

Dans le cas des fonctions d'une variable réelle qui nous occupe, on peut aussi définir les notions de continuité à gauche ou à droite.

Définition 1.2 *On dit que la fonction est continue à gauche [resp. à droite] au point $a \in I$ si :*

$$\forall \varepsilon > 0,\ \exists \eta > 0 \mid \forall x \in I \cap \left]a - \eta, a\right[,\ \left|f(x) - f(a)\right| < \varepsilon$$

$$[\textit{resp. } \forall \varepsilon > 0,\ \exists \eta > 0 \mid \forall x \in I \cap \left]a, a + \eta\right[,\ \left|f(x) - f(a)\right| < \varepsilon]$$

Dans le cas où a est l'extrémité gauche [resp. droite] de I, on ne s'intéresse qu'à la continuité à droite [resp. à gauche].

La fonction f est discontinue en $a \in I$ si, et seulement si, l'une des deux situations suivantes se produit :

- soit f n'a pas de limite à gauche ou à droite en a ;
- soit ces deux limites existent et l'une d'elles est distincte de $f(a)$.

Des définitions, on déduit immédiatement les résultats suivants.

Théorème 1.1 *La fonction f est continue [resp. continue à gauche [resp. à droite]] au point $a \in I$ si, et seulement si, $\lim\limits_{x \to a} f(x) = f(a)$ [resp. $\lim\limits_{x \to a^-} f(x) = f(a)$ [resp. $\lim\limits_{x \to a^+} f(x) = f(a)$]].*

Théorème 1.2 *Si a est un point intérieur à l'intervalle I, alors la fonction f est continue en a si, et seulement si, elle est continue à gauche et à droite en a.*

Théorème 1.3 *Si f est continue en $a \in I$, elle est alors bornée dans un voisinage de ce point.*

Le résultat précédent peut être utilisé pour montrer la discontinuité d'une fonction en un point. Par exemple la fonction f définie sur $\mathbb{R}$ par $f(0) = 0$ et $f(x) = \frac{1}{x}$ pour $x \neq 0$ n'est pas continue en 0, puisque pour tout réel $M > 0$, il existe un entier $n \geq 1$ tel que $f\left(\frac{1}{n}\right) = n > M$.

Définition 1.3 *On dit que la fonction f est continue sur I si elle est continue en tout point de I.*

Exercice 1.1 *Montrer que, pour tout entier naturel n, la fonction $f_n : x \mapsto x^n$ est continue sur $\mathbb{R}$.*

Solution 1.1 *Pour $n = 0$, f_0 est la fonction constante égale à 1. La continuité d'une fonction constante se vérifiant facilement (pour $\varepsilon > 0$ donné tout réel $\eta > 0$ convient).*
Pour $n \geq 1$, $a \in \mathbb{R}$, on peut trouver un réel $R > 0$ tel que $a \in \left]-R, R\right[$ et pour tout $x \in \left]-R, R\right[$, on a :

$$|x^n - a^n| = |x - a| \left| \sum_{k=0}^{n-1} x^{n-1-k} a^k \right| \leq nR^{n-1} |x - a|$$

et pour $\varepsilon > 0$ donné, on aura $|x^n - a^n| < \varepsilon$ dès que $|x - a| < \eta = \frac{\varepsilon}{nR^{n-1}}$.

Exercice 1.2 *Montrer que, pour tout entier naturel non nul n, la fonction $f_n : x \mapsto x^{\frac{1}{n}}$ est continue sur $\mathbb{R}^{+,*}$.*

Solution 1.2 *Pour $n \geq 1, a > 0$, on peut trouver un réel $R > 0$ tel que $a \in \left]\frac{1}{R}, R\right[$ et pour tout $x \in \left]\frac{1}{R}, R\right[$, on a :*

$$\begin{aligned} |x - a| &= \left| \left(x^{\frac{1}{n}}\right)^n - \left(a^{\frac{1}{n}}\right)^n \right| = \left| x^{\frac{1}{n}} - a^{\frac{1}{n}} \right| \left(\sum_{k=0}^{n-1} \left(x^{\frac{1}{n}}\right)^{n-1-k} \left(a^{\frac{1}{n}}\right)^k \right) \\ &\geq n \left(\frac{1}{R}\right)^{\frac{n-1}{n}} \left| x^{\frac{1}{n}} - a^{\frac{1}{n}} \right| \end{aligned}$$

donc :

$$\left| x^{\frac{1}{n}} - a^{\frac{1}{n}} \right| \leq \frac{1}{n} R^{\frac{n-1}{n}} |x - a| \leq R^{\frac{n-1}{n}} |x - a|$$

et pour $\varepsilon > 0$ donné, on aura $\left| x^{\frac{1}{n}} - a^{\frac{1}{n}} \right| < \varepsilon$ dès que $|x - a| < \eta = \frac{\varepsilon}{R^{\frac{n-1}{n}}}$.

Exercice 1.3 *Montrer que la fonction définie sur $\mathbb{R}$ par $f(0) = 0$ et $f(x) = \cos\left(\frac{1}{x}\right)$ si $x \neq 0$, n'est pas continue en 0.*

Solution 1.3 *Pour tout réel $\eta > 0$, on peut trouver un entier $n \geq 1$ tel que $x = \frac{1}{n\pi} \in \left]-\eta, \eta\right[$ et on a $|f(x) - f(0)| = 1$, ce qui prouve la discontinuité de f en 0.*

Exercice 1.4 *Montrer que la fonction caractéristique de $\mathbb{Q}$ définie par $f(x) = 1$ si $x \in \mathbb{Q}$ et $f(x) = 0$ sinon est discontinue en tout point de $\mathbb{R}$.*

Solution 1.4 *Soit a un nombre rationnel [resp. irrationnel]. Pour tout réel $\eta > 0$, on peut trouver un nombre irrationnel [resp. rationnel] x dans $]a-\eta, a+\eta[$ et on a $|f(x) - f(a)| = 1$, ce qui prouve la discontinuité de f en a.*

Exercice 1.5 *Montrer que la fonction définie sur $\mathbb{R}$ par $f(x) = x$ si $x \in \mathbb{Q}$ et $f(x) = 0$ sinon est continue en 0 et discontinue en tout point de $\mathbb{R}^*$.*

Solution 1.5 *Soit a un nombre rationnel [resp. irrationnel] non nul.*
Pour tout réel $\eta > 0$, on peut trouver un nombre irrationnel [resp. rationnel] x dans $]a-\eta, a+\eta[$ [resp. dans $]a-\eta, a+\eta[\cap \left]a - \frac{|a|}{2}, a + \frac{|a|}{2}\right[$] et on a :

$$|f(x) - f(a)| = |a| \text{ [resp. } |f(x) - f(a)| = |x| > \frac{|a|}{2}\text{]}$$

ce qui prouve la discontinuité de f en a.
Pour ce qui est de la continuité en 0, il suffit de remarquer que pour tout $\varepsilon > 0$ la condition $|x| < \eta = \varepsilon$ entraîne $|f(x) - f(0)| = |f(x)| < \varepsilon$ puisque $f(x)$ vaut x ou 0.

Exercice 1.6 *Soit n un entier naturel non nul. Étudier la continuité de la fonction f_n qui à tout réel x appartenant à $[0,1]$ associe la n-ème décimale dans le développement décimal propre si $x \in [0,1[$ et 9 si $x = 1$ (pour $x = 1$ on utilise le développement décimal illimité impropre).*

Solution 1.6 *On rappelle que pour $x \in [0,1[$, on a $f_n(x) = [10^n x] - 10[10^{n-1}x]$. Pour tout entier k compris entre 0 et $10^n - 1$ et $x \in \left[\frac{k}{10^n}, \frac{k+1}{10^n}\right[$, on a $f_n(x) = k - 10[10^{n-1}x]$. De plus avec :*

$$10\left[10^{n-1}x\right] \leq 10^n x < k+1$$

on déduit que :

$$\left[10^{n-1}x\right] \leq \frac{k}{10} \leq 10^{n-1}x < \left[10^{n-1}x\right] + 1$$

et $\left[\frac{k}{10}\right] = [10^{n-1}x]$. La fonction f_n est donc constante (et en conséquence continue) sur $\left[\frac{k}{10^n}, \frac{k+1}{10^n}\right[$ égale à $k - 10\left[\frac{k}{10}\right]$. Et avec :

$$f_n\left(\frac{k+1}{10^n}\right) = k + 1 - 10\left[\frac{k+1}{10}\right] \neq \lim_{x \to \left(\frac{k+1}{10^n}\right)^-} f_n(x) = k - 10\left[\frac{k}{10}\right]$$

on déduit que f_n est discontinue en tout point $\frac{k+1}{10^n}$ où k est compris entre 0 et $10^n - 2$. Sur le dernier intervalle $\left[\frac{10^n - 1}{10^n}, 1\right[$, on a $f_n(x) = 9$ et f_n est continue en 1.

Définition 1.4 *Si a est intérieur à I et si la fonction f est discontinue en a avec des limites à droite et à gauche en ce point, on dit alors que f a une discontinuité de première espèce en a.*

Le cas des fonctions monotones définies sur un intervalle ouvert (pour simplifier) est particulièrement intéressant.

Théorème 1.4 *Si f est une fonction monotone de l'intervalle ouvert I dans $\mathbb{R}$, alors f admet une limite à gauche et droite en tout point. Dans le cas où f est croissante, on a pour tout $x \in I$:*

$$f\left(x^-\right) = \sup_{t\in I,\ t<x} f(t) \leq f(x) \leq f\left(x^+\right) = \inf_{t\in I,\ x<t} f(t)$$

De plus pour $x < y$ dans I, on a $f(x^+) \leq f(y^-)$.

Démonstration. Quitte à remplacer f par $-f$, on peut supposer f croissante.

Pour $x \in I$, l'ensemble $A = \{f(t) \mid t \in I,\ t < x\}$ est non vide majoré par $f(x)$, il admet donc une borne supérieure $\mu = \sup\limits_{t\in I,\ t<x} f(t) \leq f(x)$.

Par définition de la borne supérieure, pour tout réel $\varepsilon > 0$, il existe $x_0 \in \left]a, x\right[$ tel que $\mu - \varepsilon < f(x_0) \leq \mu$ et avec la croissance de f, on a :

$$\forall t \in \left]x_0, x\right[,\ \mu - \varepsilon < f(x_0) \leq f(t) \leq \mu$$

On a donc ainsi montré que $\mu = \lim\limits_{t\to x^-} f(t) = f(x^-)$.

On procède de même pour l'existence de la limite à droite $f(x^+)$.

Pour $x < y$ dans I, on a :

$$f\left(x^+\right) = \inf_{t\in I,\ x<t} f(t) = \inf_{x<t<y} f(t),\ f\left(y^-\right) = \sup_{t\in I,\ t<y} f(t) = \sup_{x<t<y} f(t)$$

ce qui entraîne $f(x^+) \leq f(y^-)$. ■

Théorème 1.5 *Si f est une fonction monotone d'un intervalle ouvert I dans $\mathbb{R}$, alors l'ensemble de ses points de discontinuité est au plus dénombrable.*

Démonstration. Supposons f croissante et l'ensemble D des points de discontinuité de f non vide.

Pour tout $x \in D$, on a $f(x^-) < f(x^+)$ et on peut trouver un rationnel $r(x) \in \left]f(x^-), f(x^+)\right[$. De plus pour $x < y$ dans D avec $f(x^+) \leq f(y^-)$, on déduit que $r(x) < r(y)$. L'application r définit donc une injection de D dans $\mathbb{Q}$, il en résulte que D est dénombrable. ■

En général une fonction monotone f définie sur un intervalle I n'est pas nécessairement continue, mais nous verrons plus loin que si de plus $f(I)$ est un intervalle, alors elle est continue (théorème 2.8).

Exemple 1.1 *La fonction f définie sur $[0,1]$ par $f(0) = 0$ et $f(x) = \dfrac{1}{\left[\frac{1}{x}\right]}$ est croissante avec une infinité de points de discontinuité.*

En fait, on peut montrer le résultat suivant.

Théorème 1.6 *Si I est un intervalle ouvert, l'ensemble des points de discontinuité de première espèce de f est au plus dénombrable.*

Démonstration. Voir [40], chapitre 4, exercice 17. ■

1.2 Définition séquentielle de la continuité en un point

Une définition équivalente de la continuité en un point est donnée par le résultat suivant.

Théorème 1.7 *La fonction f est continue en $a \in I$ si, et seulement si, pour toute suite $(x_n)_{n\in\mathbb{N}}$ de points de I qui converge vers a, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$.*

Démonstration. Si f est continue en $a \in I$, alors pour tout $\varepsilon > 0$ il existe $\eta > 0$ tel que $|x-a| < \eta$ dans I entraîne $|f(x) - f(a)| < \varepsilon$ et si $(x_n)_{n\in\mathbb{N}}$ est une suite de points de I qui converge vers a, il existe alors un entier n_0 tel que $|x_n - a| < \eta$ pour tout $n \geq n_0$, ce qui implique $|f(x_n) - f(a)| < \varepsilon$. On a donc bien $\lim\limits_{n\to+\infty} f(x_n) = f(a)$.

Pour la réciproque, on raisonne par l'absurde. Si f n'est pas continue en a, il existe un réel $\varepsilon > 0$ tel que pour tout entier $n \geq 1$ on peut trouver $x_n \in I$ tel que $|x_n - a| < \frac{1}{n}$ et $|f(x_n) - f(a)| \geq \varepsilon$. On a donc ainsi une suite $(x_n)_{n\in\mathbb{N}}$ de points de I qui converge vers a pour laquelle la suite $(f(x_n))_{n\in\mathbb{N}}$ ne converge pas vers $f(a)$. ■

En fait on a le résultat plus fin suivant.

Exercice 1.7 *Montrer que $f : I \to \mathbb{C}$ est continue en $a \in I$ si, et seulement si, pour toute suite $(x_n)_{n\in\mathbb{N}}$ de points de I qui converge vers a, la suite $(f(x_n))_{n\in\mathbb{N}}$ est convergente (sans préciser que c'est vers $f(a)$).*

Solution 1.7 *On sait déjà que si f est continue en a alors pour toute suite $(x_n)_{n\in\mathbb{N}}$ de points de I qui converge vers a, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$.*
Réciproquement supposons que pour toute suite $(x_n)_{n\in\mathbb{N}}$ de points de I qui converge vers a, la suite $(f(x_n))_{n\in\mathbb{N}}$ soit convergente. Pour montrer que f est continue en a, il suffit de montrer que, dans ces conditions, la suite $(f(x_n))_{n\in\mathbb{N}}$ converge vers $f(a)$. Si $(x_n)_{n\in\mathbb{N}}$ est une suite de points de I qui converge vers a, on définit la suite $(y_n)_{n\in\mathbb{N}}$ de points de I par $y_{2n} = x_{2n}$, $y_{2n+1} = a$, cette suite converge vers a, donc la suite $(f(y_n))_{n\in\mathbb{N}}$ converge et on a :

$$f(a) = \lim_{n\to+\infty} f(y_{2n+1}) = \lim_{n\to+\infty} f(y_{2n}) = \lim_{n\to+\infty} f(x_{2n}) = \lim_{n\to+\infty} f(x_n)$$

Le théorème 1.7 est souvent utilisé pour montrer qu'une fonction n'est pas continue en un point.

Exercice 1.8 *Montrer que la fonction f définie sur $\mathbb{R}$ par $f(0) = 0$ et $f(x) = \cos(\ln(|x|))$ si $x \neq 0$, n'est pas continue en 0.*

Solution 1.8 *Si $(x_n)_{n\geq 1}$ est la suite définie par $x_n = e^{-n\pi}$ pour $n \geq 1$, on a $\lim\limits_{n\to+\infty} x_n = 0$ et la suite $(f(x_n))_{n\geq 1} = (\cos(n\pi))_{n\geq 1} = ((-1)^n)_{n\geq 1}$ est divergente.*

Exercice 1.9 *Soit $f :]0,1[\to \mathbb{R}$ la fonction définie par $f(x) = 0$ si x est irrationnel et par $f(x) = \frac{1}{q}$ si $x = \frac{p}{q}$ est rationnel où p, q sont entiers naturels non nuls premiers entre eux. Montrer que f est continue en tout point irrationnel et discontinue en tout point rationnel de $]0,1[$.*

Solution 1.9 *Un rationnel $r = \frac{p}{q} \in]0,1[\cap \mathbb{Q}$ est limite de la suite de nombres irrationnels $(x_n)_{n\geq n_0} = \left(r + \frac{\sqrt{2}}{n}\right)_{n\geq n_0}$, où n_0 est choisi assez grand pour que cette suite soit à valeurs dans $]0,1[$, et $\lim\limits_{n\to+\infty} f(x_n) = 0 \neq f(r) = \frac{1}{q}$. La fonction f n'est donc pas continue en ce point.*

Soit $\varepsilon > 0$. *Si* $a \in \left]0,1\right[\cap (\mathbb{R} \setminus \mathbb{Q})$ *et* $\eta > 0$ *est tel que* $\left]a-\eta, a+\eta\right[\subset \left]0,1\right[$, *on note :*

$$E = \{x \in \left]a-\eta, a+\eta\right[\mid f(x) > \varepsilon\}$$

Un élément de E *est nécessairement rationnel (sinon* $f(x) = 0 < \varepsilon$*), il s'écrit donc* $r = \dfrac{p}{q}$ *avec* p, q *premiers entre eux et* $f(r) = \dfrac{1}{q} > \varepsilon$ *entraîne que* E *est vide ou que* $1 \leq q < \dfrac{1}{\varepsilon}$ *et* $1 \leq p < q < \dfrac{1}{\varepsilon}$ *(*r *est strictement compris entre 0 et 1). L'ensemble* E *est donc vide ou fini. Pour* $0 < \eta' < \eta$ *assez petit on aura alors* $E \cap \left]a-\eta', a+\eta'\right[= \emptyset$*, ce qui signifie que* $0 \leq f(x) \leq \varepsilon$ *pour tout* $x \in \left]a-\eta', a+\eta'\right[$. *On a donc ainsi montré que* f *est continue en* a.

Exercice 1.10 *Soit* $f : \mathbb{R} \to \mathbb{R}$ *une fonction continue telle que* $f(ax+b) = f(x)$ *pour tout réel* x*, où* a, b *sont deux constantes réelles avec* $|a| \neq 1$. *Montrer que* f *est nécessairement constante.*

Solution 1.10 *De* $f(ax+b) = f(x)$, *on déduit que :*

$$f(a(ax+b)+b) = f(ax+b) = f(x)$$

soit :

$$f\left(a^2x + b(a+1)\right) = f(x)$$

Par récurrence, on montre alors que pour tout entier $n \geq 1$ *on a :*

$$\forall x \in \mathbb{R},\ f\left(a^n x + b\sum_{k=0}^{n-1} a^k\right) = f(x)$$

Le résultat est vrai pour $n = 1$ *et en le supposant vrai au rang* $n \geq 1$*, on a :*

$$f\left(a\left(a^n x + b\sum_{k=0}^{n-1} a^k\right) + b\right) = f\left(a^n x + b\sum_{k=0}^{n-1} a^k\right) = f(x)$$

avec :

$$a\left(a^n x + b\sum_{k=0}^{n-1} a^k\right) + b = a^{n+1}x + b\sum_{k=0}^{n} a^k$$

Comme $|a| \neq 1$*, on peut écrire :*

$$\forall n \in \mathbb{N}^*,\ \forall x \in \mathbb{R},\ f\left(a^n x + b\frac{a^n - 1}{a-1}\right) = f(x)$$

et pour $|a| < 1$*, avec la continuité de* f*, on déduit que pour tout réel* x *on a :*

$$f(x) = \lim_{n \to +\infty} f\left(a^n x + b\frac{a^n - 1}{a-1}\right) = f\left(\frac{b}{1-a}\right)$$

c'est-à-dire que f *est constante.*

Pour traiter le cas où $|a| > 1$*, on peut remarquer, en faisant le changement de variable* $t = ax + b$*, que la condition* $f(ax+b) = f(x)$ *est équivalente à* $f(t) = f\left(\dfrac{1}{a}t - \dfrac{b}{a}\right)$ *pour tout réel* t *avec* $\left|\dfrac{1}{a}\right| < 1$*, ce qui entraîne que* f *est constante.*

Une application importante du théorème 1.7 est le résultat de point fixe suivant (voir le lemme 8.1).

Théorème 1.8 *Si $f : I \to \mathbb{R}$ est telle que $f(I) \subset I$ et si la suite $(x_n)_{n\in\mathbb{N}}$ de points de I définie par la donnée de $x_0 \in I$ et la relation de récurrence $x_{n+1} = f(x_n)$ converge vers un point $\alpha \in I$ en lequel la fonction f est continue, alors $f(\alpha) = \alpha$, c'est-à-dire que α est un point fixe de f.*

Une autre application importante du théorème 1.7 est le principe de prolongement des identités [resp. des inégalités] : si f, g sont deux fonctions continues en tout point d'un intervalle I et qui coïncident [resp. telles que $f(x) \leq g(x)$] en tout point d'une partie D dense dans I (par exemple les nombres rationnels ou les nombres décimaux), alors elle sont égales [resp. $f(x) \leq g(x)$ pour tout $x \in I$].

Un exemple d'utilisation de ce principe est donné par le résultat suivant.

Exercice 1.11 *Soit $f : \mathbb{R} \to \mathbb{R}$ vérifiant l'équation fonctionnelle de Cauchy :*

$$\forall (x, y) \in \mathbb{R},\ f(x + y) = f(x) + f(y)$$

Montrer que si f est continue en un point, elle est alors continue en tout point et on a $f(x) = f(1)x$ pour tout réel x.

Solution 1.11 *Si f est continue en α, en écrivant, pour tout réel x_0, que :*

$$f(x_0 + h) - f(x_0) = f(h) = f(h + \alpha) - f(\alpha)$$

on déduit que f est continue en x_0.
On vérifie ensuite facilement que $f(r) = f(1)r$ pour tout $r \in \mathbb{Q}$ et la densité de $\mathbb{Q}$ dans $\mathbb{R}$ permet de conclure.

1.3 Prolongement par continuité

Si a est un réel adhérent à I et si f a une limite finie en ce point on peut alors prolonger f par continuité en ce point. Précisément on a le résultat suivant.

Théorème 1.9 *Si a est un réel adhérent à I n'appartenant pas à I (un point frontière) et si la fonction f a une limite ℓ en a, il existe alors un unique prolongement de f à $I \cup \{a\}$ qui est continu en a, ce prolongement est défini par $\widetilde{f}(x) = f(x)$ si $x \in I$ et $\widetilde{f}(a) = \ell$.*

Démonstration. Il est clair que la fonction $\widetilde{f}$ est un prolongement de f continu en a.

Réciproquement si g est un tel prolongement, on a $g = f$ sur I et par continuité en a, $g(a) = \lim\limits_{\substack{x\to a \\ x\neq a}} g(x) = \lim\limits_{\substack{x\to a \\ x\neq a}} f(x) = \ell = \widetilde{f}(a)$. D'où l'unicité. ■

La fonction f définie par le théorème précédent est appelée le prolongement par continuité de f en a. On le note souvent f au lieu de $\widetilde{f}$.

Exemple 1.2 *La fonction $f : x \mapsto x \sin\left(\frac{1}{x}\right)$ définie sur $\mathbb{R}^*$ se prolonge par continuité en 0 en posant $f(0) = 0$.*

Exemple 1.3 *La fonction $f : x \mapsto \frac{\sin(x)}{x}$ définie sur $\mathbb{R}^*$ se prolonge par continuité en 0 en posant $f(0) = 1$.*

Exemple 1.4 *La fonction $f : x \mapsto \sin\left(\frac{1}{x}\right)$ définie sur $\mathbb{R}^*$ ne peut pas se prolonger par continuité en 0 du fait qu'elle n'a pas de limite en ce point.*

1.4 Continuité et opérations sur les fonctions

Les résultats relatifs à la continuité et les opérations usuelles sur les fonctions sont résumés par le théorème qui suit.

Théorème 1.10 *Soient f, g deux fonctions définies sur I, à valeurs réelles ou complexes et continues en $a \in I$. Les fonctions $|f|$, $f+g$, fg, $\min(f,g)$ et $\max(f,g)$ (pour f, g à valeurs réelles) sont continues en a.*

Démonstration. La continuité de $|f|$ résulte de :

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)| \underset{x \to \alpha}{\to} 0$$

Pour la somme, la continuité se déduit de l'inégalité triangulaire :

$$|(f+g)(x) - (f+g)(a)| \leq |f(x) - f(a)| + |g(x) - g(a)| \underset{x \to \alpha}{\to} 0$$

Pour le produit, la continuité se déduit de l'inégalité triangulaire et du fait que f (ou g) est bornée au voisinage de a :

$$\begin{aligned} |(fg)(x) - (fg)(a)| &\leq |f(x)| \, |g(x) - g(a)| + |g(a)| \, |f(x) - f(a)| \\ &\leq M \, |g(x) - g(a)| + |g(a)| \, |f(x) - f(a)| \underset{x \to \alpha}{\to} 0 \end{aligned}$$

où M est un majorant de $|f|$ au voisinage de a.

Enfin la continuité de $\min(f,g)$ et $\max(f,g)$, pour f, g réelles, se déduit de :

$$\begin{cases} \min(f,g) = \dfrac{1}{2}(f + g - |f - g|) \\ \max(f,g) = \dfrac{1}{2}(f + g + |f - g|) \end{cases}$$

($\dfrac{f+g}{2}$ est le milieu de l'intervalle $[\min(f,g), \max(f,g)]$ et $|f-g|$ est sa longueur). ■

Avec la continuité de la fonction constante égale à 1 et de la fonction $x \mapsto x$ en tout point de $\mathbb{R}$, on en déduit la continuité des fonctions polynomiales.

Pour les quotients de fonctions continues, on a besoin du résultat suivant.

Lemme 1.1 *Si $f : I \to \mathbb{C}$ est continue en $a \in I$ avec $f(a) \neq 0$, il existe alors un voisinage $\mathcal{V}$ de a dans I tel que $f(x) \neq 0$ pour tout $x \in \mathcal{V}$.*

Démonstration. Comme $f(a) \neq 0$, on a $|f(a)| > 0$ et pour $\varepsilon \in]0, |f(a)|[$ on peut trouver un réel $\eta > 0$ tel que $|x - a| < \eta$ dans I entraîne

$$||f(x)| - |f(a)|| \leq |f(x) - f(a)| < \varepsilon$$

ce qui implique $|f(x)| > |f(a)| - \varepsilon > 0$ pour tout x dans ce voisinage. ■

Dans le cas des fonctions à valeurs réelles, on a le résultat plus précis suivant qui est souvent utilisé.

Lemme 1.2 *Si $f : I \to \mathbb{R}$ est continue en $a \in I$ avec $f(a) > 0$ [resp. $f(a) < 0$], il existe alors un voisinage $\mathcal{V}$ de a dans I tel que $f(x) > 0$ [resp. $f(x) < 0$] pour tout $x \in \mathcal{V}$.*

Démonstration. Supposons $f(a) > 0$. Pour $\varepsilon \in]0, f(a)[$ on peut trouver un réel $\eta > 0$ tel que $|x-a| < \eta$ dans I entraîne $|f(x) - f(a)| < \varepsilon$, ce qui implique $f(x) > f(a) - \varepsilon > 0$ pour tout x dans ce voisinage. ■

Théorème 1.11 *Si f est continue en $a \in I$ avec $f(a) \neq 0$, alors la fonction $\frac{1}{f}$ est définie dans un voisinage de a et est continue en ce point.*

Démonstration. En gardant les notations de la démonstration du lemme 1.1, on a $|f(x)| > |f(a)| - \varepsilon > 0$ pour x voisin de a et :

$$\left|\frac{1}{f(x)} - \frac{1}{f(a)}\right| = \left|\frac{f(x) - f(a)}{f(x)f(a)}\right| \leq \frac{|f(x) - f(a)|}{(|f(a)| - \varepsilon)|f(a)|} \underset{x \to \alpha}{\to} 0$$

■

De ce résultat on déduit la continuité de $\frac{f}{g}$ en a si f, g sont continues en ce point avec $g(a) \neq 0$.

De la continuité des fonctions polynomiales, on déduit la continuité des fonctions rationnelles en tout point de leur domaine de définition.

De la continuité des fonctions sin et cos sur $\mathbb{R}$, on déduit la continuité de la fonction tan sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.

Pour la composition des applications, on a le résultat suivant.

Théorème 1.12 *Si $f : I \to \mathbb{R}$ est continue en $a \in I$, J est un intervalle réel contenant $f(I)$ et $g : J \to \mathbb{C}$ est continue en $b = f(a)$, alors $g \circ f$ est continue en a.*

Démonstration. Pour $\varepsilon > 0$ donné, on peut trouver un réel $\eta > 0$ tel que $|y - b| < \eta$ dans l'intervalle J entraîne $|g(y) - g(b)| < \varepsilon$ et il existe $\eta' > 0$ tel que $|x - a| < \eta'$ dans l'intervalle I entraîne $|f(x) - f(a)| < \eta$, ce qui implique $|(g \circ f)(x) - (g \circ f)(a)| < \varepsilon$ pour tout x dans I tel que $|x - a| < \eta'$. ■

La continuité de $\frac{1}{f}$ peut se retrouver en composant la fonction f dans un voisinage de α où elle est non nulle avec $y \mapsto \frac{1}{y}$.

Exemple 1.5 *Avec les exercices 1.1 et 1.2, on déduit la continuité sur $\mathbb{R}^{+,*}$ de $x \mapsto x^r$ pour tout rationnel r.*

1.5 Propriétés globales des fonctions continues

1.5.1 Le théorème des valeurs intermédiaires

Voir le chapitre 2.

1.5.2 Continuité uniforme

Une notion importante est celle d'uniforme continuité.

Définition 1.5 *On dit que f est uniformément continue sur I si :*

$$\forall \varepsilon > 0,\ \exists \eta > 0 \mid \left((x, y) \in I^2,\ |x - y| \leq \eta\right) \Rightarrow |f(x) - f(y)| \leq \varepsilon$$

Une fonction uniformément continue sur I est évidemment continue en tout point de I, la nuance est, dans le cas de l'uniforme continuité, qu'un réel η associé à ε ne dépend que de f, I et ε.

Exemple 1.6 *Une fonction lipschitzienne (c'est-à-dire telle qu'il existe un réel $\lambda \geq 0$ avec $|f(x) - f(y)| \leq \lambda |x - y|$ pour tous x, y dans I) est uniformément continue sur I.*

Exercice 1.12 *Montrer que la fonction $x \mapsto \sqrt{x}$ est uniformément continue sur $\mathbb{R}^+$.*

Solution 1.12 *Ce résultat se déduit de :*

$$\forall (x, y) \in \mathbb{R}^+ \times \mathbb{R}^+, \ \left|\sqrt{x} - \sqrt{y}\right| \leq \sqrt{|x - y|}.$$

Cette inégalité est triviale pour $x = y$ et pour $y > x \geq 0$ (x, y jouent des rôles symétriques), on a :

$$\left|\sqrt{x} - \sqrt{y}\right|^2 = y - 2\sqrt{xy} + x < y - x.$$

On peut utiliser les suites pour traduire l'uniforme continuité.

Théorème 1.13 *Si $f : I \to \mathbb{R}$ est uniformément continue et si $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ sont deux suites d'éléments de I telles que $\lim\limits_{n\to+\infty} (u_n - v_n) = 0$, alors $\lim\limits_{n\to+\infty} (f(u_n) - f(v_n)) = 0$.*

Démonstration. Résulte de la définition. ■

Ce résultat peut être utilisé pour montrer qu'une fonction n'est pas uniformément continue sur I.

Exercice 1.13 *Montrer que la fonction $f : x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$.*

Solution 1.13 *En utilisant les suites $(u_n)_{n\in\mathbb{N}}$ et $(v_n)_{n\in\mathbb{N}}$ définies par $u_n = \sqrt{n+1}$ et $v_n = \sqrt{n}$, on a $u_n - v_n = \dfrac{1}{\sqrt{n} + \sqrt{n+1}}$, donc $\lim\limits_{n\to+\infty} (u_n - v_n) = 0$ et avec $f(u_n) - f(v_n) = 1$ on a $\lim\limits_{n\to+\infty} (f(u_n) - f(v_n)) = 1 \neq 0$, donc la fonction f n'est pas uniformément continue sur $\mathbb{R}$.*
Une démonstration directe de cette non uniforme continuité peut se faire comme suit : pour $\eta > 0$, $x = \dfrac{1}{\eta}$, $y = x + \dfrac{\eta}{2}$, on a $|x - y| < \eta$ et $|y^2 - x^2| = 1 + \dfrac{\eta^2}{4} > 1$.

Exercice 1.14 *Montrer que la fonction $f : x \mapsto \sin(x^2)$ n'est pas uniformément continue sur $\mathbb{R}$.*

Solution 1.14 *En considérant les suites introduites avec l'exercice précédent, on a :*

$$f(u_n) - f(v_n) = \sin(n+1) - \sin(n) = 2\cos\left(n + \frac{1}{2}\right)\sin\left(\frac{1}{2}\right)$$

et cette suite est divergente.

Les exercices qui précèdent nous montrent qu'une fonction continue n'est pas nécessairement uniformément continue.

Dans le cas où I est un intervalle fermé borné, nous verrons que les notions de continuité et d'uniforme continuité sont équivalentes (théorème 1.15).

Exercice 1.15 *Montrer qu'une fonction f peut très bien être uniformément continue sur tout intervalle strictement contenu dans I sans être uniformément continue sur I tout entier.*

Solution 1.15 *C'est le cas, par exemple, pour la fonction $f : x \mapsto \dfrac{1}{x}$ sur $I =]0,1]$. Elle est lipschitzienne sur tout $[a,1]$ où $0 < a < 1$, donc uniformément continue sur ces intervalles. Mais pour tout réel $\eta \in \left]0, \dfrac{1}{2}\right[$, $x = \eta$, $y = x + \dfrac{\eta}{2}$, on a $|y - x| < \eta$ avec $\left|\dfrac{1}{x} - \dfrac{1}{y}\right| = \dfrac{1}{3\eta} > \dfrac{2}{3}$.*

Exercice 1.16 *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction uniformément continue. Montrer qu'il existe deux constantes réelles a, b telles que $|f(x)| \leq a|x| + b$ pour tout réel x. En déduire que la fonction $x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$.*

Solution 1.16 *Avec l'uniforme continuité de f sur $\mathbb{R}$ on peut trouver un réel $\alpha > 0$ tel que $|f(x) - f(y)| \leq 1$ pour tout couple (x,y) de réels tels que $|x - y| \leq \alpha$. On en déduit alors par récurrence que :*

$$\forall n \in \mathbb{N},\ \forall x \in [n\alpha, (n+1)\alpha],\ |f(x) - f(0)| \leq n + 1$$

Le résultat est vrai pour $n = 0$ et en le supposant acquis pour $n \geq 0$, on a pour tout $x \in [(n+1)\alpha, (n+2)\alpha]$:

$$\begin{aligned} |f(x) - f(0)| &\leq |f(x) - f((n+1)\alpha)| + |f((n+1)\alpha) - f(0)| \\ &\leq 1 + (n+1) = n + 2 \end{aligned}$$

Si x est un réel positif ou nul, on peut alors trouver un entier naturel n tel que $x \in [n\alpha, (n+1)\alpha]$ et avec $n \leq \dfrac{x}{\alpha}$, on déduit que :

$$|f(x)| \leq |f(x) - f(0)| + |f(0)| \leq n + 1 + |f(0)| \leq \frac{x}{\alpha} + (1 + |f(0)|)$$

En raisonnant avec la fonction g définie par $g(x) = f(-x)$, on a un résultat analogue pour les réels négatifs.

Du fait qu'il n'est pas possible de trouver des réels a, b tels que $x^2 \leq ax + b$ pour tout réel positif, on déduit que la fonction $x \mapsto x^2$ n'est pas uniformément continue sur $\mathbb{R}$.

1.5.3 Continuité et compacité

Théorème 1.14 *Toute fonction définie sur un intervalle réel fermé borné $I = [a,b]$ à valeurs réelles et continue est bornée et atteint ses bornes.*

Démonstration. Soit $(y_n)_{n\in\mathbb{N}}$ une suite dans $f(I)$ avec $y_n = f(x_n)$ pour tout $n \in \mathbb{N}$. Le théorème de Bolzano-Weierstrass nous dit que de la suite $(x_n)_{n\in\mathbb{N}}$ dans le segment I on peut extraire une suite $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ qui converge vers un réel $x \in I$. Avec la continuité de f on a alors :

$$\lim_{n\to+\infty} y_{\varphi(n)} = \lim_{n\to+\infty} f\left(x_{\varphi(n)}\right) = f(x)$$

En conséquence $f(I)$ est compact. En particulier $f(I)$ est une partie non vide bornée de $\mathbb{R}$ et donc admet une borne inférieure et une borne supérieure. Notons :

$$m = \inf_{x\in I} f(x),\ M = \sup_{x\in I} f(x)$$

Il reste à montrer que m et M sont dans $f(I)$.

Par définition de la borne inférieure m, pour tout entier $n > 0$ on peut trouver x_n dans I tel que :

$$m \leq f(x_n) < m + \frac{1}{n}$$

De la suite $(x_n)_{n\in\mathbb{N}}$ ainsi définie dans le segment I on peut extraire une sous-suite $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ qui converge vers un réel $a \in I$. On a donc pour tout entier $n > 0$:

$$m \leq f\left(x_{\varphi(n)}\right) < m + \frac{1}{\varphi(n)}$$

avec $\lim\limits_{n\to+\infty} \varphi(n) = +\infty$. On a donc, avec la continuité de f :

$$f(a) = \lim_{n\to+\infty} f\left(x_{\varphi(n)}\right) = m$$

On procède de manière analogue pour la borne supérieure. ■

Un autre résultat important relatif aux fonctions continues sur un segment est le suivant.

Théorème 1.15 *Toute fonction continue sur un segment est uniformément continue.*

Démonstration. Soit f continue sur $I = [a, b]$.

Supposons f non uniformément continue sur I. Il existe alors un réel $\varepsilon > 0$ et des suites $(x_n)_{n\geq 1}$, $(y_n)_{n\geq 1}$ dans I telles que $|x_n - y_n| < \frac{1}{n}$ et $|f(x_n) - f(y_n)| > \varepsilon$ pour tout $n \geq 1$.

Avec la compacité de I, on peut extraire deux suites $\left(x_{\varphi(n)}\right)_{n\geq 1}$ et $\left(y_{\varphi(n)}\right)_{n\geq 1}$ qui convergent respectivement vers x et y dans I.

Mais avec $\left|x_{\varphi(n)} - y_{\varphi(n)}\right| < \frac{1}{\varphi(n)} \leq \frac{1}{n}$, on déduit que $|x - y| = \lim\limits_{n\to+\infty} \left|x_{\varphi(n)} - y_{\varphi(n)}\right| = 0$, soit $x = y$ et avec la continuité de f on a alors $\lim\limits_{n\to+\infty} \left|f\left(x_{\varphi(n)}\right) - f\left(y_{\varphi(n)}\right)\right| = 0$ en contradiction avec $\left|f\left(x_{\varphi(n)}\right) - f\left(y_{\varphi(n)}\right)\right| > \varepsilon$ pour tout $n \geq 1$. ■

Comme première application de ce résultat, on peut montrer que toute fonction f continue sur un segment $[a, b]$ est limite uniforme d'une suite de fonctions continues affines par morceaux.

De manière plus précise, soit f continue sur $[a, b]$. Pour tout entier $n \geq 1$ on définit une subdivision de $[a, b]$ en notant :

$$x_k = a + k\frac{b-a}{n} \quad (0 \leq k \leq n)$$

et à cette subdivision on associe la fonction f_n définie par :

$$f_n(x) = f(x_k) + \frac{x - x_k}{x_{k+1} - x_k}\left(f(x_{k+1}) - f(x_k)\right)$$

(f_n coïncide avec f aux x_k et est affine sur $[x_k, x_{k+1}]$). Cette fonction est affine par morceaux et continue sur $[a, b]$.

Lemme 1.3 *La suite $(f_n)_{n\geq 1}$ converge uniformément vers f sur $[a, b]$.*

Démonstration. La fonction f continue sur le compact $[a, b]$ y est uniformément continue, donc pour $\varepsilon > 0$ donné on peut trouver un réel $\eta > 0$ tel que si x, y dans $[a, b]$ sont tels que $|x - y| \leq \eta$ alors $|f(x) - f(y)| < \varepsilon$. Pour tout entier $n \geq \frac{b-a}{\eta}$ et tout entier k compris entre

0 et $n-1$ on a alors $x_{k+1}-x_k=\dfrac{b-a}{n}\leq\eta$. Sachant qu'un réel $x\in[a,b]$ est dans l'un des intervalles $[x_k,x_{k+1}]$, on obtient pour $n\geq\dfrac{b-a}{\eta}$:

$$\begin{aligned}|f(x)-f_n(x)| &= \left|f(x)-f(x_k)-\frac{x-x_k}{x_{k+1}-x_k}\left(f(x_{k+1})-f(x_k)\right)\right| \\ &\leq |f(x)-f(x_k)|+\frac{x-x_k}{x_{k+1}-x_k}|f(x_{k+1})-f(x_k)| \\ &\leq \varepsilon+\frac{x-x_k}{x_{k+1}-x_k}\varepsilon\leq 2\varepsilon\end{aligned}$$

ce qui prouve la convergence uniforme sur $[a,b]$ de $(f_n)_{n\geq 1}$ vers f. ■

Ce résultat peut être utilisé pour prouver, sans théorie de l'intégration, que toute fonction f continue sur un intervalle compact admet une primitive.

On vérifie tout d'abord qu'une fonction affine par morceaux et continue sur $[a,b]$ admet une primitive, ce qui n'est pas difficile.

Si, avec les notations qui précèdent, on désigne pour tout $n\geq 1$ par F_n la primitive de f_n nulle en a, on constate que la suite $(F_n')_{n\geq 1}$ converge uniformément sur $[a,b]$ vers f et que la suite $(F_n(a))_{n\geq 1}$ converge vers 0. On déduit alors que la suite $(F_n)_{n\geq 1}$ converge uniformément sur $[a,b]$ vers une fonction dérivable F et que $F'=f$, c'est-à-dire que F est une primitive de f sur $[a,b]$.

On peut alors définir l'intégrale d'une fonction f continue sur $[a,b]$ par :

$$\int_a^b f(x)\,dx=F(b)-F(a)$$

où F est une primitive de f sur cet intervalle.

Une deuxième application est relative aux fonctions périodiques.

Théorème 1.16 *Toute fonction continue sur $\mathbb{R}$ périodique de période $2T>0$ et à valeurs réelles est uniformément continue.*

Démonstration. Toute fonction f continue sur $\mathbb{R}$ périodique de période $2T$ est uniformément continue sur le compact $J=[-T-1,T+1]$. Donc pour tout $\varepsilon>0$, on peut trouver un réel $\eta\in\,]0,1[$ tel que :

$$(t,x)\in J^2,\ |t-x|\leq\eta\implies|f(t)-f(x)|\leq\varepsilon$$

Pour $x\in[-T,T]$ et $t\in\mathbb{R}$ tels que $|t-x|\leq\eta$ on a nécessairement $t\in[-T-1,T+1]$ ($\eta\in\,]0,1[$) et $|f(t)-f(x)|\leq\varepsilon$.

Pour $(t,x)\in\mathbb{R}^2$ tels que $|t-x|\leq\eta$ et $n\in\mathbb{Z}$ tel que $x-2nT\in[-T,T]$ ($n=E\left(\dfrac{x+T}{2T}\right)$) on a $|(t-2nT)-(x-2nT)|\leq\eta$ et :

$$|f(t)-f(x)|=|f(t-2nT)-f(x-2nT)|\leq\varepsilon$$

On a donc ainsi prouvé que f est uniformément continue sur $\mathbb{R}$. ■

Une troisième application importante est le résultat suivant de continuité d'une fonction définie par une intégrale sur un segment.

Théorème 1.17 *Soit f une fonction à valeurs réelles définie et continue sur $I \times [a,b]$, où I est un intervalle réel non réduit à un point et $a < b$ dans $\mathbb{R}$. La fonction φ définie sur I par :*

$$\forall x \in I,\ \varphi(x) = \int_a^b f(x,t)\,dt$$

est continue sur I.

Démonstration. On se fixe un réel x_0 dans I et un intervalle $[\alpha, \beta]$ contenu dans I, contenant x_0, avec $\alpha < \beta$.

La fonction f étant continue sur le compact $K = [\alpha, \beta] \times [a,b]$ y est uniformément continue, donc pour tout réel $\varepsilon > 0$ on peut trouver un réel $\eta > 0$ tel que $|f(x,t) - f(u,v)| < \varepsilon$ pour tous (x,t) et (u,v) dans K tels que $\|(x,t) - (u,v)\|_\infty < \eta$.

Pour x dans I tel que $|x - x_0| < \eta$, on a $\|(x,t) - (x_0,t)\|_\infty < \eta$ pour tout $t \in [a,b]$ et :

$$|\varphi(x) - \varphi(x_0)| \leq \int_a^b |f(x,t) - f(x_0,t)|\,dt < (b-a)\,\varepsilon$$

ce qui prouve la continuité de φ en x_0. ■

1.6 Dérivabilité en un point, dérivabilité sur I

Définition 1.6 *On dit que la fonction f est dérivable en $a \in I$ si la fonction ;*

$$\begin{array}{rccc} \tau_a : & I \setminus \{a\} & \to & \mathbb{C} \\ & x & \mapsto & \dfrac{f(x) - f(a)}{x-a} \end{array}$$

admet une limite finie en a.

Quand cette limite existe, elle est unique, on la note $f'(a)$ et on dit que c'est le nombre dérivé de f en a.

De manière équivalente, on peut dire que f est dérivable en a si, et seulement si, elle admet un développement limité d'ordre 1 en a :

$$f(x) = a_0 + a_1(x-a) + o(x-a)$$

et dans ce cas $\mathrm{a}_0 = f(a)$, $a_1 = f'(a)$.

De la définition du nombre dérivé on déduit facilement le résultat suivant.

Théorème 1.18 *Si f est dérivable en $a \in I$ elle est alors continue en ce point.*

La réciproque de ce résultat est fausse comme le montre l'exemple de la fonction $x \mapsto |x|$ au voisinage de 0.

Définition 1.7 *On dit que la fonction f est dérivable à gauche [resp. à droite] en $a \in I$ si la fonction τ_a admet une limite à gauche [resp. à droite] finie en a.*

Quand cette limite existe, elle est unique, on la note $f'_g(a)$ [resp. $f'_d(a)$] et on dit que c'est le nombre dérivé à gauche [resp. à droite] de f en a.

Là encore des définitions on déduit facilement les résultats suivants.

Théorème 1.19 *Si f est dérivable à gauche et à droite en $a \in \overset{\circ}{I}$ alors elle est continue en ce point.*

Théorème 1.20 *Si $a \in \overset{\circ}{I}$, alors f est dérivable en a si, et seulement si, elle dérivable à gauche et à droite en ce point avec $f'_g(a) = f'_d(a)$. Cette valeur commune est alors égale à $f'(a)$.*

Définition 1.8 *Si D est l'ensemble des points $x \in I$ où $f : I \to \mathbb{R}$ (ou $\mathbb{C}$) est dérivable, on définit alors la fonction dérivée de f sur D par $x \mapsto f'(x)$.*

Cet ensemble D peut être vide et dans ce cas la fonction dérivée n'est pas définie. C'est le cas par exemple pour la fonction caractéristique de $\mathbb{Q}$ qui est discontinue en tout point de $\mathbb{R}$ (exemple 1.4) et en conséquence ne peut être dérivable.

Une fonction peut très bien être continue et nulle part dérivable sur I. Pour construire une telle fonction, on désigne par φ la fonction 2-périodique sur $\mathbb{R}$ définie par :

$$\forall x \in [-1,1], \ \varphi(x) = |x|$$

Lemme 1.4 *Pour tous réels x, y on a :*

$$|\varphi(x) - \varphi(y)| \leq |x - y|$$

Démonstration. Avec $\varphi(x) \in [0,1]$ pour tout réel x, on déduit que pour x, y dans $\mathbb{R}$ tels que $|x - y| \geq 1$, on a $|\varphi(x) - \varphi(y)| \leq 1 \leq |x - y|$.

On suppose donc que $|x - y| < 1$ et comme x, y jouent des rôles symétriques, on peut même supposer que $0 < y - x < 1$. On distingue alors les cas suivants :

– si x, y sont dans $[-1,1]$, alors :

$$|\varphi(x) - \varphi(y)| = ||x| - |y|| \leq |x - y|$$

– si $x \in [-1,1]$ et $y \in [1,2]$, avec $0 < y - x < 1$ on a nécessairement $x \in [0,1]$ et $y - 2 \in [-1,0]$, ce qui donne :

$$\left\{ \begin{array}{l} \varphi(x) - \varphi(y) = \varphi(x) - \varphi(y-2) = x + y - 2 \leq y - x \\ \varphi(y) - \varphi(x) = \varphi(y-2) - \varphi(x) = 2 - y - x \leq y - x \end{array} \right.$$

soit $|\varphi(x) - \varphi(y)| \leq |x - y|$;

– dans le cas général, toujours avec $0 < y - x < 1$, en notant n la partie entière de $\dfrac{x+1}{2}$, on a :

$$\left\{ \begin{array}{l} -1 \leq x - 2n < 1 \\ -1 \leq x - 2n < y - 2n < 1 + x - 2n < 2 \end{array} \right.$$

et avec ce qui précède :

$$|\varphi(x) - \varphi(y)| = |\varphi(x - 2n) - \varphi(y - 2n)| \leq |x - y|$$

■

La proposition qui suit nous sera également utile.

Lemme 1.5 *Soit $x \in \mathbb{R}$. Si l'intervalle $\left[x, x + \dfrac{1}{2}\right[$ ne contient pas d'entiers, alors :*

$$\left|\varphi\left(x + \frac{1}{2}\right) - \varphi(x)\right| = \frac{1}{2}$$

Démonstration. Si n est la partie entière de $\dfrac{x+1}{2}$, alors $-1 < x-2n < 1$ (l'inégalité stricte à gauche est justifiée par $x \notin \mathbb{Z}$) et comme il n'y a pas d'entiers dans $\left[x-2n, x-2n+\dfrac{1}{2}\right[$, on a soit $-1 < x-2n < 0$ et $-\dfrac{1}{2} < x-2n+\dfrac{1}{2} \leq 0$, soit $x-2n > 0$ et $\dfrac{1}{2} < x-2n+\dfrac{1}{2} \leq 1$, ce qui donne :

$$\varphi\left(x+\frac{1}{2}\right)-\varphi(x)=\varphi\left(x-2n+\frac{1}{2}\right)-\varphi(x-2n)=-\frac{1}{2}$$

dans le premier cas et :

$$\varphi\left(x+\frac{1}{2}\right)-\varphi(x)=\varphi\left(x-2n+\frac{1}{2}\right)-\varphi(x-2n)=\frac{1}{2}$$

dans le second. ■

On définit la fonction de Van der Waerden par :

$$\forall x \in \mathbb{R},\ f(x)=\sum_{n=0}^{+\infty}\left(\frac{3}{4}\right)^{n}\varphi(4^{n}x)$$

Avec $0 \leq \left(\dfrac{3}{4}\right)^{n}\varphi(4^{n}x) \leq \left(\dfrac{3}{4}\right)^{n}$, on déduit que cette série de fonction est uniformément convergente et avec la continuité de φ que la somme f est continue sur $\mathbb{R}$.

Théorème 1.21 *La fonction f n'est dérivable en aucun point de $\mathbb{R}$.*

Démonstration. Soient $x \in \mathbb{R}$ et $m \in \mathbb{N}\setminus\{0\}$.

L'intervalle $\left[4^{m}x-\dfrac{1}{2}, 4^{m}x+\dfrac{1}{2}\right[$ qui est de longueur 1 contient exactement un entier et cet entier est dans l'un seulement des deux intervalles $\left[4^{m}x-\dfrac{1}{2}, 4^{m}x\right[$ ou $\left[4^{m}x, 4^{m}x+\dfrac{1}{2}\right[$. On note alors $\varepsilon_m=-1$ si le premier intervalle ne contient pas d'entier et $\varepsilon_m=1$ si c'est le second. En posant $h_m=\dfrac{\varepsilon_m}{2}\dfrac{1}{4^{m}}$, on a :

$$\frac{f(x+h_m)-f(x)}{h_m}=\sum_{n=0}^{+\infty}\left(\frac{3}{4}\right)^{n}\frac{\varphi(4^{n}(x+h_m))-\varphi(4^{n}x)}{h_m}$$

Pour $n > m$, $4^{n}h_m=2\varepsilon_m 4^{n-m-1}$ est un entier pair et $\varphi(4^{n}x+4^{n}h_m)-\varphi(4^{n}x)=0$.

Pour $n=m$, $4^{m}h_m=\dfrac{\varepsilon_m}{2}$ et le lemme précédent nous dit que :

$$|\varphi(4^{m}x+4^{m}h_m)-\varphi(4^{m}x)|=\frac{1}{2}$$

Enfin pour $0 \leq n < m$, avec le lemme 1.4 on a :

$$|\varphi(4^{n}(x+h_m))-\varphi(4^{n}x)| \leq 4^{n}h_m$$

Il en résulte que :

$$\left|\frac{f(x+h_m)-f(x)}{h_m}\right|=\left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^{n}\frac{\varphi(4^{n}(x+h_m))-\varphi(4^{n}x)}{h_m}\right|$$
$$\geq 3^{m}-\sum_{n=0}^{m-1}3^{n}=\frac{3^{m}+1}{2}$$

et $\lim\limits_{m\to+\infty}\left|\dfrac{f(x+h_m)-f(x)}{h_m}\right| = +\infty$ avec $\lim\limits_{m\to+\infty} h_m = 0$, ce qui implique que f ne peut être dérivable en x. ■

Un autre exemple de telle fonction est donné par la fonction de Weierstrass définie par :

$$\forall x \in \mathbb{R},\ f(x) = \sum_{n=0}^{+\infty} \frac{\cos(p^n x)}{2^n}$$

où $p \geq 6$ est un entier pair. On peut même montrer que cette fonction n'est monotone sur aucun intervalle non réduit à un point (voir [33]).

Si une fonction f est dérivable sur un intervalle I, sa dérivée f' n'est pas nécessairement continue comme le montre l'exemple de la fonction f définie par $f(x) = x^2 \sin\left(\dfrac{1}{x}\right)$ pour $x \neq 0$ dans $\mathbb{R}$ et $f(0) = 0$. On a $f'(x) = 2x\sin\left(\dfrac{1}{x}\right) - \cos\left(\dfrac{1}{x}\right)$ pour $x \neq 0$ et :

$$\left|\frac{f(x)}{x}\right| = \left|x\sin\left(\frac{1}{x}\right)\right| \leq |x| \underset{x\to 0}{\to} 0$$

donc $f'(0) = 0$. Mais f' n'est pas continue en 0 car la fonction $g : x \mapsto \cos\left(\dfrac{1}{x}\right)$ n'a pas de limite en 0 (ce qui se déduit de $g\left(\dfrac{1}{k\pi}\right) = (-1)^k$ pour tout entier $k \geq 1$).

Définition 1.9 *On dit qu'une fonction f est de classe $\mathcal{C}^1$ (ou continûment dérivable) sur I si elle est dérivable en tout point de I et si la fonction dérivée f' est continue sur cet intervalle.*

Exercice 1.17 *Soit $f : [0,1] \to \mathbb{R}$ continue, dérivable en 0 avec $f(0) = 0$. Montrer que :*

$$\lim_{n\to+\infty} \sum_{k=1}^{n} f\left(\frac{1}{n+k}\right) = f'(0)\ln(2)$$

Solution 1.17 *Soient $\varepsilon > 0$ et $\eta > 0$ tels que :*

$$0 < x < \eta \ \Rightarrow\ \left|\frac{f(x)}{x} - f'(0)\right| < \varepsilon$$

Pour tout entier $n > \dfrac{1}{\eta}$, on a $0 < \dfrac{1}{n+k} < \eta$ pour tout entier k compris entre 1 et n et donc :

$$\left|f\left(\frac{1}{n+k}\right) - \frac{1}{n+k}f'(0)\right| < \frac{1}{n+k}\varepsilon$$

ce qui entraîne :

$$\left|\sum_{k=1}^{n} f\left(\frac{1}{n+k}\right) - \left(\sum_{k=1}^{n}\frac{1}{n+k}\right)f'(0)\right| < \left(\sum_{k=1}^{n}\frac{1}{n+k}\right)\varepsilon < \varepsilon$$

En considérant que :

$$\lim_{n\to+\infty}\sum_{k=1}^{n}\frac{1}{n+k} = \lim_{n\to+\infty}\frac{1}{n}\sum_{k=1}^{n}\frac{1}{1+\frac{k}{n}} = \int_0^1 \frac{dx}{1+x} = \ln(2)$$

on aboutit au résultat.
Par exemple pour $f(x) = x^2$, on obtient :

$$\lim_{n \to +\infty} \sum_{k=1}^{n} \frac{1}{(n+k)^2} = 0$$

et pour $f(x) = \arctan(x)$:

$$\lim_{n \to +\infty} \sum_{k=1}^{n} \arctan\left(\frac{1}{n+k}\right) = \ln(2)$$

1.7 Opérations sur les fonctions dérivables

Pour toute fonction f et tout point a de I, on note $\tau_a(f)$ la fonction définie sur $I \setminus \{a\}$ par $\tau_a(f)(x) = \dfrac{f(x) - f(a)}{x - a}$.

Les résultats importants relatifs aux opérations algébriques sur les fonctions dérivables sont résumés avec le théorème qui suit.

Théorème 1.22 *Soient f, g deux fonctions de I dans $\mathbb{R}$ (ou $\mathbb{C}$) dérivables en $a \in I$.*

1. *Pour tous réels λ, μ la fonction $\lambda f + \mu g$ est dérivable en a avec :*
$$(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a)$$
2. *La fonction fg est dérivable en a avec :*
$$(fg)'(a) = f'(a)g(a) + f(a)g'(a)$$
(formule de Leibniz).
3. *Si $g(a) \neq 0$, alors la fonction g ne s'annule pas dans un voisinage de a, les fonctions $\dfrac{1}{g}$ et $\dfrac{f}{g}$ qui sont définies dans un tel voisinage sont dérivables en a avec :*
$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g^2(a)}, \ \left(\frac{f}{g}\right)'(a) = \frac{g(a)f'(a) - f(a)g'(a)}{g^2(a)}$$

Démonstration.

1. Résulte de :
$$\tau_a(\lambda f + \mu g) = \lambda \tau_a(f) + \mu \tau_a(g)$$
2. Résulte de :
$$\tau_a(fg)(x) = f(x)\tau_a(g) + g(a)\tau_a(f)$$
et de la continuité de a.
3. Si $g(a) \neq 0$, on a $g(x) \neq 0$ pour tout x dans un voisinage de a du fait de la continuité de g en ce point. Avec :
$$\tau_a\left(\frac{1}{g}\right)(x) = -\frac{1}{g(x)g(a)}\tau_a(g)$$
pour $x \neq a$, on déduit par passage à la limite quand x tend vers a que $\left(\dfrac{1}{g}\right)'(a) = -\dfrac{g'(a)}{g^2(a)}$.
La formule de Leibniz permet d'obtenir le deuxième point.

■

Pour ce qui est de la composition et du passage à l'inverse, on a les résultats suivants.

Théorème 1.23 *Soient I, J deux intervalles réels $f : I \to J$ une fonction dérivable en $a \in I$ et $g : J \to \mathbb{R}$ (ou $\mathbb{C}$) une fonction dérivable en $b = f(a)$. La fonction $g \circ f$ est définie sur I et dérivable en a avec :*

$$(g \circ f)'(a) = g'(f(a))\, f'(a)$$

Démonstration. On définit la fonction τ_b sur J par :

$$\tau_b(y) = \begin{cases} \dfrac{g(y) - g(b)}{y - b} \text{ si } y \neq b \\ g'(b) \text{ si } y = b \end{cases}$$

et pour $x \neq a$ dans I, on a :

$$\frac{g \circ f(x) - g \circ f(a)}{x - a} = \tau_b(f(x)) \frac{f(x) - f(a)}{x - a}$$

(pour $f(x) \neq f(a)$ c'est clair et pour $f(x) = f(a)$, les deux membres de cette égalité sont nuls). Faisant tendre x vers a on obtient le résultat du fait de la continuité de f en a, de celle de τ_b en b et de la définition de $f'(a)$. ■

Théorème 1.24 *Soit $f : I \to \mathbb{R}$ une fonction continue strictement monotone dérivable en a. La fonction réciproque f^{-1} est dérivable en $f(a)$ si, et seulement si, $f'(a) \neq 0$ et dans ce cas on a :*

$$\left(f^{-1}\right)'(f(a)) = \frac{1}{f'(a)}$$

Démonstration. On rappelle que si $f : I \to \mathbb{R}$ est continue strictement monotone, c'est alors un homéomorphisme de I sur $f(I)$ (théorème 2.10).

Si f^{-1} est dérivable en $f(a)$, on a alors :

$$1 = \left(f^{-1} \circ f\right)'(a) = \left(f^{-1}\right)'(f(a)) \cdot f'(a)$$

et nécessairement $f'(a) \neq 0$.

Supposons $f'(a) \neq 0$ et notons $b = f(a)$. Pour tout $y \neq b$ dans J on a $f^{-1}(y) \neq a$ et :

$$\frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \left(\frac{f(f^{-1}(y)) - f(a)}{f^{-1}(y) - a}\right)^{-1}$$

et avec la continuité de f^{-1}, on a :

$$\lim_{y \to b} \frac{f(f^{-1}(y)) - f(a)}{f^{-1}(y) - a} = f'(a)$$

ce qui entraîne :

$$\lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{1}{f'(a)}$$

■

De ce résultat, on déduit les dérivées des fonctions trigonométriques et hyperboliques inverses, à savoir :

– $\forall x \in \left]-1,1\right[,\ \arcsin'(x) = \dfrac{1}{\sqrt{1-x^2}}$;
– $\forall x \in \left]-1,1\right[,\ \arccos'(x) = -\dfrac{1}{\sqrt{1-x^2}}$;
– $\forall x \in \mathbb{R},\ \arctan'(x) = \dfrac{1}{1+x^2}$;
– $\forall x \in \mathbb{R},\ \operatorname{argsh}'(x) = \dfrac{1}{\sqrt{1+x^2}}$;
– $\forall x \in \left]1,+\infty\right[,\ \operatorname{argch}'(x) = \dfrac{1}{\sqrt{x^2-1}}$;
– $\forall x \in \left]-1,1\right[,\ \operatorname{argth}'(x) = \dfrac{1}{1-x^2}$.

La fonction $\arcsin + \arccos$ étant de dérivée nulle sur $\left]-1,1\right[$ est constante et la valeur de cette fonction en $x=0$, nous donne :

$$\forall x \in \left]-1,1\right[,\ \arcsin(x) + \arccos(x) = \frac{\pi}{2}$$

De même en remarquant que la dérivée de la fonction $x \mapsto \arctan(x) + \arctan\left(\dfrac{1}{x}\right)$ est nulle sur $\mathbb{R}^*$ et en prenant les valeurs en -1 et 1, on déduit que :

$$\forall x \in \mathbb{R}^*,\ \arctan(x) + \arctan\left(\frac{1}{x}\right) = \frac{x}{|x|}\frac{\pi}{2}$$

Exercice 1.18 *Étudier la dérivabilité de la fonction f définie sur $\mathbb{R}$ par :*

$$f(x) = \begin{cases} x^2\left|\cos\left(\dfrac{\pi}{x}\right)\right| & si\ x \neq 0 \\ 0\ si\ x = 0 \end{cases}$$

Solution 1.18 *Il est clair que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^* \setminus \left\{\dfrac{2}{2k+1} \mid k \in \mathbb{Z}\right\}$.*
Avec :

$$\left|\frac{f(x)}{x}\right| \leq |x| \underset{x\to 0}{\to} 0$$

on déduit que f est dérivable en 0 avec $f'(0) = 0$.
Soit $k = 2p$ un entier naturel pair et $x_k = \dfrac{2}{2k+1}$. Pour $\dfrac{\pi}{x} \in \left]k\pi, \dfrac{\pi}{2} + k\pi\right[$, on a $\cos\left(\dfrac{\pi}{x}\right) > 0$ et pour $\dfrac{\pi}{x} \in \left]\dfrac{\pi}{2} + k\pi, (k+1)\pi\right[$, $\cos\left(\dfrac{\pi}{x}\right) < 0$, de sorte que :
si $\dfrac{2}{2k+1} < x < \dfrac{1}{k}$ pour $k \neq 0$, ou $x > 2$ pour $k = 0$, alors :

$$\frac{f(x) - f(x_k)}{x - x_k} = \frac{x^2\cos\left(\dfrac{\pi}{x}\right)}{x - x_k} \underset{x\to x_k^+}{\to} \left(x^2\cos\left(\frac{\pi}{x}\right)\right)'_{|x=x_k} = \pi$$

et si $\dfrac{1}{k+1} < x < \dfrac{2}{2k+1}$, alors :

$$\frac{f(x) - f(x_k)}{x - x_k} = -\frac{x^2\cos\left(\dfrac{\pi}{x}\right)}{x - x_k} \underset{x\to x_k^+}{\to} -\left(x^2\cos\left(\frac{\pi}{x}\right)\right)'_{|x=x_k} = -\pi$$

La fonction f est donc dérivable à droite et à gauche en x_{2p} avec :

$$f'_g(x_{2p}) = -\pi, \ f'_d(x_{2p}) = \pi$$

On vérifie de même que f est dérivable à droite et à gauche en x_{2p+1}, pour tout entier naturel p, avec :

$$f'_g(x_{2p+1}) = -\pi, \ f'_d(x_{2p+1}) = \pi$$

Ces dérivées à droite et à gauche étant distinctes, on en déduit que f n'est pas dérivable en x_k pour tout $k \in \mathbb{N}$. Avec la parité de la fonction f, on déduit que ce résultat est encore valable pour tout $k \in \mathbb{Z}$.

1.8 Extremums et dérivation

Théorème 1.25 *Soient I un intervalle réel d'intérieur non vide, f une fonction à valeurs réelles définie sur I dérivable en un point a intérieur à I. Si f admet un extremum local en a alors $f'(a) = 0$.*

Démonstration. On suppose que la fonction f admet un maximum local en a.

Le point a étant intérieur à I, la fonction $\varphi : t \mapsto f(a+t)$ est définie sur un voisinage ouvert de 0 et en écrivant que :

$$\begin{cases} f'(a) = \lim\limits_{\substack{t \to 0 \\ t>0}} \dfrac{f(a+t) - f(a)}{t} \leq 0 \\ f'(a) = \lim\limits_{\substack{t \to 0 \\ t<0}} \dfrac{f(a+t) - f(a)}{t} \geq 0 \end{cases}$$

on déduit que $f'(a) = 0$. ■

Deux conséquences importantes de ce théorème sont le théorème de Rolle (théorème 3.4) et le théorème de Darboux (théorème 1.27).

La démonstration du théorème précédent contient plus précisément le résultat suivant.

Théorème 1.26 *Si I est un intervalle réel d'intérieur non vide, f une fonction à valeurs réelles définie sur I dérivable à gauche et à droite en un point a intérieur à I et qui admet un maximum [resp. minimum] local en ce point, alors $f'_g(a) \geq 0$ et $f'_d(a) \leq 0$ [resp. $f'_g(a) \leq 0$ et $f'_d(a) \geq 0$].*

La réciproque de ce théorème est fausse comme le montre l'exemple de la fonction $x \mapsto x^3$ au voisinage de 0.

1.9 Le théorème de Darboux

On a vu que si f est une fonction dérivable sur un intervalle I, alors sa dérivée f' n'est pas nécessairement continue et pourtant cette dérivée vérifie la propriété des valeurs intermédiaires.

Théorème 1.27 (Darboux) *Si f est une fonction à valeurs réelles définie et dérivable sur un intervalle I, alors sa fonction dérivée f' vérifie la propriété des valeurs intermédiaires.*

Démonstration. Soient $a < b$ dans I. Si $f'(a) = f'(b)$ il n'y a alors rien à montrer.

On suppose donc que $f'(a) < f'(b)$ et on se donne $\lambda \in \left]f'(a), f'(b)\right[$. On définit la fonction $\varphi : [a,b] \to \mathbb{R}$ par :

$$\forall x \in [a,b], \ \varphi(x) = f(x) - \lambda x$$

Cette fonction est continue sur le compact $[a,b]$, elle est donc minorée et atteint sa borne inférieure, c'est-à-dire qu'il existe $c \in [a,b]$ tel que $\varphi(c) = \inf_{x \in [a,b]} \varphi(x)$. Si $c = a$ [resp. $c = b$], alors $\dfrac{\varphi(x) - \varphi(a)}{x-a} \geq 0$ [resp. $\dfrac{\varphi(b) - \varphi(x)}{b-x} \leq 0$] pour tout $x \in \left]a,b\right[$ et en passant à limite quand x tend vers a [resp. vers b] par valeurs supérieures [resp. inférieures], on déduit que $\varphi'(a) \geq 0$ [resp. $\varphi'(b) \leq 0$] ce qui est en contradiction avec $\varphi'(a) = f'(a) - \lambda < 0 < \varphi'(b) = f'(b) - \lambda$. On a donc $c \in \left]a,b\right[$ et $\varphi'(c) = 0$, soit $f'(c) = \lambda$. ■

Le théorème de Darboux nous montre qu'une fonction peut très bien vérifier la propriété des valeurs intermédiaires sans être continue.

Par exemple la fonction f définie par $f(x) = x^2 \sin\left(\dfrac{1}{x}\right)$ prolongée par continuité en 0 est dérivable sur $\mathbb{R}$ de dérivée f' vérifiant la propriété des valeurs intermédiaires et non continue en 0.

Corollaire 1.1 *Il existe des fonctions définies sur un intervalle réel qui n'admettent pas de primitive.*

Démonstration. Si f admet une primitive F sur I, elle doit vérifier la propriété des valeurs intermédiaires. Une fonction ne vérifiant pas cette propriété, par exemple une fonction en escaliers, n'admet donc pas de primitive sur I. ■

En fait, on peut vérifier directement qu'une fonction en escalier non constante n'admet pas de primitives.

Du théorème de Darboux et du théorème 2.8, on déduit le résultat suivant.

Théorème 1.28 *Une fonction convexe et dérivable de I dans $\mathbb{R}$ est continûment dérivable.*

Démonstration. Si f est convexe et dérivable de I dans $\mathbb{R}$, sa dérivée f' est alors croissante. Cette dérivée vérifiant la propriété des valeurs intermédiaires, on déduit du théorème 2.8 qu'elle est continue. ■

2

Théorème des valeurs intermédiaires

2.1 Le théorème des valeurs intermédiaires

Le théorème des valeurs intermédiaires, pour les fonctions d'une variable réelle, peut être vu comme une conséquence du théorème de la borne supérieure sur $\mathbb{R}$ qui nous dit que toute partie non vide et majorée dans $\mathbb{R}$ admet une borne supérieure.

Théorème 2.1 *Soient I un intervalle réel non réduit à un point, f une fonction continue de I dans $\mathbb{R}$ et $a < b$ deux réels dans I tels que $f(a) f(b) < 0$. Dans ces conditions, il existe au moins un réel $\alpha \in \left]a, b\right[$ tel que $f(\alpha) = 0$.*

Démonstration. Supposons que $f(a) < 0 < f(b)$ (quitte à remplacer f par $-f$, on s'y ramène).

Soit :

$$\mathcal{A} = \{x \in [a, b] \mid f(x) \leq 0\}$$

$\mathcal{A}$ est une partie non vide ($a \in \mathcal{A}$) majorée (par b) de $\mathbb{R}$, elle admet donc une borne supérieure $\alpha \in [a, b]$.

La fonction f étant continue à droite en a et à gauche en b, on peut trouver un réel $\eta \in \left]0, \frac{b-a}{2}\right[$ tel que :

$$\begin{array}{l} \forall x \in [a, a+\eta],\ f(x) < 0 \\ \forall x \in [b-\eta, b],\ f(x) > 0 \end{array}$$

On a alors $a < a + \eta \leq \alpha \leq b - \eta < b$ et $\alpha \in \left]a, b\right[$. Il existe donc un entier n_0 strictement positif tel que $\left[\alpha - \frac{1}{n_0}, \alpha + \frac{1}{n_0}\right] \subset \left]a, b\right[$ et par définition de la borne supérieure α, pour tout entier $n \geq n_0$ il existe $x_n \in \mathcal{A}$ tels que :

$$\alpha - \frac{1}{n} < x_n \leq \alpha$$

En posant $y_n = \alpha + \frac{1}{n}$, on a $y_n \in [a, b] - \mathcal{A}$ et :

$$x_n \leq \alpha < y_n$$

On a alors $\alpha = \lim\limits_{n \to +\infty} x_n = \lim\limits_{n \to +\infty} y_n$ avec, pour tout entier $n \geq n_0$, $f(x_n) \leq 0$ ($x_n \in \mathcal{A}$) et $f(y_n) > 0$ ($y_n \in [a, b] - \mathcal{A}$). Avec la continuité de f en α, on en déduit que :

$$f(\alpha) = \lim_{n \to +\infty} f(x_n) \leq 0,\ f(\alpha) = \lim_{n \to +\infty} f(y_n) \geq 0$$

et $f(\alpha) = 0$. ■

Remarque 2.1 *Si, avec les hypothèses du théorème précédent, la fonction f est de plus strictement monotone, elle est alors injective et la solution α de l'équation $f(x) = 0$ dans l'intervalle $[a, b]$ est unique.*

Corollaire 2.1 *Soient I un intervalle réel non réduit à un point, f une fonction continue de I dans $\mathbb{R}$ et a, b deux réels dans I tels que $f(a) < f(b)$. Dans ces conditions, pour tout réel $\lambda \in \left]f(a), f(b)\right[$, il existe au moins un réel α strictement compris entre a et b tel que $f(\alpha) = \lambda$.*

Démonstration. Il suffit d'applique le théorème des valeurs intermédiaires à la fonction $g : x \mapsto f(x) - \lambda$. ■

Le théorème des valeurs intermédiaires est équivalent au résultat suivant.

Théorème 2.2 *Si I est un intervalle réel et f une fonction continue de I dans $\mathbb{R}$, alors $f(I)$ est un intervalle.*

Démonstration. Dire que $f(I)$ est un intervalle équivaut à dire que pour tous $y = f(a) < z = f(b)$ dans $f(I)$, le segment $[y, z]$ est contenu dans $f(I)$, donc tout $\lambda \in \left]f(a), f(b)\right[$ est dans $f(I)$ et il existe α strictement compris entre a et b tel que $f(\alpha) = \lambda$.

Réciproquement, pour f continue sur I, $y = f(a) < z = f(b)$ dans $f(I)$ avec $a \neq b$ dans I et tout $\lambda \in \left]y, z\right[$, le théorème des valeurs intermédiaires nous dit qu'il existe α compris entre a et b tel que $\lambda = f(\alpha) \in f(I)$, ce qui prouve que $f(I)$ est un intervalle. ■

Remarque 2.2 *Si la fonction f n'est pas continue en tout point de I, alors $f(I)$ n'est pas nécessairement un intervalle comme le montre l'exemple de la fonction f définie sur $I = [0, 2]$ par $f(x) = 1$ si $0 \leq x \leq 1$ et $f(x) = 2$ si $1 < x \leq 2$ (cette fonction est continue sur $I \setminus \{1\}$ avec $f(I) = \{1, 2\}$).*

On peut donner une autre démonstration du théorème des valeurs intermédiaires, en utilisant le théorème des segments emboîtés. Cette démonstration ayant l'avantage de fournir une méthode d'approximation d'une solution de l'équation $f(x) = 0$, à savoir la méthode de dichotomie (voir le paragraphe 8.2).

Les connexes (et convexes) de $\mathbb{R}$ étant les intervalles, le théorème des valeurs intermédiaires est un cas particulier du résultat qui suit.

On rappelle qu'une partie $\mathcal{C}$ d'un espace normé $(E, \|\cdot\|)$ est connexe si la condition $\mathcal{C} \subset \mathcal{O}_1 \cup \mathcal{O}_2$ avec $\mathcal{O}_1, \mathcal{O}_2$ ouverts de E tels que $\mathcal{C} \cap \mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$ entraîne $\mathcal{C} \cap \mathcal{O}_1 = \emptyset$ (et donc $\mathcal{C} \subset \mathcal{O}_2$) ou $\mathcal{C} \cap \mathcal{O}_2 = \emptyset$ (et donc $\mathcal{C} \subset \mathcal{O}_1$), c'est-à-dire que $\mathcal{C}$ ne peut s'écrire comme réunion disjointe de deux ouverts de $\mathcal{C}$ (pour la topologie induite).

Un ensemble connexe par arcs (i. e. pour tout $(a, b) \in \mathcal{C}$ il existe $\gamma : [0, 1] \to \mathcal{C}$ continue telle que $\gamma(0) = a$ et $\gamma(1) = b$) est connexe et dans le cas d'un ouvert non vide, la réciproque est vraie.

Théorème 2.3 *Si E, F sont deux espaces vectoriels normés et f une application continue de E dans F, alors pour tout connexe $\mathcal{C}$ de E l'image $f(\mathcal{C})$ est connexe dans F.*

Démonstration. Soit $\mathcal{C}$ une partie connexe de E et $\mathcal{C}' = f(\mathcal{C})$. Supposons qu'il existe deux ouverts de F, $\mathcal{O}_1, \mathcal{O}_2$, tels que $\mathcal{C}' \subset \mathcal{O}_1 \cup \mathcal{O}_2$ et $\mathcal{C}' \cap \mathcal{O}_1 \cap \mathcal{O}_2 = \emptyset$. On a alors :

$$\mathcal{C} \subset f^{-1}(\mathcal{C}') \subset f^{-1}(\mathcal{O}_1) \cup f^{-1}(\mathcal{O}_2)$$

avec $f^{-1}(\mathcal{O}_1), f^{-1}(\mathcal{O}_2)$ ouverts dans E puisque f est continue et $\mathcal{C} \cap f^{-1}(\mathcal{O}_1) \cap f^{-1}(\mathcal{O}_2) = \emptyset$, ce qui entraîne $\mathcal{C} \cap f^{-1}(\mathcal{O}_1) = \emptyset$ ou $\mathcal{C} \cap f^{-1}(\mathcal{O}_2) = \emptyset$ et donc $\mathcal{C}' \cap \mathcal{O}_1 = \emptyset$ ou $\mathcal{C}' \cap \mathcal{O}_2 = \emptyset$. L'ensemble $\mathcal{C}'$ est donc connexe dans F. ■

2.2 Quelques applications du théorème des valeurs intermédiaires

Le résultat qui suit est intéressant dans l'étude des points fixes.

Théorème 2.4 *Une fonction continue de $[a, b]$ dans $[a, b]$ admet au moins un point fixe.*

Démonstration. Voir le lemme 8.2. ■

En utilisant le fait qu'une fonction continue sur un compact est bornée et atteint ses bornes, on en déduit le résultat suivant.

Théorème 2.5 (Première formule de la moyenne) *Soient f, g deux fonctions à valeurs réelles définies et continues sur $I = [a, b]$, la fonction g étant de signe constant. Il existe un réel $c \in [a, b]$ tel que :*

$$\int_a^b f(x)\, g(x)\, dx = f(c) \int_a^b g(x)\, dx$$

Démonstration. La fonction f étant continue sur le compact I est bornée et atteint ses bornes, il existe donc deux réels α, β dans I tels que $m = \inf\limits_{x \in I} f(x) = f(\alpha)$ et $M = \sup\limits_{x \in I} f(x) = f(\beta)$. En supposant g à valeurs positives ou nulles, on a $mg(x) \leq f(x)\, g(x) \leq Mg(x)$ pour tout $x \in I$ et :

$$m \int_a^b g(x)\, dx \leq \int_a^b f(x)\, g(x)\, dx \leq M \int_a^b g(x)\, dx$$

Si $\displaystyle\int_a^b g(x)\, dx = 0$, alors $\displaystyle\int_a^b f(x)\, g(x)\, dx = 0$ et n'importe quel point c convient.

Si $\gamma = \displaystyle\int_a^b g(x)\, dx \neq 0$ alors $\gamma > 0$, $\delta = \dfrac{1}{\gamma} \displaystyle\int_a^b f(x)\, g(x)\, dx \in [f(\alpha), f(\beta)]$ et le théorème des valeurs intermédiaires nous dit qu'il existe un réel c compris entre α et β tel que $\delta = f(c)$, ce qui donne la formule de la moyenne. ■

On a aussi la version discrète du premier théorème de la moyenne donnée par l'exercice qui suit. Ce résultat peut être utilisé pour obtenir une estimation de l'erreur dans la méthode des rectangles à gauche.

Exercice 2.1 *Soit φ est une fonction continue sur $[a, b]$ et $(x_k)_{1 \leq k \leq n}$ une suite de points de $[a, b]$ avec $n \geq 2$. Montrer qu'il existe un réel c dans $[a, b]$ tel que :*

$$\sum_{k=1}^{n-1} \varphi(x_k) = n \cdot \varphi(c)$$

Solution 2.1 *En désignant par m [resp. M] la borne inférieure [resp. supérieure] de φ sur $[a, b]$, on a :*

$$m \leq \frac{1}{n} \sum_{k=0}^{n-1} \varphi(x_k) \leq M$$

et le théorème des valeurs intermédiaires nous dit qu'il existe $c \in [a, b]$ tel que $\dfrac{1}{n} \sum\limits_{k=1}^{n-1} \varphi(x_k) = \varphi(c)$.

En utilisant le théorème d'intégration par parties et le théorème des valeurs intermédiaires, on peut montrer la seconde formule de la moyenne qui suit.

Théorème 2.6 (Seconde formule de la moyenne) *Soient f une fonction à valeurs réelles positives de classe $\mathcal{C}^1$ et décroissante sur $I = [a, b]$ et g une fonction continue sur I. Il existe un réel $c \in I$ tel que :*

$$\int_a^b f(x)\, g(x)\, dx = f(a) \int_a^c g(x)\, dx$$

Démonstration. Soit G la primitive de g nulle en a. Une intégration par parties nous donne :

$$\int_a^b f(x)\, g(x)\, dx = f(b)\, G(b) - \int_a^b f'(x)\, G(x)\, dx$$

La fonction G étant continue sur le compact I est bornée et on peut noter $m = \inf\limits_{x \in I} G(x)$ et $M = \sup\limits_{x \in I} G(x)$. Comme f est à valeurs positives, on a :

$$mf(b) \leq f(b)\, G(b) \leq Mf(b)$$

Si de plus f est de classe $\mathcal{C}^1$ et décroissante, on a alors $f'(x) \leq 0$ pour tout $x \in I$ et :

$$Mf'(x) \leq f'(x)\, G(x) \leq mf'(x)$$

qui donne par intégration :

$$M(f(b) - f(a)) \leq \int_a^b f'(x)\, G(x)\, dx \leq m(f(b) - f(a))$$

Il en résulte que :

$$mf(b) - m(f(b) - f(a)) \leq \int_a^b f(x)\, g(x)\, dx \leq Mf(b) - M(f(b) - f(a))$$

soit :

$$mf(a) \leq \int_a^b f(x)\, g(x)\, dx \leq Mf(a).$$

Si $f(a) = 0$, on a alors $\displaystyle\int_a^b f(x)\, g(x)\, dx = 0$ et la formule est vérifiée pour tout $c \in I$, sinon on a $f(a) > 0$ et $m \leq \dfrac{1}{f(a)} \displaystyle\int_a^b f(x)\, g(x)\, dx \leq M$, ce qui entraîne l'existence de $c \in I$ tel que $\dfrac{1}{f(a)} \displaystyle\int_a^b f(x)\, g(x)\, dx = G(c)$ (théorème des valeurs intermédiaires pour la fonction continue G). ■

Remarque 2.3 *Cette seconde formule de la moyenne est encore valable pour f, g continues par morceaux sur I, la fonction f étant décroissante à valeurs positives (voir [33]).*

Exercice 2.2

1. *Soit f une application continue de $[0, 1]$ dans $\mathbb{R}$ telle que $f(0) = f(1)$.*
   Montrer que pour tout entier naturel non nul n il existe un réel x_n dans $\left[0, 1 - \dfrac{1}{n}\right]$ tel que $f(x_n) = f\left(x_n + \dfrac{1}{n}\right)$.

2. *Ce résultat est-il valable si on remplace* $\frac{1}{n}$ *par un réel* $\alpha \in \left]0,1\right[$ *qui n'est pas l'inverse d'un entier.*
3. *Montrer que si une voiture parcourt* 100 *km en une heure, il existe alors un intervalle de temps égal à une demi-heure pendant lequel la voiture a parcouru* $50\,km$.

Solution 2.2

1. *Pour tout* $n \geq 1$ *on désigne par* f_n *la fonction définie sur* $I_n = \left[0, 1 - \frac{1}{n}\right]$ *par* $f_n(x) = f\left(x + \frac{1}{n}\right) - f(x)$. *Si cette fonction ne s'annule jamais sur* I_n, *du fait de sa continuité, on déduit du théorème des valeurs intermédiaires qu'elle garde un signe constant sur cet intervalle. Supposons que* $f_n(x) > 0$ *pour tout* $x \in I_n$. *On a alors :*

$$f(1) - f(0) = \sum_{k=0}^{n-1} \left(f\left(\frac{k+1}{n}\right) - f\left(\frac{k}{n}\right) \right) = \sum_{k=0}^{n-1} f_n\left(\frac{k}{n}\right) > 0$$

ce qui contredit $f(0) = f(1)$. *La fonction* f_n *s'annule donc au moins une fois sur* I_n.

2. *Ce résultat n'est plus valable si on remplace* $\frac{1}{n}$ *par un réel* $\alpha \in \left]0,1\right[$ *qui n'est pas l'inverse d'un entier. En effet, si* f *est la fonction définie et continue sur* $[0,1]$ *par* $f(x) = \cos\left(\frac{2\pi x}{\alpha}\right) - \lambda x$, *où* λ *est choisi tel que* $f(0) = f(1)$ *(prendre* $\lambda = \cos\left(\frac{2\pi}{\alpha}\right) - 1$*), alors pour tout réel* $x \in [0, 1-\alpha]$, *on a :*

$$f(x+\alpha) - f(x) = -\lambda\alpha = \alpha\left(1 - \cos\left(\frac{2\pi}{\alpha}\right)\right) \neq 0$$

puisque $\frac{1}{\alpha}$ *n'est pas entier.*

3. *On considère la fonction* f *définie sur* $[0,1]$ *par* $f(x) = d(x) - 100x$ *où* $d(x)$ *est le nombre de kilomètres parcourus en* x *heures. Cette application est continue avec* $f(0) = f(1) = 0$. *Le résultat précédent, pour* $n = 2$, *nous dit qu'il existe* $x_2 \in \left[0, \frac{1}{2}\right]$ *tel que* $f(x_2) = f\left(x_2 + \frac{1}{2}\right)$, *soit* $d\left(x_2 + \frac{1}{2}\right) - d(x_2) = 50$ *et l'intervalle de temps* $\left[x_2, x_2 + \frac{1}{2}\right]$ *répond à la question.*

Exercice 2.3 *Montrer que si* $f : \mathbb{R} \to \mathbb{R}$ *est une fonction périodique continue et non constante de période* $T > 0$, *il existe alors un réel* a *tel que* $f\left(a + \frac{T}{2}\right) = f(a)$.

Solution 2.3 *La fonction* g *définie sur* $\mathbb{R}$ *par* $g(x) = f\left(x + \frac{T}{2}\right) - f(x)$ *est continue avec :*

$$\left\{ \begin{array}{l} g(0) = f\left(\frac{T}{2}\right) - f(0) \\ g\left(\frac{T}{2}\right) = f(T) - f\left(\frac{T}{2}\right) = f(0) - f\left(\frac{T}{2}\right) = -g(0) \end{array} \right.$$

on déduit alors du théorème des valeurs intermédiaires qu'il existe un réel a *compris entre* 0 *et* $\frac{T}{2}$ *tel que* $g(a) = 0$.

Exercice 2.4 *On désigne par Γ le cercle unité dans le plan complexe et par f une fonction continue de Γ dans $\mathbb{R}$. Montrer qu'il existe deux points diamétralement opposés dans Γ ayant même image par f.*

Solution 2.4 *La fonction g définie sur $\mathbb{R}$ par $g(x) = f(e^{ix})$ étant continue 2π-périodique et à valeurs réelles, il existe $\alpha \in [0, \pi]$ tel que $g(\alpha + \pi) = g(\alpha)$, soit $f\left(e^{i(\alpha+\pi)}\right) = f(e^{i\alpha})$, les points $e^{i\alpha}$ et $e^{i(\alpha+\pi)} = -e^{i\alpha}$ étant diamétralement opposés.*

Exercice 2.5 *On se place dans l'espace $E = \mathcal{C}^0([-1,1], \mathbb{R})$ muni de la norme $\|\cdot\|_\infty$ définie par :*

$$\forall f \in E,\ \|f\|_\infty = \sup_{x\in[-1,1]} |f(x)|$$

1. *Montrer que pour tout entier naturel $n \geq 1$, la fonction :*
$$T_n : x \mapsto \frac{1}{2^{n-1}} \cos(n \arccos(x))$$
est polynomiale de degré n et calculer $\|T_n\|_\infty$.
2. *Montrer que, pour toute fonction polynomiale de degré n, $P : x \mapsto x^n + \sum_{k=0}^{n-1} a_k x^k$, on a $\|P\|_\infty \geq \frac{1}{2^{n-1}}$.*

Solution 2.5 *On rappelle que la fonction* arccos *réalise un homéomorphisme de $[-1,1]$ sur $[0,\pi]$.*

1. *Pour $n =$ et $x \in [-1,1]$, on a $T_1(x) = \cos(\arccos(x)) = x$.*
Pour $n = 2$ et $x \in [-1,1]$, on a :
$$2T_2(x) = \cos(2\arccos(x)) = 2\cos^2(\arccos(x)) - 1 = 2x^2 - 1$$
soit $T_2(x) = x^2 - \frac{1}{2}$.
Supposons que pour $n \geq 2$ et k compris entre 1 et 2, T_k soit la restriction à $[-1,1]$ d'un polynôme unitaire de degré k.
Avec :
$$\cos((n+1)\theta) + \cos((n-1)\theta) = 2\cos(n\theta)\cos(\theta)$$
où on a noté $\theta = \arccos(x)$, on déduit que :
$$2^n T_{n+1}(x) + 2^{n-2} T_{n-1}(x) = 2^n x T_n(x)$$
soit :
$$T_{n+1}(x) = xT_n(x) - \frac{1}{4} T_{n-1}(x)$$
et T_{n+1} est la restriction à $[-1,1]$ d'un polynôme T_{n+1} unitaire de degré $n+1$.
Pour $n \geq 1$, on a $|T_n(x)| \leq \frac{1}{2^{n-1}}$ pour tout $x \in [-1,1]$ et la valeur $\frac{1}{2^{n-1}}$ est atteinte pour $x = 1$, donc $\|T_n\|_\infty = \sup_{x\in[-1,1]} |T_n(x)| = \frac{1}{2^{n-1}}$.

2. *Supposons qu'il existe un polynôme unitaire P de degré n tel que $\|P\|_\infty < \frac{1}{2^{n-1}}$. Pour tout k compris entre 0 et n, on a alors :*
$$-\frac{1}{2^{n-1}} < P\left(\cos\left(\frac{k\pi}{n}\right)\right) < \frac{1}{2^{n-1}}$$
et :
$$-\frac{1+(-1)^k}{2^{n-1}} < (P-T_n)\left(\cos\left(\frac{k\pi}{n}\right)\right) = P\left(\cos\left(\frac{k\pi}{n}\right)\right) - \frac{(-1)^k}{2^{n-1}} < \frac{1-(-1)^k}{2^{n-1}}$$
Donc :
$$(P-T_n)\left(\cos\left(\frac{k\pi}{n}\right)\right) < 0$$
pour k est pair et :
$$(P-T_n)\left(\cos\left(\frac{k\pi}{n}\right)\right) > 0$$
pour k est impair. On a donc :
$$(P-T_n)\left(\cos\left(\frac{k\pi}{n}\right)\right)(P-T_n)\left(\cos\left(\frac{(k+1)\pi}{n}\right)\right) < 0$$
pour $k = 0, \cdots, n$ et le théorème des valeurs intermédiaires nous dit que $P - T_n$ s'annule en au moins n points distincts de $[-1, 1]$. Les polynômes P et T_n étant unitaires de degré n, le polynôme $P - T_n$ est de degré au plus égal à $n - 1$ et donc nul puisqu'il a trop de racines distinctes.
On a donc $P = T_n$ et $T_n(1) = P(1) = \frac{1}{2^{n-1}}$, ce qui contredit $\|P\|_\infty < \frac{1}{2^{n-1}}$.
On a donc $\|P\|_\infty \geq \frac{1}{2^{n-1}}$ pour tout polynôme P unitaire de degré n.

2.3 Réciproque du théorème des valeurs intermédiaires

La réciproque du théorème des valeurs intermédiaires n'est pas vraie, c'est-à-dire qu'une fonction peut vérifier la propriété des valeurs intermédiaires sans être continue.

Plus précisément, on dit qu'une fonction f définie sur un intervalle réel I et à valeurs réelles vérifie la propriété des valeurs intermédiaires si pour tout intervalle J contenu dans I, $f(J)$ est un intervalle.

Le théorème de Darboux (théorème 1.27) nous dit qu'une fonction dérivée vérifie la propriété des valeurs intermédiaires et il existe des fonctions dérivables de dérivée non continue.

Un autre exemple est donné par le résultat qui suit.

Exercice 2.6 *Montrer que la fonction f définie sur $\mathbb{R}$ par $f(x) = \sin\left(\frac{1}{x}\right)$ si $x \neq 0$ et $f(0) = 0$ vérifie la propriété des valeurs intermédiaires sans être continue.*

Solution 2.6 *En considérant la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_n = \frac{1}{n\pi + \frac{\pi}{2}}$, on a $\lim\limits_{n\to+\infty} x_n = 0$ et la suite $(f(x_n))_{n\in\mathbb{N}} = ((-1)^n)_{n\in\mathbb{N}}$ n'a pas de limite, ce qui prouve que f n'est pas continue en 0.*
Si J est un intervalle réel ne contenant pas 0, alors f est continue sur J et $f(J)$ est un

intervalle.
Si J est un intervalle contenant 0 non réduit à un point (sinon $J = f(J) = \{0\}$), on considère les suites $(x_n)_{n\geq n_0}$, $(y_n)_{n\geq n_0}$, définies par $x_n = \dfrac{1}{2n\pi + \frac{\pi}{2}}$, $y_n = \dfrac{1}{(2n+1)\pi + \frac{\pi}{2}}$, où n_0 est un entier assez grand pour que ces suites soient à valeurs dans J, et on a $f(x_n) = 1$, $f(y_n) = -1$ pour tout $n \geq n_0$ et :

$$[-1,1] \supset f(J) \supset f([y_n, x_n]) \supset [-1,1]$$

c'est-à-dire que $f(J) = [-1,1]$.
En définitive, la fonction f vérifie la propriété des valeurs intermédiaires.

Ce qu'il manque à une fonction vérifiant la propriété des valeurs intermédiaires sur un intervalle compact $[a,b]$ pour être continue est donné par le résultat suivant.

Théorème 2.7 *Si f est une fonction à valeurs réelles définie sur un intervalle I vérifiant la propriété des valeurs intermédiaires alors f est continue si, et seulement si, pour tout réel y, l'ensemble $f^{-1}\{y\}$ est fermé dans I.*

Démonstration. Si f est continue, on sait qu'elle vérifie la propriété des valeurs intermédiaires et l'image réciproque du fermé $\{y\}$ par f est fermé.

Réciproquement supposons que $f : I \to \mathbb{R}$ vérifie la propriété des valeurs intermédiaires et que $f^{-1}\{y\}$ est fermé dans I pour tout $y \in \mathbb{R}$. On se donne $x_0 \in I$ et $\varepsilon > 0$. Les ensembles $F_1 = f^{-1}\{f(x_0) - \varepsilon\}$ et $F_2 = f^{-1}\{f(x_0) + \varepsilon\}$ sont des fermés (éventuellement vides) de I qui ne contiennent pas x_0, donc x_0 est dans l'ouvert $\mathcal{O} = (I \setminus F_1) \cap (I \setminus F_2)$ et il existe $\eta > 0$ tel que $J = I \cap]x_0 - \eta, x_0 + \eta[\subset \mathcal{O}$. Pour tout $x \in J$ on a $f(x) \neq f(x_0) \pm \varepsilon$ et $f(J)$ est un intervalle qui contient $f(x_0)$, nécessairement $f(J) \subset]f(x_0) - \varepsilon, f(x_0) + \varepsilon[$, c'est-à-dire que $|f(x) - f(x_0)| < \varepsilon$ pour tout $x \in I \cap]x_0 - \eta, x_0 + \eta[$. On a donc ainsi montré que f est continue en tout point de I. ■

Remarque 2.4 *On peut affaiblir les hypothèses dans le théorème précédent en remplaçant l'hypothèse, $f^{-1}\{y\}$ est fermé dans I pour tout réel y, par $f^{-1}\{y\}$ est fermé dans I pour tout rationnel y (voir [11], exercice 4.1).*

Dans le cas des fonctions monotones, on a le résultat suivant, où $I =]a,b[$ est un intervalle ouvert avec $-\infty \leq a < b \leq +\infty$.

Théorème 2.8 *Si f est une fonction monotone de I dans $\mathbb{R}$ telle que $f(I)$ soit un intervalle, alors elle est continue sur I.*

Démonstration. On suppose que f est croissante.

Il s'agit de montrer que pour tout $x \in I$ on a $\lim\limits_{t \to x} f(t) = f(x)$. On sait déjà que f admet une limite à gauche et à droite en x avec $f(x^-) \leq f(x) \leq f(x^+)$. Il s'agit donc de montrer que $f(x^-) = f(x^+) = f(x)$.

Supposons que $f(x^-) < f(x)$, pour $\varepsilon > 0$ donné, on peut alors trouver $x_0 \in]a,x[$ tel que :

$$f\left(x^-\right) - \varepsilon < f(x_0) \leq f\left(x^-\right) < f(x).$$

Mais si de plus $f(I)$ est un intervalle alors tout $\lambda \in]f(x^-), f(x)[$ étant dans $]f(x_0), f(x)[$ s'écrit $\lambda = f(x_1)$ avec $x_1 \in]x_0, x[$ et on a alors $\lambda = f(x_1) \leq f(x^-)$ en contradiction avec $\lambda > f(x^-)$. On a donc $f(x^-) = f(x)$.

On montre de manière analogue que $f(x) = f(x^+)$.

Les limites à droite et à gauche en x sont donc égales à $f(x)$, ce qui prouve la continuité de f en x. ■

2.4 Fonctions réciproques

Si f est une application injective de I dans $\mathbb{R}$, elle définit alors une bijection de I sur $f(I)$ et on peut définir sa fonction réciproque notée f^{-1} par :

$$\left(y \in f(I) \text{ et } x = f^{-1}(y)\right) \Leftrightarrow (x \in I \text{ et } y = f(x))$$

Un cas particulièrement intéressant est celui des fonctions strictement monotones.

Théorème 2.9 *Si f est une application strictement monotone de I dans $\mathbb{R}$, elle réalise alors une bijection de I sur $f(I)$ d'inverse strictement monotone et de même sens de variation que f.*

Démonstration. En remplaçant éventuellement f par $-f$, on peut supposer que f est strictement croissante. Si $x \neq y$ dans I on a $x < y$ ou $y < x$ (l'ordre de $\mathbb{R}$ est total) ce qui entraîne $f(x) < f(y)$ ou $f(y) < f(x)$, soit $f(x) \neq f(y)$ dans tous les cas. La fonction f est donc injective et elle réalise une bijection de I sur $f(I)$.
Il est facile de vérifier que la fonction réciproque f^{-1} est également strictement croissante. ■

Sans hypothèse de continuité pour f, l'image $f(I)$ n'est pas nécessairement un intervalle comme le montre l'exemple de la fonction f définie sur $[0, 2]$ par $f(x) = x$ pour $0 \leq x \leq 1$ et $f(x) = x + 1$ pour $1 < x \leq 2$.

En réalité si f est strictement monotone et $f(I)$ est un intervalle, alors la fonction f est nécessairement continue ainsi que la fonction f^{-1}.

Le théorème qui suit est à la base des définitions de fonctions réciproques classiques comme les fonctions racines n-ème, la fonction exponentielle, les fonctions trigonométriques et hyperboliques inverses.

Théorème 2.10 *Si f est une application continue et strictement monotone de I dans $\mathbb{R}$, alors $f(I)$ est un intervalle de même nature que I et f est une bijection de I sur $f(I)$ d'inverse f^{-1} continue strictement monotone de même sens de variation que f.*

Démonstration. Si f est strictement monotone, elle est alors injective. Si de plus elle est continue alors $J = f(I)$ est un intervalle (théorème des valeurs intermédiaires). La fonction réciproque f^{-1} est alors strictement monotone de J sur $I = f^{-1}(J)$ qui est un intervalle, elle est donc continue.
Il nous reste à montrer que $J = f(I)$ est de même nature que I. En notant $-\infty \leq a < b \leq +\infty$ les extrémités de I et en supposant f strictement croissante, J est un intervalle d'extrémités :

$$\alpha = \inf(f(I)) = \lim_{\substack{x \to a \\ x > a}} f(x), \ \beta = \sup(f(I)) = \lim_{\substack{x \to b \\ x < b}} f(x)$$

Si $a \in I$ alors $\alpha = f(a) \in J$ et de même pour b.

Si $a \notin I$ alors $\alpha \notin J$ (il suffit de raisonner avec f^{-1} qui a les mêmes propriétés que f) et $\alpha < f(x)$ pour tout $x \in J$. De même pour b.

Les intervalles I et J sont donc de même nature. ■

Remarque 2.5 *Sans hypothèse de stricte monotonie, l'intervalle $f(I)$ n'est pas nécessairement de même nature que I comme le montre l'exemple de la fonction f définie par $f(x) = |x|$ sur l'intervalle $I =]-1, 1[$ ($f(I) = [0, 1[$).*

Une application continue, bijective de I sur J et d'inverse continu est appelée homéomorphisme.

Ce résultat nous permet de définir la fonction exponentielle comme l'inverse de la fonction logarithme définie comme la primitive sur $\mathbb{R}^{+,*}$ nulle en 1 de la fonction $x \mapsto \frac{1}{x}$.

De $\ln'(x) > 0$ sur $\mathbb{R}^{+,*}$ on déduit que ln est strictement croissante sur $\mathbb{R}^{+,*}$, puis avec $\ln(2) > \ln(1) = 0$, $\ln(2^n) = n\ln(2)$ que cette fonction n'est pas bornée et $\lim_{x\to+\infty} \ln(x) = +\infty$.

Enfin avec $\ln\left(\frac{1}{x}\right) = -\ln(x)$, on déduit que $\lim_{x\to 0} \ln(x) = -\infty$.

La fonction ln est donc continue strictement croissante de $\mathbb{R}^{+,*}$ sur $\mathbb{R}$, c'est donc un homéomorphisme de $\mathbb{R}^{+,*}$ sur $\mathbb{R}$ et sa fonction réciproque est la fonction exponentielle $x \mapsto e^x$. Cette fonction est donc définie par :

$$(x \in \mathbb{R} \text{ et } y = e^x) \Leftrightarrow \big(y \in \mathbb{R}^{+,*} \text{ et } x = \ln(y)\big)$$

Pour tout entier naturel non nul n, avec la continuité et la stricte croissance de la fonction $x \mapsto x^n$ de $\mathbb{R}^+$ sur $\mathbb{R}^+$, on aboutit à la définition de la fonction racine n-ème :

$$\big(x \in \mathbb{R}^+ \text{ et } y = \sqrt[n]{x}\big) \Leftrightarrow \big(y \in \mathbb{R}^+ \text{ et } x = y^n\big)$$

Avec les mêmes arguments, on définit les fonctions trigonométriques inverses :

$$\left\{\begin{array}{l} (x \in [-1,1] \text{ et } y = \arcsin(x)) \Leftrightarrow \left(y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \text{ et } x = \sin(y)\right) \\ (x \in [-1,1] \text{ et } y = \arccos(x)) \Leftrightarrow (y \in [0,\pi] \text{ et } x = \cos(y)) \\ (x \in \mathbb{R} \text{ et } y = \arctan(x)) \Leftrightarrow \left(y \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\text{ et } x = \tan(y)\right) \end{array}\right.$$

et les fonctions hyperboliques inverses :

$$\left\{\begin{array}{l} (x \in \mathbb{R} \text{ et } y = \operatorname{argsh}(x)) \Leftrightarrow (y \in \mathbb{R} \text{ et } x = \operatorname{sh}(y)) \\ (x \in [1,+\infty[\text{ et } y = \operatorname{argch}(x)) \Leftrightarrow (y \in \mathbb{R}^+ \text{ et } x = \operatorname{ch}(y)) \\ (x \in \mathbb{R} \text{ et } y = \operatorname{argth}(x)) \Leftrightarrow (y \in \mathbb{R} \text{ et } x = \operatorname{th}(y)) \end{array}\right.$$

On a vu qu'une fonction strictement monotone est injective, mais en général la réciproque est fausse comme le montre l'exemple de la fonction f définie sur $[0,2]$ par $f(x) = x$ pour $0 \leq x \leq 1$ et $f(x) = \frac{1-x}{2}$ pour $1 < x \leq 2$ (faire un graphique). Mais dans le cas où f est continue, il y a équivalence. De manière précise, on a le résultat suivant.

Théorème 2.11 *Soit f une fonction continue de I dans $\mathbb{R}$. Cette fonction f est injective si, et seulement si, elle est strictement monotone.*

Démonstration. Si f est strictement monotone, on a déjà vu qu'elle est injective (qu'elle soit continue ou non).

Pour la réciproque, on propose trois démonstrations.

Supposons f continue et injective de I dans $\mathbb{R}$.

Première démonstration. S'il existe $x < y < z$ dans I tels que $f(x) < f(z) < f(y)$ alors tout réel u dans $]f(z), f(y)[\subset]f(x), f(y)[$ va s'écrire $u = f(c) = f(d)$ avec $c \in]y,z[$ et $d \in]x,y[$, donc $c \neq d$, ce qui contredit l'injectivité de f. La fonction f est donc strictement monotone.

Deuxième démonstration. Pour $a < b$ dans I, on a $f(a) \neq f(b)$. Supposons que $f(a) < f(b)$. Nous allons alors montrer que f est strictement croissante. Pour $x < y$ dans I, l'application :

$$\varphi : t \mapsto \varphi(t) = f(ta + (1-t)x) - f(tb + (1-t)y)$$

est continue sur $[0,1]$ et ne s'annule jamais sur cet intervalle (comme f est injective, $\varphi(t) = 0$ équivaut à $t(b-a) + (1-t)(y-x) = 0$ encore équivalent à $t(b-a) = (1-t)(y-x) = 0$ qui est impossible). Cette fonction garde donc un signe constant sur $[0,1]$ et en particulier $\varphi(0) = f(x) - f(y)$ est de même signe que $\varphi(1) = f(a) - f(b)$, c'est-à-dire que $f(x) < f(y)$.

Troisième démonstration. On utilise la fonction g définie par $g(x,y) = f(x) - f(y)$ sur la partie Δ de $\mathbb{R}^2$ définie par :

$$\Delta = \left\{(x,y) \in I^2 \mid x < y\right\}.$$

Cette fonction est continue sur Δ qui est convexe donc connexe, ce qui implique que $g(\Delta)$ est connexe dans $\mathbb{R}^*$ (f est injective), c'est donc un intervalle de $\mathbb{R}^*$ et on a soit $g(\Delta) \subset \mathbb{R}^{+,*}$ et f est strictement croissante, soit $g(\Delta) \subset \mathbb{R}^{-,*}$ et f est strictement décroissante. ■

3

Théorèmes de Rolle et des accroissements finis

3.1 Le théorème de Rolle sur un espace vectoriel normé

Pour ce paragraphe, on se donne un espace vectoriel normé $(E, \|\cdot\|)$.

Le théorème de Rolle est basé sur les deux théorèmes suivants, relatifs à des problèmes d'extremum.

Théorème 3.1 *Si K est un compact de E et f une fonction continue K dans $\mathbb{R}$, elle est alors bornée et atteint ses bornes, c'est-à-dire qu'il existe α et β dans K tels que :*

$$f(\alpha) = \inf_{x \in K} f(x), \ f(\beta) = \sup_{x \in K} f(x)$$

Théorème 3.2 *Si $\mathcal{O}$ est un ouvert non vide de E et f une fonction de $\mathcal{O}$ dans $\mathbb{R}$ différentiable en un point $\alpha \in \mathcal{O}$ et admettant un extremum local en α alors $df(\alpha) = 0$.*

Théorème 3.3 (Rolle) *Soient K un compact de E d'intérieur non vide, f une fonction continue de K dans $\mathbb{R}$ différentiable sur l'intérieur de K et constante sur la frontière $K \setminus \overset{\circ}{K}$ de K. Il existe alors un élément $c \in \overset{\circ}{K}$ tel que $df(c) = 0$.*

Démonstration. Si f est constante, sa différentielle est alors nulle.

On suppose donc f non constante.

La fonction f étant continue sur le compact K est bornée et atteint ses bornes, c'est-à-dire qu'il existe α, β dans K tels que $f(\alpha) = \inf_{x \in K} f(x)$ et $f(\beta) = \sup_{x \in K} f(x)$. Si α, β sont dans $\mathrm{Fr}(K) = K \setminus \overset{\circ}{K}$, on a alors $f(\alpha) \leq f(x) \leq f(\beta) = f(\alpha)$ pour tout $x \in K$ et f est constante contrairement à l'hypothèse de départ, on a donc $\alpha \in \overset{\circ}{K}$ ou $\beta \in \overset{\circ}{K}$, ce qui entraîne $df(\alpha) = 0$ ou $df(\beta) = 0$. ■

La classique version « réelle » de ce théorème est la suivante.

Théorème 3.4 (Rolle) *Si f est une fonction à valeurs réelles définie sur un intervalle compact $[a, b]$ non réduit à un point, continue sur cet intervalle et dérivable sur l'intervalle ouvert $]a, b[$ avec $f(a) = f(b)$, il existe alors un point $c \in]a, b[$ tel que $f'(c) = 0$.*

Remarque 3.1 *Il n'y a pas unicité d'un point c tel que $f'(c) = 0$ (figure 3.1).*

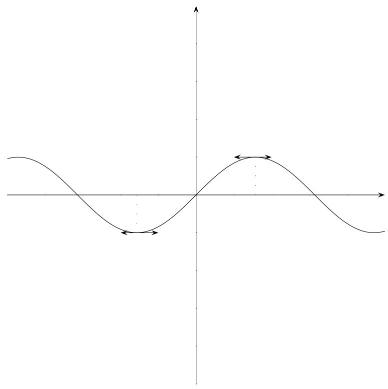

Figure 3.1 – $y = \sin(x)$

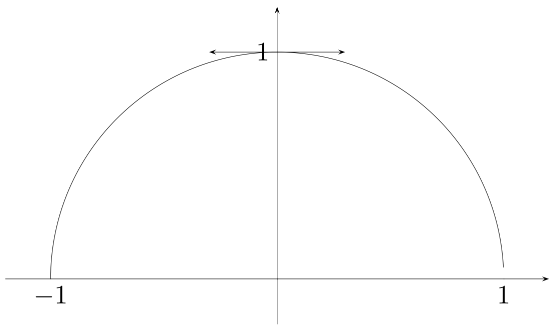

Figure 3.2 – $y = \sqrt{1 - x^2}$

Remarque 3.2 *La fonction $x \mapsto \sqrt{1 - x^2}$ sur $[-1, 1]$ nous donne un exemple de situation où f n'est pas dérivable au bord (figure 3.2).*

Remarque 3.3 *Le théorème n'est plus vrai si f n'est pas continue au bord comme le montre l'exemple de la fonction f définie par $f(x) = x$ sur $]0, 1]$ et $f(0) = 1$ (figure 3.3).*

Remarque 3.4 *Le théorème n'est plus vrai si f n'est pas dérivable sur $]a, b[$ tout entier comme le montre l'exemple de la fonction f définie par $f(x) = |x|$ sur $[-1, 1]$ (figure 3.4).*

Corollaire 3.1 *Si f est une fonction dérivable sur un intervalle réel I et à valeurs réelles telle que f' admette exactement $p \geq 0$ racines réelles distinctes, alors f a au plus $p+1$ racines réelles distinctes.*

Démonstration. Si f a $p+2$ racines réelles $\lambda_1 < \lambda_2 < \cdots < \lambda_{p+2}$, le théorème de Rolle nous assure l'existence d'au moins une racine réelle sur chaque intervalle $]\lambda_k, \lambda_{k+1}[$ pour k compris entre 1 et $p + 1$, ce qui donne au moins $p + 1$ racines distinctes pour f'. ■

Le théorème de Rolle pour les fonctions d'une variable réelle est encore valable sur une demi-droite fermée. Précisément on a le résultat suivant.

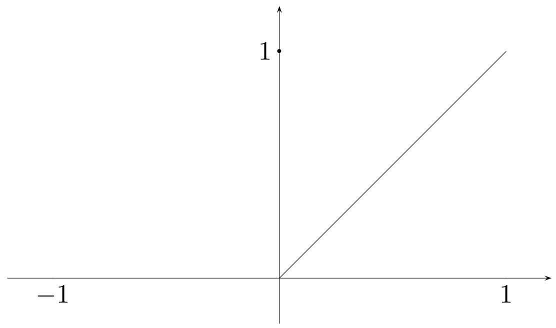

FIGURE 3.3 –

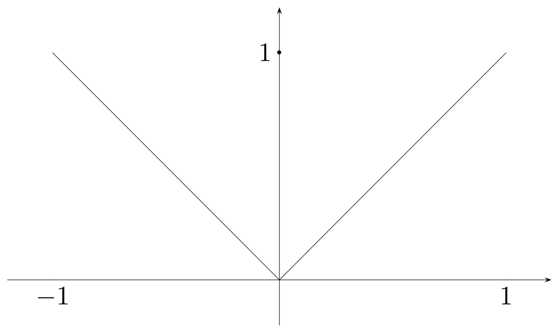

FIGURE 3.4 – $y = |x|$

Théorème 3.5 (Rolle) *Si $f : [a, +\infty[\to \mathbb{R}$ est continue sur cet intervalle et dérivable sur l'intervalle ouvert $]a, +\infty[$ avec $\lim\limits_{x \to +\infty} f(x) = f(a)$, il existe alors un point $c \in]a, +\infty[$ tel que $f'(c) = 0$.*

Démonstration. Le changement de variable $t = e^{-x}$ nous ramène à un intervalle compact. On définit donc la fonction g sur $[0, e^{-a}]$ par :

$$g(t) = \begin{cases} f(-\ln(t)) \text{ si } t \in]0, e^{-a}] \\ f(a) \text{ si } t = 0 \end{cases}$$

Cette fonction est continue sur $]0, e^{-a}]$ comme composée de fonctions continues et avec $\lim\limits_{t \to 0} g(t) = \lim\limits_{x \to +\infty} f(x) = f(a)$, on déduit qu'elle est continue en a. Elle est dérivable sur $]0, e^{-a}[$ avec $g'(t) = -\dfrac{f'(-\ln(t))}{t}$.

Enfin avec $g(0) = g(e^{-a}) = f(a)$, on peut utiliser le théorème de Rolle sur $[0, e^{-a}]$ pour dire qu'il existe $d \in]0, e^{-a}[$ tel que $g'(d) = 0$ et $c = -\ln(d) \in]a, +\infty[$ est tel que $f'(c) = 0$.

On peut aussi utiliser la fonction g définie sur $\left[0, \dfrac{\pi}{2}\right]$ par :

$$g(t) = \begin{cases} f(a + \tan(t)) \text{ si } t \in \left[0, \dfrac{\pi}{2}\right[\\ f(a) \text{ si } t = \dfrac{\pi}{2} \end{cases}$$

Cette fonction est continue sur $\left[0, \frac{\pi}{2}\right[$ comme composée de fonctions continues et avec $\lim_{t \to \frac{\pi}{2}} g(t) = \lim_{x \to +\infty} f(x) = f(a)$, on déduit qu'elle est continue en $\frac{\pi}{2}$. Elle est dérivable sur $\left]0, \frac{\pi}{2}\right[$ avec $g'(t) = (1 + \tan^2(t)) f'(a + \tan(t))$. Enfin avec $g(0) = g\left(\frac{\pi}{2}\right) = f(a)$, on peut utiliser le théorème de Rolle sur $\left[0, \frac{\pi}{2}\right]$ pour dire qu'il existe $d \in \left]0, \frac{\pi}{2}\right[$ tel que $g'(d) = 0$ et $c = a + \tan(d) \in \left]a, +\infty\right[$ est tel que $f'(c) = 0$. ■

On a également le résultat suivant pour les fonctions définies sur $\mathbb{R}$.

Théorème 3.6 (Rolle) *Si $f : \mathbb{R} \to \mathbb{R}$ est dérivable avec $\lim_{x \to -\infty} f(x) = \lim_{x \to +\infty} f(x)$, il existe alors un réel c tel que $f'(c) = 0$.*

Démonstration. Le changement de variable $t = \arctan(x)$ nous ramène à un intervalle compact.

On définit la fonction g sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ par :

$$g(t) = \begin{cases} f(\tan(t)) \text{ si } t \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\\ \ell = \lim_{x \to \pm\infty} f(x) \text{ si } t = \pm\frac{\pi}{2} \end{cases}$$

Cette fonction est continue sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ comme composée de fonctions continues et avec $\lim_{t \to \pm\frac{\pi}{2}} g(t) = \lim_{x \to \pm\infty} f(x) = \ell$, on déduit qu'elle est continue en $\pm\frac{\pi}{2}$. Elle est dérivable sur $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ avec $g'(t) = (1 + \tan^2(t)) f'(\tan(t))$. Le théorème de Rolle sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ nous dit alors qu'il existe $d \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ tel que $g'(d) = 0$ et $c = \tan(d)$ est tel que $f'(c) = 0$. ■

La version itérée suivante du théorème de Rolle est souvent utile (voir le problème de l'interpolation de Lagrange et l'étude des polynômes orthogonaux).

Théorème 3.7 (Rolle) *Si f est une fonction à valeurs réelles de classe $\mathcal{C}^m$ sur un intervalle réel I, où m est un entier naturel, qui s'annule en $m+1$ points de I distincts, il existe alors un point c dans I tel que $f^{(m)}(c) = 0$.*

Démonstration. Si $m = 0$ le résultat est évident. On suppose donc que m est non nul.

Si a, b sont deux racines distinctes de f, le théorème de Rolle nous dit alors qu'entre ces deux racines il existe une racine de f'. On en déduit que la fonction f' admet m racines distinctes dans I. Une récurrence finie nous permet alors de montrer que la dérivée d'ordre m, $f^{(m)}$ admet au moins une racine dans I. ■

3.2 Quelques applications du théorème de Rolle

Le théorème de Rolle est important pour ses nombreuses applications.

3.2.1 Quelques exercices classiques

Exercice 3.1 *Soient f, g deux fonctions à valeurs réelles non nulles, continues sur $[a, b]$, dérivables sur $]a, b[$, avec $f(a) g(b) = f(b) g(a)$. Montrer qu'il existe un réel $c \in]a, b[$ tel que $\frac{f'(c)}{f(c)} = \frac{g'(c)}{g(c)}$.*

Solution 3.1 *La fonction h définie sur $[a,b]$ par $h(x) = \dfrac{f(x)}{g(x)}$ est continue sur $[a,b]$, dérivable sur $]a,b[$ avec :*

$$\begin{cases} h'(x) = \dfrac{g(x) f'(x) - f(x) g'(x)}{g(x)^2} \\ h(a) = h(b) \end{cases}$$

Le théorème de Rolle nous dit alors qu'il existe un réel $c \in]a,b[$ tel que $h'(c) = 0$, ce qui équivaut à $\dfrac{f'(c)}{f(c)} = \dfrac{g'(c)}{g(c)}$.
Pour g constante égale à 1, on retrouve le théorème de Rolle classique.

Exercice 3.2 *Montrer que si $f : [0,1] \to \mathbb{R}$ est telle que $\displaystyle\int_0^1 f(t)\,dt = \frac{1}{2}$, alors f a un point fixe dans $]0,1[$.*

Solution 3.2 *La fonction g définie sur $[0,1]$ par $g(x) = \displaystyle\int_0^x f(t)\,dt - \frac{x^2}{2}$ est continue sur $[0,1]$, dérivable sur $]0,1[$ avec $g(0) = g(1) = 0$. Le théorème de Rolle nous dit alors qu'il existe $c \in]0,1[$ tel que $g'(c) = 0$, ce qui signifie $f(c) = c$.*

3.2.2 Sur les racines de polynômes réels

Théorème 3.8 *Si P est un polynôme réel de degré $n \geq 2$ scindé sur $\mathbb{R}$ alors il en est de même de son polynôme dérivé.*
Précisément si $\lambda_1 < \lambda_2 < \cdots < \lambda_p$ sont les racines réelles distinctes de P avec $p \geq 2$, la racine λ_j étant de multiplicité $m_j \geq 1$ ($\sum\limits_{j=1}^{p} m_j = n$), alors le polynôme dérivé P' admet les réels λ_j pour racines de multiplicités respectives $m_j - 1$, pour $1 \leq j \leq p$ (une multiplicité nulle signifie que λ_j n'est pas racine de P') et des racines simples $\mu_j \in]\lambda_j, \lambda_{j+1}[$ pour $1 \leq j \leq p-1$.

Démonstration. Pour $j \in \{1, \cdots, p\}$ tel que $m_j \geq 2$, λ_j est racine d'ordre $m_j - 1$ du polynôme P'. Ce qui donne $\sum\limits_{j=1}^{p} (m_j - 1) = n - p$ racines réelles pour P'. D'autre part, le théorème de Rolle nous dit que pour tout j dans $\{1, \cdots, p-1\}$ il existe $\mu_j \in]\lambda_j, \lambda_{j+1}[$ tel que $P'(\mu_j) = 0$, ce qui donne $p-1$ racines réelles supplémentaires et distinctes pour P'. On a donc un total de $n-1$ racines réelles pour P' et les μ_j sont nécessairement simples. ■

On peut remarquer que toutes les racines de P' sont dans l'intervalle $[\lambda_1, \lambda_p]$.

De manière plus générale, si P est un polynôme non constant à coefficients complexes, alors les racines du polynôme dérivé P' sont dans l'enveloppe convexe de l'ensemble des racines de P (théorème de Lucas).

Exercice 3.3 *Soient $n \geq 2$, a, b réels et $P(x) = x^n + ax + b$. Montrer que si n est pair alors P a 0, 1 ou 2 racines réelles et si n est impair alors P a 1, 2 ou 3 racines réelles.*

Solution 3.3 *Supposons n pair. On a alors $P''(x) = n(n-1)x^{n-2} > 0$ pour tout réel non nul x et P' est strictement croissante sur $\mathbb{R}$ de degré impair, elle s'annule donc une fois (théorème des valeurs intermédiaires) et une seule (P' est injective). Avec le théorème de Rolle on déduit alors que P s'annule au plus 2 fois.*
Supposons n impair. Alors P' est strictement décroissante sur $]-\infty, 0[$, strictement croissante sur $]0, +\infty[$, avec $P'(0) = a$. Il en résulte que P' a 2 racines réelles $-\rho$ et $\rho > 0$ si $a < 0$, 0 pour unique racine réelle si $a = 0$ et pas de racine réelle si $a > 0$. Avec la théorème de Rolle, on déduit alors que P a au plus 3 racines réelles.

On sait qu'un polynôme réel de degré n a au plus n racines réelles sur un intervalle I. Plus généralement, on a le résultat suivant.

Exercice 3.4 *Montrer que pour tout entier naturel n et toutes suites de réels $(a_k)_{0\leq k\leq n}$ et $(\lambda_k)_{0\leq k\leq n}$, les a_k étant non tous nuls et les λ_k deux à deux distincts, la fonction f_n définie par :*

$$f_n(x) = \sum_{k=0}^{n} a_k x^{\lambda_k}$$

a au plus n racines réelles distinctes dans $\mathbb{R}^{+,*}$. On dit que la famille de fonctions $\left(x^{\lambda_k}\right)_{0\leq k\leq n}$ est un système de Tchebychev (ou système de Haar ou encore système unisolvent) dans $\mathcal{C}^0\left(\mathbb{R}^{+,*}\right)$.

Solution 3.4 *On procède par récurrence sur $n \geq 0$.*

Pour $n = 0$, $f_0(x) = a_0 x^{\lambda_0}$ n'a pas de racine dans $\mathbb{R}^{+,*}$ puisque a_0 est non nul.

Supposons le résultat acquis au rang $n \geq 0$. Si la fonction $f_{n+1} = \sum\limits_{k=0}^{n+1} a_k x^{\lambda_k}$ a plus de $n+1$ racines distinctes dans $\mathbb{R}^{+,*}$, il en est alors de même de la fonction :

$$g_{n+1}(x) = x^{-\lambda_j} f_{n+1}(x) = \sum_{k=0}^{n+1} a_k x^{\lambda_k - \lambda_j}$$

où j compris entre 0 et $n+1$ est choisi tel que $a_j \neq 0$. Le théorème de Rolle nous dit alors que la fonction dérivée :

$$g'_{n+1}(x) = \sum_{k=0}^{n+1} (\lambda_k - \lambda_j) a_k x^{\lambda_k - \lambda_j - 1} = \sum_{\substack{k=0 \\ k\neq j}}^{n+1} (\lambda_k - \lambda_j) a_k x^{\lambda_k - \lambda_j - 1}$$

a plus de n racines distinctes dans $\mathbb{R}^{+,*}$ et en conséquence tous les $(\lambda_k - \lambda_j) a_k$ pour $k \neq j$ sont nuls (hypothèse de récurrence), ce qui entraîne $f_{n+1}(x) = a_j x^{\lambda_j}$, mais cette fonction ne s'annule jamais sur $\mathbb{R}^{+,*}$. On aboutit donc à une impossibilité.

Exercice 3.5 *Montrer que pour tout entier n, on a :*

$$\forall x \in \mathbb{R},\ \arctan^{(n+1)}(x) = \frac{P_n(x)}{(1+x^2)^{n+1}}$$

où P_n est un polynôme de degré n avec n racines réelles distinctes.

Solution 3.5 *Pour $n = 0$, on a* $\arctan'(x) = \dfrac{P_0(x)}{1+x^2}$ *avec $P_0(x) = 1$ sans racine réelle.*

En supposant le résultat acquis au rang n, la fonction f_n définie par $f_n(x) = \dfrac{P_n(x)}{(1+x^2)^{n+1}}$ est nulle en $\pm\infty$ et en n points distincts, on déduit alors des théorèmes de Rolle (classique sur un compact et généralisé sur un intervalle fermé de longueur infinie) que sa dérivée s'annule en $n+1$ points distincts, cette dérivée s'écrivant $\dfrac{P_{n+1}(x)}{(1+x^2)^{n+2}}$, où :

$$P_{n+1}(x) = \left(1+x^2\right) P'_n(x) - 2(n+1) x P_n(x)$$

est un polynôme de degré égal à $n+1$ (il a $n+1$ racines, ou alors on peut calculer son coefficient dominant). D'où le résultat.

En fait, en utilisant une décomposition en éléments simples dans $\mathbb{C}(X)$, on peut montrer que ces racines sont les $x_k = -\operatorname{cotan}\left(\dfrac{k\pi}{n+1}\right)$ avec k compris entre 1 et n.

3.2.3 Racines des polynômes de Legendre, de Laguerre et d'Hermite

Pour tout $n \in \mathbb{N}$, on note $\pi_{2n}(x) = (x^2-1)^n$ et $L_n = \pi_{2n}^{(n)}$.
Les polynômes L_n sont les polynômes de Legendre sur $[-1,1]$.
Pour $n = 0$, on a $L_0 = 1$.
Pour $n \geq 1$ le polynôme :

$$\pi_{2n}(x) = \sum_{k=0}^{n} (-1)^{n-k} C_n^k x^{2k}$$

est de degré $2n$ et sa dérivée d'ordre n :

$$L_n(x) = \sum_{\frac{n}{2} \leq k \leq n} (-1)^{n-k} C_n^k \frac{(2k)!}{(2k-n)!} x^{2k-n}$$

est un polynôme de degré n, de la parité de n.

Théorème 3.9 *Pour $n \geq 1$, le polynôme L_n admet n racines réelles distinctes dans l'intervalle $]-1,1[$.*

Démonstration. Pour $n = 1$, $L_1(x) = 2x$ s'annule en 0.

Pour $n \geq 2$, on vérifie par récurrence sur $k \in \{0, \cdots, n-1\}$, que le polynôme $\pi_{2n}^{(k)}$ s'annule en -1, 1 et en k points distincts de $]-1,1[$.

Le polynôme π_{2n} admettant -1 et 1 comme racines d'ordre n, le résultat est vrai pour $k = 0$.

Supposons le acquis pour $k-1 \in \{0, \cdots, n-2\}$ $(n \geq 2)$. La fonction $\pi_{2n}^{(k-1)}$ est nulle en $-1 < t_1 < \cdots < t_{k-1} < 1$ et avec le théorème de Rolle on déduit que sa dérivée $\pi_{2n}^{(k)}$ s'annule en k points distincts de $]-1,1[$. D'autre part, -1 et 1 étant racines d'ordre n de π_{2n}, elles sont aussi racines d'ordre $n-k > 0$ de $\pi_{2n}^{(k)}$.

En appliquant le théorème de Rolle à la fonction $\pi_{2n}^{(n-1)}$ qui est nulle en $n+1$ points distincts $-1 < t_1 < \cdots < t_{n-1} < 1$, on déduit que $L_n = \pi_{2n}^{(n)}$ s'annule en n points distincts de $]-1,1[$. ■

Soit $\alpha > -1$. Pour tout entier naturel n, on définit le polynôme $L_{\alpha,n}$ par $(x^{n+\alpha}e^{-x})^{(n)} = L_{\alpha,n}(x)\, x^\alpha e^{-x}$. Les polynômes $L_{\alpha,n}$ sont les polynômes de Laguerre sur $]0,+\infty[$. Il est facile de vérifier que $L_{\alpha,n}$ est un polynôme de degré n.

En effet, on a $L_{\alpha,0}(x) = 1$, $L_{\alpha,1}(x) = 1 + \alpha - x$. Supposons, pour $n \geq 1$, que $L_{\alpha,n}$ est polynomiale de degré n pour tout réel $\alpha > -1$. Avec :

$$\begin{aligned}
\left(x^{n+1+\alpha}e^{-x}\right)^{(n+1)} &= (n+1+\alpha)\left(x^{n+\alpha}e^{-x}\right)^{(n)} - \left(x^{n+(1+\alpha)}e^{-x}\right)^{(n)} \\
&= (n+1+\alpha)\, L_{\alpha,n}(x)\, x^\alpha e^{-x} - L_{\alpha+1,n}(x)\, x^{\alpha+1} e^{-x} \\
&= ((n+1+\alpha)\, L_{\alpha,n}(x) - x L_{\alpha+1,n}(x))\, x^\alpha e^{-x}
\end{aligned}$$

on déduit que $L_{\alpha,n+1}(x) = (n+1+\alpha)\, L_{\alpha,n}(x) - x L_{\alpha+1,n}(x)$ est polynomiale de degré $n+1$.

Théorème 3.10 *Pour tout réel $\alpha > -1$ et tout entier $n \geq 1$, le polynôme $L_{\alpha,n}$ admet n racines réelles distinctes dans $]0,+\infty[$.*

Démonstration. Pour $\alpha > -1$ et $n = 1$, on a $L_{\alpha,1}(x) = 1+\alpha-x$ de degré 1 nul en $1+\alpha > 0$. En supposant le résultat acquis au rang n pour tout $\alpha > -1$, la fonction f_n définie sur $[0,+\infty[$ par $f_n(x) = (x^{n+\alpha+1}e^{-x})^{(n)} = L_{\alpha+1,n}(x)\, x^{\alpha+1} e^{-x}$ est nulle en 0, $+\infty$ et en n points distincts de $]0,+\infty[$, on déduit alors des théorèmes de Rolle (classique sur un compact et généralisé sur $[0,+\infty[$) que sa dérivée $f_n'(x) = L_{\alpha,n+1}(x)\, x^\alpha e^{-x}$ s'annule en $n+1$ points distincts de $]0,+\infty[$. D'où le résultat au rang $n+1$. ■

Pour tout $n \in \mathbb{N}$, on définit le polynôme H_n par $\left(e^{-x^2}\right)^{(n)} = H_n(x)\, e^{-x^2}$. Les polynômes H_n sont les polynômes d'Hermite sur $\mathbb{R}$. Il est facile de vérifier que H_n est un polynôme de degré n.

Théorème 3.11 *Pour $n \geq 1$, le polynôme H_n admet n racines réelles distinctes.*

Démonstration. Pour $n = 1$, on a $H_1(x) = -2x$ de degré 1 nul en 0. En supposant le résultat acquis au rang n, la fonction f_n définie par $f_n(x) = H_n(x)\, e^{-x^2}$ est nulle en $\pm\infty$ et en n points distincts, on déduit alors des théorèmes de Rolle (classique sur un compact et généralisé sur un intervalle fermé de longueur infinie) que sa dérivée $f_n'(x) = H_{n+1}(x)\, e^{-x^2}$ s'annule en $n+1$ points distincts. D'où le résultat au rang $n+1$. ■

En fait ces résultat sont vrais pour toute famille de polynômes orthogonaux et on peut les démontrer en utilisant uniquement les propriétés d'orthogonalité et le théorème des valeurs intermédiaires (voir [38]).

3.2.4 Majoration de l'erreur dans l'interpolation de Lagrange

Soient $I = [a, b]$ un intervalle réel fermé borné avec $a < b$, n un entier naturel non nul et $(x_i)_{0 \leq i \leq n}$ une suite de réels deux à deux distincts dans I. À toute fonction f définie sur I et à valeurs réelles on associe le polynôme d'interpolation de Lagrange $L_n(f)$ défini par :

$$\begin{cases} L_n(f) \in \mathbb{R}_n[x] \\ L_n(f)(x_i) = f(x_i) \ \ (0 \leq i \leq n) \end{cases}$$

Un tel polynôme est uniquement déterminé par f. On peut l'écrire sous la forme :

$$L_n(f) = \sum_{i=0}^{n} f(x_i)\, L_{n,i}$$

avec :

$$L_{n,i}(x) = \prod_{j=0,\, j\neq i}^{n} \frac{x - x_j}{x_i - x_j} \ \ (0 \leq i \leq n)$$

Dans le cas où la fonction f est de classe $\mathcal{C}^{n+1}$ sur I, on peut donner une expression de l'erreur d'interpolation $f - L_n(f)$ en tout point de l'intervalle I. Précisément on a le résultat suivant où, pour $n \geq 1$, π_{n+1} est la fonction polynomiale définie par :

$$\pi_{n+1}(x) = \prod_{i=0}^{n} (x - x_i)$$

Théorème 3.12 *Soit f une fonction de classe $\mathcal{C}^{n+1}$ sur l'intervalle I. Pour tout x dans I il existe un point c_x appartenant à I tel que :*

$$f(x) - L_n(f)(x) = \frac{1}{(n+1)!}\pi_{n+1}(x)\, f^{(n+1)}(c_x)$$

Démonstration. Si x est l'un des points x_i, on a alors $f(x) - L_n(f)(x) = \pi_{n+1}(x) = 0$ et tout point $c_x \in I$ convient.

On se donne donc un point x dans $I \setminus \{x_0, \cdots, x_n\}$. On désigne par P_x le polynôme d'interpolation de Lagrange associé à la fonction f et aux points $x_0, \cdots, x_n, x$. Ce polynôme est défini par :

$$\begin{cases} P_x \in \mathbb{R}_{n+1}[t] \\ P_x(x_i) = f(x_i) \ \ (0 \leq i \leq n) \\ P_x(x) = f(x) \end{cases}$$

On vérifie facilement que :

$$P_x = L_n(f) + \frac{f(x) - L_n(f)(x)}{\pi_{n+1}(x)} \pi_{n+1}$$

La fonction $g_x = f - P_x$ est alors de classe $\mathcal{C}^{n+1}$ sur l'intervalle I, nulle en $n+2$ points distincts (x et les x_i), le théorème de Rolle itéré nous dit alors qu'il existe un point $c_x \in I$ tel que $g_x^{(n+1)}(c_x) = 0$, ce qui compte tenu de :

$$P_x^{(n+1)} = \frac{f(x) - L_n(f)(x)}{\pi_{n+1}(x)} (n+1)!$$

s'écrit :

$$f^{(n+1)}(c_x) - \frac{f(x) - L_n(f)(x)}{\pi_{n+1}(x)} (n+1)! = 0$$

ou encore :

$$f(x) - L_n(f)(x) = \frac{1}{(n+1)!} \pi_{n+1}(x) f^{(n+1)}(c_x)$$

■

Une démonstration analogue nous permet d'obtenir une majoration de l'erreur dans l'interpolation d'Hermite (voir [38]).

3.2.5 Convexité

Voir le paragraphe 5.3.

3.2.6 Le théorème de Darboux

On peut donner une démonstration du théorème de Darboux qui utilise le théorème 2.11 et le théorème de Rolle.

Théorème 3.13 (Darboux) *Si f est une fonction à valeurs réelles définie et dérivable sur un intervalle I, alors sa fonction dérivée f' vérifie la propriété des valeurs intermédiaires.*

Démonstration. Soient $a < b$ dans I. Si $f'(a) = f'(b)$ il n'y a alors rien à montrer.

On suppose donc que $f'(a) < f'(b)$ et on se donne $\lambda \in]f'(a), f'(b)[$. On définit la fonction $\varphi : [a, b] \to \mathbb{R}$ par :

$$\forall x \in [a, b], \ \varphi(x) = f(x) - \lambda x.$$

Cette fonction est dérivable sur $[a, b]$ avec $\varphi'(a) < 0 < \varphi'(b)$ et en conséquence elle ne peut être monotone sur I (une fonction monotone dérivable sur un intervalle a une dérivée de signe constant). Le théorème 2.11 nous dit alors que φ n'est pas injective, c'est-à-dire qu'il existe $x < y$ dans I tels que $\varphi(x) = \varphi(y)$ et le théorème de Rolle nous dit qu'il existe $c \in]x, y[$ tel que $\varphi'(c) = 0$, ce qui équivaut à $f'(c) = \lambda$. ■

3.3 Théorème et inégalité des accroissements finis

Le théorème de Rolle pour les fonctions d'une variable réelle est équivalent au théorème des accroissements finis qui suit où $[a,b]$ est un intervalle non réduit à un point.

Théorème 3.14 (Accroissements finis) *Si f est une fonction définie sur $[a,b]$, à valeurs réelles, continue sur $[a,b]$ et dérivable sur $]a,b[$, il existe alors un point c dans $]a,b[$ tel que $f(b)-f(a)=f'(c)(b-a)$.*

Démonstration. On se ramène aux hypothèses du théorème de Rolle en considérant la fonction g définie sur $[a,b]$ par :

$$g(x)=f(x)-f(a)-\lambda(x-a)$$

où la constante réelle λ est telle que $g(b)=g(a)$ (soit $\lambda=\dfrac{f(b)-f(a)}{b-a}$).

Le théorème de Rolle appliqué à la fonction g nous assure de l'existence d'un réel $c\in\,]a,b[$ tel que $g'(c)=0$, ce qui équivaut à $f(b)-f(a)=f'(c)(b-a)$. ■

Cette formule est encore valable en permutant les rôles de a et b. L'hypothèse $a<b$ n'est donc pas essentielle.

On dispose aussi de la version suivante, un peu plus générale, du théorème des accroissements finis.

Théorème 3.15 (généralisé des accroissements finis) *Si f,g sont deux fonctions définies sur $[a,b]$, à valeurs réelles, continues sur $[a,b]$ et dérivables sur $]a,b[$, il existe alors un point c dans $]a,b[$ tel que $(f(b)-f(a))g'(c)=(g(b)-g(a))f'(c)$.*

Démonstration. On se ramène aux hypothèses du théorème de Rolle en considérant la fonction h définie sur $[a,b]$ par :

$$h(x)=\lambda g(x)-\mu f(x)$$

où les constantes réelles λ,μ sont choisies telles que $h(a)=h(b)$, soit $\lambda g(a)-\mu f(a)=\lambda g(b)-\mu f(b)$, ou encore $\lambda(g(b)-g(a))=\mu(f(b)-f(a))$. On peut prendre $\lambda=f(b)-f(a)$ et $\mu=g(b)-g(a)$.

Le théorème de Rolle appliqué à la fonction h nous assure de l'existence d'un réel $c\in\,]a,b[$ tel que $h'(c)$, ce qui équivaut à $(f(b)-f(a))g'(c)=(g(b)-g(a))f'(c)$. ■

Prenant $g(x)=x$, on retrouve le théorème classique des accroissements finis.

L'exercice qui suit nous permet de retrouver les deux théorèmes précédents.

Exercice 3.6 *Soient f,g,h trois fonctions à valeurs réelles, continues sur $[a,b]$, dérivables sur $]a,b[$ et φ la fonction définie sur $[a,b]$ par :*

$$\varphi(x)=\det\begin{pmatrix} f(x) & g(x) & h(x)\\ f(a) & g(a) & h(a)\\ f(b) & g(b) & h(b)\end{pmatrix}$$

Montrer qu'il existe un réel $c\in\,]a,b[$ tel que $\varphi'(c)=0$.
Quels résultats obtient-on pour $h(x)=1$?

Solution 3.6 *La fonction φ est continue sur $[a,b]$ dérivable sur $]a,b[$ et avec le caractère 3-linéaire alterné du déterminant, on a :*

$$\left\{ \begin{array}{l} \varphi'(x) = \det \begin{pmatrix} f'(x) & g'(x) & h'(x) \\ f(a) & g(a) & h(a) \\ f(b) & g(b) & h(b) \end{pmatrix} \\ \varphi(a) = \varphi(b) = 0 \end{array} \right.$$

Le théorème de Rolle nous dit alors qu'il existe un réel $c \in]a,b[$ tel que $\varphi'(c) = 0$.
Pour $h = 1$, on obtient :

$$\varphi'(c) = \det \begin{pmatrix} f'(c) & g'(c) & 0 \\ f(a) & g(a) & 1 \\ f(b) & g(b) & 1 \end{pmatrix} = f'(c)(g(a) - g(b)) - g'(c)(f(a) - f(b)) = 0$$

soit $(f(b) - f(a))\, g'(c) = (g(b) - g(a))\, f'(c)$.
C'est le théorème généralisé des accroissements finis.

Avec l'exercice qui suit, on propose une version un peu plus générale du théorème des accroissements finis.

Exercice 3.7 *Soient $(f_k)_{1\leq k\leq n}$ et $(g_k)_{1\leq k\leq n}$ deux familles de fonctions à valeurs réelles, continues sur $[a,b]$, dérivables sur $]a,b[$ et telles que $g_k(a) \neq g_k(b)$ pour tout k compris entre 1 et n. Monter qu'il existe un réel c dans $]a,b[$ tel que :*

$$\sum_{k=1}^{n} f_k'(c) = \sum_{k=1}^{n} g_k'(c) \frac{f_k(b) - f_k(a)}{g_k(b) - g_k(a)}$$

Solution 3.7 *On considère la fonction φ définie sur $[a,b]$ par :*

$$\varphi(x) = \sum_{k=1}^{n} \left(f_k(x) - f_k(a) - \lambda_k (g_k(x) - g_k(a))\right)$$

où les constantes λ_k sont choisies telles que $\varphi(a) = \varphi(b)$. On peut prendre :

$$\lambda_k = \frac{f_k(b) - f_k(a)}{g_k(b) - g_k(a)} \quad (1 \leq k \leq n)$$

Cette fonction est continue sur $[a,b]$, dérivable sur $]a,b[$ avec $\varphi(a) = \varphi(b)$. Le théorème de Rolle nous dit alors qu'il existe un réel c dans $]a,b[$ tel que $\varphi'(c)$, ce qui donne le résultat annoncé.

Le théorème des accroissements finis est encore valable pour les fonctions de plusieurs variables à valeurs réelles. Précisément, on a le résultat suivant.

Théorème 3.16 *Soit f une fonction à valeurs réelles définie et différentiable sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$. Si a, b sont deux points distincts de $\mathcal{O}$ tels que le segment $[a,b]$ soit contenu dans $\mathcal{O}$, il existe alors un point $c \in]a,b[$ tel que :*

$$f(b) - f(a) = df(c)(b-a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(c)(b_k - a_k)$$

Démonstration. On se ramène au cas des fonctions d'une variable réelle en considérant la fonction g définie sur $[0,1]$ par :

$$g(t) = f(a + t(b-a))$$

Cette fonction est bien définie sur $[0,1]$ du fait que $[a,b] \subset \mathcal{O}$ et elle continue sur cet intervalle, dérivable sur $]0,1[$ avec :

$$g'(t) = df(a + t(b-a))(b-a) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(a + t(b-a))(b_k - a_k)$$

Le théorème de Rolle nous dit alors qu'il existe un réel θ dans $]0,1[$ tel que $g(1) - g(0) = g'(\theta)$, ce qui s'écrit $f(b) - f(a) = df(c)(b-a)$ avec $c = a + \theta(b-a)$ dans $]a,b[$. ■

Le théorème des accroissements finis n'est pas valable pour les fonctions à valeurs dans $\mathbb{C}$ (ou dans $\mathbb{R}^n$ avec $n \geq 2$) comme le montre l'exemple de la fonction $x \mapsto e^{ix}$ sur $[0,2\pi]$.

Toutefois on a le résultat suivant.

Théorème 3.17 (Inégalité des accroissements finis) *Si f est une fonction définie sur$[a,b]$, à valeurs dans un espace préhilbertien $(E, \langle \cdot \mid \cdot \rangle)$, continue sur $[a,b]$ et dérivable sur $]a,b[$, il existe alors un point $c \in]a,b[$ tel que $\|f(b) - f(a)\| \leq \|f'(c)\|(b-a)$.*

Démonstration. On se ramène au cas des fonctions à valeurs réelles en introduisant la fonction g définie sur $[a,b]$ par $g(x) = \langle f(x) \mid f(b) - f(a) \rangle$.

Cette fonction est continue sur $[a,b]$, dérivable sur $]a,b[$ avec $g'(x) = \langle f'(x) \mid f(b) - f(a) \rangle$.

Le théorème des accroissements finis nous assure de l'existence d'un réel $c \in]a,b[$ tel que :

$$\|f(b) - f(a)\|^2 = g(b) - g(a) = (b-a)\langle f'(c) \mid f(b) - f(a) \rangle$$

puis avec l'inégalité de Cauchy-Schwarz, on déduit que :

$$\|f(b) - f(a)\|^2 \leq (b-a)\|f'(c)\| \|f(b) - f(a)\|$$

ce qui entraîne $\|f(b) - f(a)\| \leq \|f'(c)\|(b-a)$. ■

Dans le cas où la fonction dérivée f' est majorée sur $]a,b[$, en notant M un majorant de cette fonction, on a, pour tout couple (x,y) de points de $[a,b]$:

$$\|f(y) - f(x)\| \leq M|y - x|$$

c'est-à-dire que f est lipschitzienne sur $[a,b]$.

Réciproquement si f est M-lipschitzienne sur $[a,b]$, f' est alors majorée par M sur $]a,b[$.

On dispose également de la version suivante plus générale de l'inégalité des accroissements finis.

Théorème 3.18 *Soient f une fonction définie sur $[a,b]$, à valeurs dans un espace vectoriel normé $(E, \|\cdot\|)$ et g une fonction définie sur $[a,b]$ à valeurs dans $\mathbb{R}$, ces fonctions étant continues sur $[a,b]$ et dérivables sur $]a,b[$. Si $\|f'(x)\| \leq g'(x)$ pour tout $x \in]a,b[$, on a alors $\|f(b) - f(a)\| \leq g(b) - g(a)$.*

Démonstration. On suppose dans un premier temps que $\|f'(x)\| < g'(x)$ pour tout $x \in]a,b[$.

On se fixe un réel $\alpha \in]a,b[$ et on note :

$$E = \{x \in [\alpha,b] \mid \|f(x) - f(\alpha)\| > g(x) - g(\alpha)\}$$

Cet ensemble est ouvert dans $[\alpha, b]$ comme image réciproque d'un ouvert par une application continue.

En le supposant non vide, on note γ la borne inférieure de E.

On a $\gamma \neq b$ du fait que E qui est ouvert ne peut être réduit à $\{b\}$.

Si $\gamma = \alpha$, par définition de la borne inférieure, pour tout réel $\varepsilon > 0$ assez petit il existe alors un réel $x \in E \cap]\alpha, \alpha + \varepsilon[$ (E est ouvert), donc $\dfrac{\|f(x) - f(\alpha)\|}{x - \alpha} > \dfrac{g(x) - g(\alpha)}{x - \alpha}$ qui par passage à la limite quand ε tend vers 0 donne $\|f'(\alpha)\| \geq g'(\alpha)$ en contradiction avec $\|f'(\alpha)\| < g'(\alpha)$.

On a donc $\gamma \in]\alpha, b[$, $\gamma \notin E$ puisque E est ouvert et :

$$\|f(\gamma) - f(\alpha)\| \leq g(\gamma) - g(\alpha), \ \|f'(\gamma)\| < g'(\gamma)$$

ce qui entraîne pour $x > \gamma$ voisin de γ :

$$\frac{\|f(x) - f(\gamma)\|}{x - \gamma} < \frac{g(x) - g(\gamma)}{x - \gamma}$$

et :

$$\left\{ \begin{array}{l} \|f(x) - f(\gamma)\| < g(x) - g(\gamma) \\ \|f(\gamma) - f(\alpha)\| \leq g(\gamma) - g(\alpha) \end{array} \right.$$

qui par addition donne :

$$\|f(x) - f(\alpha)\| \leq \|f(x) - f(\gamma)\| + \|f(\gamma) - f(\alpha)\| < g(x) - g(\alpha)$$

c'est-à-dire $x \notin E$, en contradiction avec le fait que γ est la borne inférieure de E.

En définitive l'ensemble E est vide et $\|f(x) - f(\alpha)\| \leq g(x) - g(\alpha)$ pour tout $x \in [\alpha, b]$.

En particulier, on a $\|f(b) - f(\alpha)\| \leq g(b) - g(\alpha)$ et faisant tendre α vers a on obtient $\|f(b) - f(a)\| \leq g(b) - g(a)$.

Le cas général s'obtient en remplaçant la fonction g par la fonction $g_\varepsilon : x \mapsto g(x) + \varepsilon x$ avec $\varepsilon > 0$ quelconque.

On a $\|f'(x)\| < g'_\varepsilon(x) = g'(x) + \varepsilon$, donc $\|f(b) - f(a)\| \leq g(b) - g(a) + \varepsilon(b - a)$ et en faisant tendre ε vers 0 on a le résultat annoncé. ■

Dans le cas où f' est majorée par M sur $]a, b[$, en prenant $g(x) = Mx$, on retrouve l'inégalité classique des accroissements finis.

3.4 Quelques applications du théorème des accroissements finis

Comme le théorème de Rolle, le théorème des accroissements finis est important pour ses nombreuses applications.

Les intervalles considérés sont supposés non réduit à un point.

3.4.1 Sens de variation d'une fonction

Théorème 3.19 *Si f est une fonction à valeurs réelles dérivable sur un intervalle réel I, alors f est croissante sur I si, et seulement si, $f'(x) \geq 0$ pour tout x dans I.*

Démonstration. Si f est croissante sur I alors pour $x \neq y$ dans I le taux d'accroissement $\dfrac{f(y) - f(x)}{y - x}$ est positif ou nul et en passant à la limite quand y tend vers x, on déduit que $f'(x) \geq 0$.

Réciproquement si $f'(x) \geq 0$ pour tout x dans I, en utilisant le théorème des accroissements finis on déduit alors que pour tout $y > x$ dans I on a $f(y) - f(x) = f'(z)(y-x) \geq 0$. ■

Si le théorème de Rolle n'est pas lié à la connexité de l'intervalle $[a,b]$ on peut remarquer que le résultat précédent l'est.

Si on considère la fonction $f : x \mapsto -\frac{1}{x}$ sur $\mathbb{R}^*$, on a $f'(x) > 0$ et f n'est pas croissante sur $\mathbb{R}^*$.

Du théorème précédent, on déduit les résultats classiques suivants.

Corollaire 3.2 *Soient f, g deux fonctions dérivables sur un intervalle réel I.*

1. *La fonction f est décroissante sur I si, et seulement si, $f'(x) \leq 0$ pour tout x dans I.*
2. *La fonction f est constante sur I si, et seulement si, $f'(x) = 0$ pour tout x dans I.*
3. *Si $f'(x) > 0$ [resp. $f'(x) < 0$] pour tout x dans I, alors la fonction f est strictement croissante [resp. strictement décroissante] sur I.*
4. *Si $f'(x) \leq g'(x)$ pour tout x dans $I = [a,b]$, alors :*
$$\forall x \in [a,b], \ f(x) - f(a) \leq g(x) - g(a).$$
5. *Si $m \leq f'(x) \leq M$ pour tout x dans $I = [a,b]$, alors :*
$$\forall x \in [a,b], \ m(x-a) \leq f(x) - f(a) \leq M(x-a).$$

Démonstration.

1. On applique le théorème précédent à la fonction $-f$.
2. Si $f' = 0$, alors la fonction f est à la fois croissante et décroissante, donc constante. La réciproque est évidente.
3. Si $f'(x) > 0$ pour tout x dans I, on sait déjà que la fonction f est croissante. Si il existe $a < b$ dans I tels $f(a) = f(b)$, on a alors :
$$\forall x \in [a,b], \ f(a) \leq f(x) \leq f(b) = f(a)$$
et la fonction f est constante sur $[a,b]$ d'intérieur non vide ce qui entraîne $f' = 0$ sur $[a,b]$ en contradiction avec $f' > 0$.
4. Résulte de la croissance de la fonction $h = g - f$.
5. On applique ce qui précède à (mx, f) et (f, Mx).

■

On peut remarquer que le point 3. de ce corollaire implique le théorème précédent. Les deux résultats sont donc équivalents.

En effet, en notant, pour tout réel $m > 0$, $g_m(x) = f(x) + m$, l'hypothèse $f'(x) \geq 0$ pour tout x dans I entraîne $g'_m(x) > 0$ pour tout x dans I, donc g_m est strictement croissante et pour $x < y$ dans I, on a $f(x) + m < f(y) + m$ pour tout $m > 0$ ce qui entraîne $f(x) \leq f(y)$ en faisant tendre m vers 0.

3.4.2 Limites et dérivation

Théorème 3.20 *Soit f une fonction à valeurs réelles continue sur $[a,b]$ et dérivable sur $]a,b[\setminus\{c\}$ où c est un point de $]a,b[$. Si la fonction dérivée f' a une limite ℓ en c, alors f est dérivable en c avec $f'(c)=\ell$.*

Démonstration. Comme $c\in\,]a,b[$, il existe un réel $\delta>0$ tel que $]c-\delta,c+\delta[\,\subset\,]a,b[$.

Pour tout $x\in\,]c-\delta,c+\delta[\setminus\{c\}$ il existe un réel d_x strictement compris entre c et x tel que $\dfrac{f(x)-f(c)}{x-c}=f'(d_x)$ et avec $\lim\limits_{x\to c}d_x=c$, $\lim\limits_{t\to c}f'(t)=\ell$, on déduit que :

$$\lim_{x\to c}\frac{f(x)-f(c)}{x-c}=\lim_{x\to c}f'(d_x)=\ell$$

c'est-à-dire que f est dérivable en c avec $f'(c)=\ell$. ■

On peut remarquer que la fonction f' est également continue en c.

Exercice 3.8 *Montrer que la fonction f définie par $f(0)=0$ et $f(x)=e^{-\frac{1}{x^2}}$ pour $x\neq 0$ est indéfiniment dérivable sur $\mathbb{R}$ avec $f^{(n)}(0)=0$ pour tout entier naturel n.*

Solution 3.8 *Cette fonction est continue sur $\mathbb{R}$, dérivable sur $\mathbb{R}^*$ avec $f'(x)=\dfrac{2}{x^3}e^{-\frac{1}{x^2}}$ pour tout $x\in\mathbb{R}^*$, $\lim\limits_{x\to 0}f'(x)=0$, elle est donc dérivable en 0 de dérivée nulle et f' est continue sur $\mathbb{R}$.*
En supposant, pour $n\geq 1$, que f est de classe $\mathcal{C}^n$ sur $\mathbb{R}$ avec $f^{(n)}(0)=0$, $f^{(n)}(x)=P_n\left(\dfrac{1}{x}\right)e^{-\frac{1}{x^2}}$ pour $x\neq 0$, où P_n est une fonction polynomiale de degré $3n$, on déduit que f est de classe $\mathcal{C}^{n+1}$ sur $\mathbb{R}^*$ avec :

$$f^{(n+1)}(x)=\left(\frac{2}{x^3}P_n\left(\frac{1}{x}\right)-\frac{1}{x^2}P_n'\left(\frac{1}{x}\right)\right)e^{-\frac{1}{x^2}}=P_{n+1}\left(\frac{1}{x}\right)e^{-\frac{1}{x^2}},$$

le polynôme P_{n+1} étant de degré $3n+3$. Puis avec $\lim\limits_{x\to 0}f^{(n+1)}(x)=0$, on déduit que f est de classe $\mathcal{C}^{n+1}$ sur $\mathbb{R}$ avec $f^{(n+1)}(0)=0$.
On dispose ainsi d'une fonction de classe $\mathcal{C}^\infty$ qui n'est pas développable en série entière au voisinage de 0.

Théorème 3.21 (L'Hospital) *Soient f,g deux fonctions à valeurs réelles continues sur un intervalle ouvert I, dérivables sur $I\setminus\{c\}$ avec $g'(x)\neq 0$ pour tout $x\in I\setminus\{c\}$ où $c\in I$. Si $\lim\limits_{x\to c}\dfrac{f'(x)}{g'(x)}=\ell$, on a alors $\lim\limits_{x\to c}\dfrac{f(x)-f(c)}{g(x)-g(c)}=\ell$.*

Démonstration. De $g'(t)\neq 0$ pour tout $t\in I\setminus\{c\}$ on déduit avec le théorème des accroissements finis que $g(x)\neq g(c)$ pour tout $x\in I\setminus\{c\}$.

Le théorème des accroissements finis généralisé appliqué aux fonctions f,g sur l'intervalle compact d'extrémités x,c, où $x\in I\setminus\{c\}$ permet d'écrire $\dfrac{f(x)-f(c)}{g(x)-g(c)}=\dfrac{f'(d_x)}{g'(d_x)}$ avec d_x strictement compris entre x et c. Puis avec $\lim\limits_{x\to c}d_x=c$, $\lim\limits_{t\to c}\dfrac{f'(t)}{g'(t)}=\ell$, on déduit que :

$$\lim_{x\to c}\frac{f(x)-f(c)}{g(x)-g(c)}=\lim_{x\to c}\frac{f'(d_x)}{g'(d_x)}=\ell$$

$\blacksquare$

Si f, g sont dérivables en c avec $g'(c) \neq 0$ la règle de l'Hospital n'est pas utile. Il suffit en effet d'écrire que :

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

et d'utiliser la définition du nombre dérivé.

Remarque 3.5 *La réciproque du théorème précédent est fausse.*
Considérons par exemple la fonction f définie par $f(x) = x^2 \sin\left(\frac{1}{x}\right)$ sur $\mathbb{R}^*$, prolongé par continuité en 0 et la fonction g définie par $g(x) = x$ sur $[-1, 1]$.
La fonction f est dérivable avec $f'(0) = 0$, $f'(x) = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)$ pour $x \neq 0$. On a $\lim\limits_{x \to 0} \frac{f(x) - f(0)}{g(x) - g(0)} = \lim\limits_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0$ et $\frac{f'(x)}{g'(x)}$ n'a pas de limite en 0.

3.4.3 Intégration et dérivation

Si f est une fonction dérivable sur $[a, b]$, il n'est pas toujours possible d'écrire que $\int_a^b f'(x)\,dx = f(b) - f(a)$.

Dans le cas où la dérivée est Riemann-intégrable, en utilisant les sommes de Riemann, on a le résultat suivant.

Théorème 3.22 *Si f est une fonction dérivable sur $[a, b]$ avec f' Riemann-intégrable sur $[a, b]$, on a alors :*

$$\int_a^b f'(x)\,dx = f(b) - f(a)$$

Démonstration. Pour tout entier naturel non nul n, on désigne par $\sigma_n = (a_i)_{0 \leq i \leq n}$ la subdivision de pas constant $\frac{b-a}{n}$ où $a_i = a + i\frac{b-a}{n}$, pour $0 \leq i \leq n - 1$.

En utilisant le théorème des accroissements finis, on peut écrire pour tout $n \geq 1$:

$$\begin{aligned} f(b) - f(a) &= \sum_{i=0}^{n-1} \left(f(a_{i+1}) - f(a_i)\right) = \sum_{i=0}^{n-1} f'(c_i)(a_{i+1} - a_i) \\ &= \frac{b-a}{n} \sum_{k=0}^{n-1} f'(c_i) \end{aligned}$$

où $c_i \in \left]a_i, a_{i+1}\right[$ et on reconnaît là une somme de Riemann pour la fonction f' qui tend vers $\int_a^b f'(x)\,dx$ quand n tend vers l'infini. $\blacksquare$

Le résultat qui suit est souvent utilisé pour montrer la dérivabilité d'une fonction définie par une intégrale sur un segment.

Théorème 3.23 *Soient I est un intervalle réel non réduit à un point, $a < b$ dans $\mathbb{R}$ et f une fonction à valeurs réelles définie et continue sur $I \times [a, b]$ telle que la dérivée partielle $\frac{\partial f}{\partial x}(x, t)$*

existe en tout point de $I \times [a,b]$ *, la fonction* $\dfrac{\partial f}{\partial x}$ *étant continue sur* $I \times [a,b]$ *.*
La fonction φ *définie sur* I *par :*

$$\forall x \in I,\ \varphi(x) = \int_a^b f(x,t)\,dt$$

est dérivable sur I *de dérivée :*

$$\varphi'(x) = \int_a^b \frac{\partial f}{\partial x}(x,t)\,dt$$

Démonstration. On se fixe un réel x_0 dans I et un réel $r > 0$ tel que l'intervalle $[x_0 - r, x_0 + r]$ soit contenu dans I.

Pour tout réel h tel que $0 < |h| < r$, on a :

$$\begin{aligned}\tau(h) &= \frac{\varphi(x_0+h) - \varphi(x_0)}{h} - \int_a^b \frac{\partial f}{\partial x}(x_0,t)\,dt \\ &= \int_a^b \left(\frac{f(x_0+h,t) - f(x_0,t)}{h} - \frac{\partial f}{\partial x}(x_0,t)\right) dt\end{aligned}$$

En utilisant le théorème des accroissements finis, on peut écrire pour tout $t \in [a,b]$:

$$\frac{f(x_0+h,t) - f(x_0,t)}{h} - \frac{\partial f}{\partial x}(x_0,t) = \frac{\partial f}{\partial x}(x_0 + \theta_{h,t}, t) - \frac{\partial f}{\partial x}(x_0,t)$$

où $\theta_{h,t} \in \,]0,1[$. La fonction $\dfrac{\partial f}{\partial x}$ qui est continue sur le compact $K = [x_0 - r, x_0 + r] \times [a,b]$ y est uniformément continue, donc pour tout réel $\varepsilon > 0$ on peut trouver un réel $\eta \in \,]0,r[$ tel que $\left|\dfrac{\partial f}{\partial x}(x,t) - \dfrac{\partial f}{\partial x}(u,v)\right| < \varepsilon$ pour tous (x,t) et (u,v) dans K tels que $\|(x,t)-(u,v)\|_\infty < \eta$.
Pour tout réel h tel que $0 < |h| < \eta$, on a $\|(x_0 + \theta_{h,t}h, t) - (x_0,t)\|_\infty = |\theta_{h,t}h| < |h| < \eta$ pour tout $t \in [a,b]$, ce qui entraîne :

$$\left|\frac{\partial f}{\partial x}(x_0 + \theta_{h,t}h, t) - \frac{\partial f}{\partial x}(x_0,t)\right| < \varepsilon$$

et en conséquence :

$$|\tau(h)| \leq \int_a^b \left|\frac{f(x_0+h,t) - f(x_0,t)}{h} - \frac{\partial f}{\partial x}(x_0,t)\right| dt \leq (b-a)\,\varepsilon$$

Ce qui prouve la dérivabilité de φ en x_0 et donne la valeur de $\varphi'(x_0)$. ■

On peut, par exemple, utiliser ce résultat pour calculer la valeur de l'intégrale de Gauss $\displaystyle\int_0^{+\infty} e^{-t^2}dt.$

Pour ce faire, on considère les fonctions F et G définies sur $\mathbb{R}^+$ par :

$$F(x) = \left(\int_0^x e^{-t^2}dt\right)^2,\ G(x) = \int_0^1 \frac{e^{-x^2(t^2+1)}}{t^2+1}dt$$

Ces fonctions sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^+$ avec :

$$F'(x) = 2e^{-x^2}\int_0^x e^{-t^2}dt \text{ et } G'(x) = -2xe^{-x^2}\int_0^1 e^{-x^2t^2}dt$$

(la fonction $(x,t) \mapsto \dfrac{e^{-x^2\left(t^2+1\right)}}{t^2+1}$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ et on intègre sur un segment).

Le changement de variable $y = xt$, pour $x > 0$, dans $G'(x)$ donne :

$$G'(x) = -2e^{-x^2} \int_0^x e^{-y^2} dy = -F'(x)$$

ce résultat étant encore valable pour $x = 0$. On a donc :

$$\forall x \in \mathbb{R}^+, \ F'(x) + G'(x) = 0$$

ce qui entraîne :

$$\forall x \in \mathbb{R}^+, \ F(x) + G(x) = F(0) + G(0) = \int_0^1 \frac{dt}{t^2+1} = \frac{\pi}{4}$$

Puis avec :

$$0 \leq G(x) \leq e^{-x^2} \int_0^1 \frac{dt}{t^2+1} = \frac{\pi}{4} e^{-x^2} \underset{x \to +\infty}{\to} 0$$

on déduit que $\lim\limits_{x \to +\infty} F(x) = \dfrac{\pi}{4} - \lim\limits_{x \to +\infty} G(x) = \dfrac{\pi}{4}$, soit :

$$\int_0^{+\infty} e^{-t^2} dt = \frac{\sqrt{\pi}}{2}$$

Corollaire 3.3 *Soient I un intervalle réel non réduit à un point, $a < b$ dans $\mathbb{R}$ et f une fonction à valeurs réelles définie et continue sur $I \times [a,b]$ telle que la dérivée partielle $\dfrac{\partial f}{\partial x}(x,t)$ existe en tout point de $I \times [a,b]$, la fonction $\dfrac{\partial f}{\partial x}$ étant continue sur $I \times [a,b]$.*
Si u, v sont deux fonctions dérivables de I dans $[a,b]$, alors la fonction ψ définie sur I par :

$$\forall x \in I, \ \psi(x) = \int_{u(x)}^{v(x)} f(x,t)\, dt$$

est dérivable sur I de dérivée :

$$\psi'(x) = f(x, v(x))\, v'(x) - f(x, u(x))\, u'(x) + \int_{u(x)}^{v(x)} \frac{\partial f}{\partial x}(x,t)\, dt$$

Démonstration. En écrivant que :

$$\psi(x) = \int_a^{v(x)} f(x,t)\, dt - \int_a^{u(x)} f(x,t)\, dt$$

il suffit de considérer les fonctions de la forme :

$$\alpha(x) = \int_a^{v(x)} f(x,t)\, dt$$

où v est une fonction dérivable de I dans $[a,b]$.

On se fixe un réel x_0 dans I et un réel $r > 0$ tel que l'intervalle $[x_0 - r, x_0 + r]$ soit contenu dans I et pour tout x dans $]x_0 - r, x_0 + r[$ on peut écrire que :

$$\alpha(x) = \int_a^{v(x_0)} f(x,t)\, dt + \int_{v(x_0)}^{v(x)} f(x,t)\, dt = \beta(x) + \gamma(x)$$

Le théorème précédent nous dit que la fonction β est dérivable sur $]x_0 - r, x_0 + r[$ de dérivée :

$$\beta'(x) = \int_a^{v(x_0)} \frac{\partial f}{\partial x}(x,t)\,dt$$

ce résultat étant en particulier valable pour $x = x_0$. Il nous suffit donc de montrer que la fonction γ est dérivable en x_0 de dérivée :

$$\gamma'(x_0) = f(x_0, v(x_0))\, v'(x_0)$$

Pour ce faire, on considère pour h tel que $0 < |h| < r$, la quantité :

$$\begin{aligned}\tau(h) &= \frac{\gamma(x_0+h) - \gamma(x_0)}{h} - f(x_0, v(x_0))\, v'(x_0) \\ &= \int_{v(x_0)}^{v(x_0+h)} \frac{f(x_0+h,t)}{h} dt - f(x_0, v(x_0))\, v'(x_0)\end{aligned}$$

et il s'agit de montrer que $\lim\limits_{h\to 0} \tau(h) = 0$.

Par définition du nombre dérivé, on peut écrire que :

$$v'(x_0) = \frac{v(x_0+h) - v(x_0)}{h} + \delta(h)$$

avec $\lim\limits_{h\to 0} \delta(h) = 0$ et :

$$\tau(h) = \int_{v(x_0)}^{v(x_0+h)} \frac{f(x_0+h,t) - f(x_0, v(x_0))}{h} dt - f(x_0, v(x_0))\,\delta(h)$$

La fonction f étant continue sur le compact $K = [x_0 - r, x_0 + r] \times [a, b]$ y est uniformément continue, donc pour tout réel $\varepsilon > 0$ on peut trouver un réel $\eta \in]0, r[$ tel que $|f(x,t) - f(x',t')| < \varepsilon$ pour tous (x,t) et (x',t') dans K tels que $\|(x,t) - (x',t')\|_\infty < \eta$. De plus avec la continuité de la fonction v en x_0, on peut trouver un réel $\eta' \in]0, r[$ tel que $|v(x_0+h) - v(x_0)| < \eta$ dès que $|h| < \eta'$. Pour $0 < |h| < \min(\eta, \eta')$, on a alors en notant J_h l'intervalle d'extrémités $v(x_0)$ et $v(x_0+h)$:

$$\begin{aligned}&\left|\int_{v(x_0)}^{v(x_0+h)} \frac{f(x_0+h,t) - f(x_0, v(x_0))}{h} dt\right| \\ &\leq \left|\frac{v(x_0+h) - v(x_0)}{h}\right| \sup_{t\in J_h} |f(x_0+h,t) - f(x_0, v(x_0))| \\ &\leq |v'(x_0) - \delta(h)|\,\varepsilon\end{aligned}$$

ce qui donne en définitive pour $0 < |h| < \min(\eta, \eta')$:

$$|\tau(h)| \leq |v'(x_0)|\,\varepsilon + |\delta(h)|\,(1+\varepsilon)$$

et $|\tau(h)| \leq |v'(x_0)|\,\varepsilon + \varepsilon(1+\varepsilon)$ pour $h > 0$ assez petit, ce qui prouve que $\lim\limits_{h\to 0} \tau(h) = 0$. ■

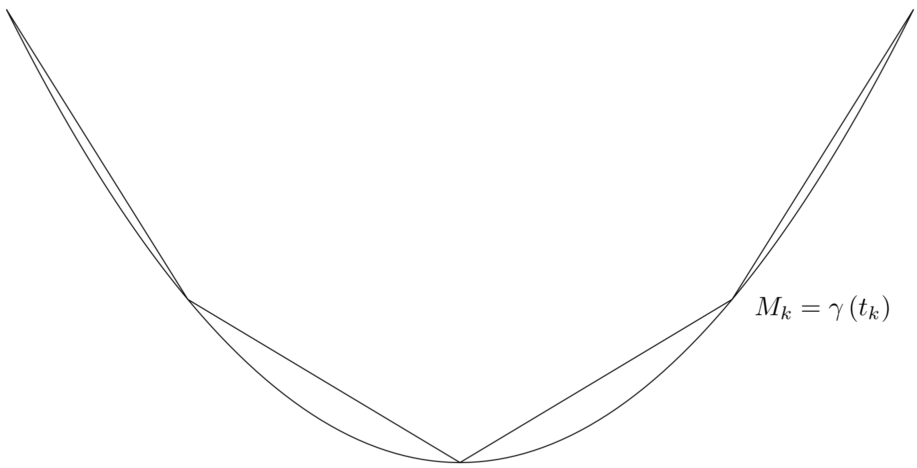

FIGURE 3.5 –

3.4.4 Arcs rectifiables, longueur d'un arc paramétré

Soit γ un arc géométrique paramétré par une application continue $\gamma : [a,b] \to \mathcal{P}$.

Un tel arc est dit compact ($\gamma([a,b])$ est compact dans $\mathcal{P}$ comme image d'un compact par une application continue).

À toute subdivision du segment $[a,b]$:

$$\sigma = \left\{ (t_0, t_1, ..., t_p) \in \mathbb{R}^{p+1} \mid a = t_0 < t_1 < ... < t_p = b \right\}$$

on associe la ligne polygonale continue γ_σ de sommets $M_k = \gamma(t_k)$ $(0 \leq k \leq p)$ représentée par la figure (3.5).

Une telle ligne polygonale peut être définie par la paramétrisation $\gamma_\sigma : [a,b] \to \mathcal{P}$, avec :

$$\forall k \in \{0, ..., p-1\}, \ \forall t \in [t_k, t_{k+1}], \ \gamma_\sigma(t) = (1-t) M_k + t M_{k+1}$$

La longueur de γ_σ est alors naturellement définie par :

$$L(\gamma_\sigma) = \sum_{i=0}^{p-1} \left\| \overrightarrow{M_i M_{i+1}} \right\| = \sum_{i=0}^{p-1} \| \gamma(t_{i+1}) - \gamma(t_i) \|$$

On désigne par $\Sigma_{a,b}$ l'ensemble de toutes les subdivisions du segment $[a,b]$.

Définition 3.1 *On dit que l'arc paramétré continu $\gamma : [a,b] \to \mathbb{R}^n$ est rectifiable si l'ensemble $\{L(\gamma_\sigma) \mid \sigma \in \Sigma_{a,b}\}$ admet une borne supérieure et dans ce cas, on note :*

$$L(\gamma) = \sup_{\sigma \in \Sigma_{a,b}} L(\gamma_\sigma)$$

la longueur de l'arc $(\gamma, [a,b])$.

Si $(\theta, [\alpha, \beta])$ est un arc paramétré continu équivalent à $(\gamma, [a,b])$, il existe alors un homéomorphisme $\varphi : [a,b] \to [\alpha, \beta]$ tel que $\gamma = \theta \circ \varphi$ et le changement de paramètre φ permet de réaliser une bijection de l'ensemble $\Sigma_{\alpha,\beta}$ des subdivisions $\Sigma_{a,b}$ de $[\alpha, \beta]$ sur l'ensemble des subdivisions de $[a,b]$ (si φ est décroissante cette bijection inverse alors l'ordre des points des subdivisions) et on a $L(\gamma,[a,b]) = L(\theta,[\alpha,\beta])$.

La longueur d'un arc géométrique continu (quand elle est définie) ne dépend donc pas du choix d'une paramétrisation. De manière précise, on peut donner la définition suivante.

Définition 3.2 *Soit γ un arc géométrique compact et continu. On dit qu'il est rectifiable si pour toute paramétrisation $\gamma : [a,b] \to \mathcal{P}$, $(\gamma,[a,b])$ est rectifiable. La longueur de γ est alors la longueur de $(\gamma,[a,b])$ et on la note $L(\gamma)$.*

Théorème 3.24 *Si γ est un arc géométrique compact de classe $\mathcal{C}^k$ avec $k \geq 1$, paramétré par $\gamma : [a,b] \to \mathcal{P}$, il est alors rectifiable et sa longueur est donnée par :*

$$L(\gamma) = \int_a^b \|\gamma'(t)\| \, dt$$

Démonstration. Soit σ une subdivision de $[a,b]$. On a :

$$\begin{aligned} L(\gamma_\sigma) &= \sum_{k=0}^{p-1} \|\gamma(t_{k+1}) - \gamma(t_k)\| = \sum_{k=0}^{p-1} \left\| \int_{tk}^{t_{k+1}} \gamma'(t)\, dt \right\| \\ &\leq \sum_{i=0}^{p-1} \int_{tk}^{t_{k+1}} \|\gamma'(t)\| \, dt = \int_a^b \|\gamma'(t)\| \, dt < +\infty \end{aligned}$$

donc :

$$L(\gamma) = \sup_{\sigma \in \Sigma_{a,b}} L(\gamma_\sigma) \leq \int_a^b \|\gamma'(t)\| \, dt < +\infty$$

La fonction γ' étant continue sur le compact $[a,b]$, pour tout $\varepsilon > 0$, on peut trouver un réel $\eta > 0$ tel que :

$$|t - t'| < \eta \Rightarrow \|\gamma'(t) - \gamma'(t')\| < \frac{\varepsilon}{2(b-a)}$$

En prenant une subdivision σ de pas $h = \max\limits_{0 \leq k \leq p-1} (t_{k+1} - t_k)$, on a :

$$\begin{aligned} \left| \int_{t_k}^{t_{k+1}} \|\gamma'(t)\| \, dt - \|\gamma(t_{k+1}) - \gamma(t_k)\| \right| \leq & \left| \int_{t_k}^{t_{k+1}} \|\gamma'(t)\| \, dt - (t_{k+1} - t_k) \|\gamma'(t_i)\| \right| \\ & + \left| (t_{k+1} - t_k) \|\gamma'(t_k)\| - \|\gamma(t_{k+1}) - \gamma(t_k)\| \right| \end{aligned}$$

avec :

$$\begin{aligned} I_1 &= \left| \int_{t_k}^{t_{k+1}} \|\gamma'(t)\| \, dt - (t_{k+1} - t_k) \|\gamma'(t_i)\| \right| \\ &= \left| \int_{t_k}^{t_{k+1}} (\|\gamma'(t)\| - \|\gamma'(t_k)\|) \, dt \right| \\ &\leq \left| \int_{t_k}^{t_{k+1}} \|\gamma'(t) - \|\gamma'(t_k)\|\| \, dt \right| \leq (t_{k+1} - t_k) \frac{\varepsilon}{2(b-a)} \end{aligned}$$

D'autre part, on a :

$$\begin{aligned} I_2 &= \left| \|\gamma(t_{k+1}) - \gamma(t_k)\| - (t_{k+1} - t_k) \|\gamma'(t_k)\| \right| \\ &\leq \|\gamma(t_{k+1}) - \gamma(t_k) - (t_{k+1} - t_k) \gamma'(t_k)\| = \|\varphi_k(t_{k+1}) - \varphi_k(t_k)\| \end{aligned}$$

où on a noté $\varphi_k(t) = \gamma(t) - t\gamma'(t_k)$.

En écrivant que :

$$\|\varphi_k'(t)\| = \|\gamma'(t) - \gamma'(t_k)\| < \frac{\varepsilon}{2(b-a)}$$

sur $[t_k, t_{k+1}]$ et en utilisant le théorème des accroissements finis, on déduit que :

$$I_2 = \|\varphi_k(t_{k+1}) - \varphi_k(t_k)\| \leq (t_{k+1} - t_k)\frac{\varepsilon}{2(b-a)}$$

En définitive, on a :

$$I_1 + I_2 \leq (t_{k+1} - t_k)\frac{\varepsilon}{b-a}$$

et par addition :

$$\left|\int_a^b \|\gamma'(t)\|\,dt - L(\gamma_\sigma)\right| < \varepsilon$$

C'est-à-dire que pour tout $\varepsilon > 0$, on peut trouver une subdivision σ telle que :

$$\int_a^b \|\gamma'(t)\|\,dt - \varepsilon \leq L(\gamma) \leq \int_a^b \|\gamma'(t)\|\,dt$$

ce qui prouve que γ est rectifiable et que $L(\gamma) = \displaystyle\int_a^b \|\gamma'(t)\|\,dt$. ∎

3.4.5 Points fixes attractifs et répulsifs

Voir le paragraphe 8.3.3.

3.4.6 Majoration de l'erreur dans la méthode de Simpson

Théorème 3.25 *Soit f une fonction à valeurs réelles de classe $\mathcal{C}^4$ sur un intervalle $[a,b]$. On a :*

$$\left|\int_a^b f(x)\,dx - \frac{b-a}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right)\right| \leq \frac{M_4}{2880}(b-a)^5$$

où $M_4 = \sup_{x\in[a,b]} \left|f^{(4)}(x)\right|$.

Démonstration. On se ramène à l'intervalle $[-1,1]$ en utilisant le changement de variable $x = \frac{a+b}{2} + \frac{b-a}{2}t$ avec $t \in [-1,1]$, ce qui conduit à introduire la fonction g définie sur $[-1,1]$ par :

$$g(t) = f\left(\frac{a+b}{2} + \frac{b-a}{2}t\right)$$

L'erreur dans la méthode de Simpson sur $[a,b]$:

$$E(f) = \int_a^b f(x)\,dx - \frac{b-a}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right)$$

s'écrit alors :

$$E(f) = \frac{b-a}{2}\left(\int_{-1}^1 g(t)\,dt - \frac{1}{3}(g(-1) + 4g(0) + g(1))\right)$$

On désigne par φ la fonction définie sur $[0,1]$ par :

$$\forall x \in [0,1]\,,\ \varphi(x) = \int_{-x}^{x} g(t)\,dt - \frac{x}{3}\left(g(-x) + 4g(0) + g(x)\right)$$

(erreur dans la méthode de Simpson sur $[-x,x]$). Cette fonction est de classe $\mathcal{C}^4$ sur $[0,1]$ avec :

$$\begin{cases} \varphi'(x) = \dfrac{2}{3}\left(g(-x) + g(x)\right) - \dfrac{4}{3}g(0) - \dfrac{x}{3}\left(g'(x) - g'(-x)\right), \\ \varphi''(x) = \dfrac{1}{3}\left(g'(x) - g'(-x)\right) - \dfrac{x}{3}\left(g''(x) + g''(-x)\right) \\ \varphi'''(x) = -\dfrac{x}{3}\left(g'''(x) - g'''(-x)\right) \end{cases}$$

Avec le théorème des accroissements finis on peut écrire, pour $x > 0$:

$$\varphi'''(x) = -2\frac{x^2}{3}g^{(4)}(c_x)$$

où $c_x \in \left]-x, x\right[$, ce qui donne $|\varphi'''(x)| \leq 2\dfrac{x^2}{3}L_4$, où :

$$L_4 = \sup_{x\in[-1,1]} \left|g^{(4)}(x)\right| = \left(\frac{b-a}{2}\right)^4 M_4$$

On en déduit, tenant compte de $\varphi(0) = \varphi'(0) = \varphi''(0) = 0$, que :

$$\begin{cases} |\varphi''(x)| = \left|\displaystyle\int_0^x \varphi'''(t)\,dt\right| \leq \dfrac{2}{3}L_4\displaystyle\int_0^x t^2 dt = 2\dfrac{x^3}{9}L_4 \\ |\varphi'(x)| = \left|\displaystyle\int_0^x \varphi''(t)\,dt\right| \leq \dfrac{2}{9}L_4\displaystyle\int_0^x t^3 dt = \dfrac{x^4}{18}L_4 \\ |\varphi(x)| = \left|\displaystyle\int_0^x \varphi'(t)\,dt\right| \leq \dfrac{1}{18}L_4\displaystyle\int_0^x t^4 dt = \dfrac{x^5}{90}L_4 \end{cases}$$

Et pour $x = 1$, on obtient :

$$\begin{aligned} |\varphi(1)| &= \left|\int_{-1}^{1} g(t)\,dt - \frac{1}{3}\left(g(-1) + 4g(0) + g(1)\right)\right| \\ &\leq \frac{1}{90}L_4 = \frac{1}{90}\left(\frac{b-a}{2}\right)^4 M_4, \end{aligned}$$

ce qui donne :

$$|E(f)| \leq \frac{1}{90}\left(\frac{b-a}{2}\right)^5 M_4 = \frac{(b-a)^5}{2880}M_4$$

■

Cette méthode de démonstration est encore valable pour obtenir une majoration l'erreur de quadrature dans la méthode du point milieu ou la méthode du trapèze.

3.4.7 Suites de fonctions dérivables

Théorème 3.26 *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions continues de $[a,b]$ dans $\mathbb{R}$, dérivables sur $\left]a,b\right[$, qui converge simplement vers une fonction f. S'il existe une constante $M > 0$ telle que $|f_n'(x)| \leq M$ pour tout n et tout x dans $\left]a,b\right[$, alors la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément et f est continue.*

Démonstration. En utilisant le théorème des accroissements finis, on voit que les fonctions f_n et f sont M-lipschitzienne. En effet, on a :

$$\forall n \in \mathbb{N},\ \forall (x,y) \in [a,b]^2,\ |f_n(x) - f_n(y)| \leq M|x-y|$$

et en passant à la limite quand n tend vers l'infini, on obtient :

$$\forall (x,y) \in \mathbb{R}^2,\ |f(x) - f(y)| \leq M|x-y|$$

Ce qui entraîne déjà la continuité de f.

Pour tout $\varepsilon > 0$ il existe une subdivision $a = x_0 < x_1 < \cdots < x_p = b$ telle que $x_{k+1} - x_k < \varepsilon$ et de la convergence de chacune des suites $(f_n(x_k))_{n\in\mathbb{N}}$ vers $f(x_k)$, on déduit qu'il existe un entier n_0 tel que :

$$\forall k \in \{0,1,\cdots,p\},\ \forall n \geq n_0,\ |f_n(x_k) - f(x_k)| < \varepsilon.$$

Pour tout $x \in [a,b]$, il existe un entier k compris entre 0 et $p-1$ tel que x soit dans $[x_k, x_{k+1}]$ et pour $n \geq n_0$, on a :

$$\begin{aligned}|f_n(x) - f(x)| &\leq |f_n(x) - f_n(x_k)| + |f_n(x_k) - f(x_k)| + |f(x_k) - f(x)| \\ &\leq M|x-x_k| + \varepsilon + M|x-x_k| \leq (2M+1)\varepsilon\end{aligned}$$

ce qui permet de conclure à la convergence uniforme de $(f_n)_{n\in\mathbb{N}}$ vers f sur $[a,b]$. ■

Théorème 3.27 *Soit $(f_n)_{n\in\mathbb{N}}$ une suite de fonctions dérivables de $[a,b]$ dans $\mathbb{R}$ telle que la suite $(f'_n)_{n\in\mathbb{N}}$ converge uniformément sur $[a,b]$ vers une fonction g. S'il existe $x_0 \in [a,b]$ tel que la suite $(f_n(x_0))_{n\in\mathbb{N}}$ soit convergente alors la suite $(f_n)_{n\in\mathbb{N}}$ converge uniformément vers une fonction dérivable f et $f' = g$.*

Démonstration. Avec le théorème des accroissements finis, on peut écrire pour n, m entiers naturels et $x \in [a,b]$:

$$\begin{aligned}|f_n(x) - f_m(x)| &\leq |(f_n - f_m)(x) - (f_n - f_m)(x_0)| + |f_n(x_0) - f_m(x_0)| \\ &\leq \|f'_n - f'_m\|_\infty |x-x_0| + |f_n(x_0) - f_m(x_0)|\end{aligned}$$

où on note pour toute fonction h continue sur $[a,b]$, $\|h\|_\infty = \sup\limits_{x\in[a,b]} |h(x)|$. Il en résulte que :

$$\|f_n - f_m\|_\infty \leq (b-a)\|f'_n - f'_m\|_\infty + |f_n(x_0) - f_m(x_0)|$$

ce qui permet, d'après les hypothèses, de conclure que la suite $(f_n)_{n\in\mathbb{N}}$ vérifie le critère de Cauchy uniforme sur $[a,b]$ et donc qu'elle converge uniformément vers une fonction f.

Pour x, y dans $[a,b]$ et $n \in \mathbb{N}$ on peut écrire :

$$\begin{aligned}|f(x) - f(y) - (x-y)g(x)| \leq\ & |f(x) - f(y) - (f_n(x) - f_n(y))| \\ &+ |f_n(x) - f_n(y) - (x-y)f'_n(x)| \\ &+ |(x-y)(f'_n(x) - g(x))|\end{aligned}$$

avec :

$$\begin{aligned}|f(x) - f(y) - (f_n(x) - f_n(y))| &= \lim_{m\to+\infty} |(f_m - f_n)(x) - (f_m - f_n)(y)| \\ &\leq \lim_{m\to+\infty} \|f'_m - f'_n\|_\infty |x-y| \\ &\leq \|g - f'_n\|_\infty |x-y|\end{aligned}$$

et :

$$\left|(x-y)\left(f_n'(x)-g(x)\right)\right| \leq \|g-f_n'\|_\infty |x-y|$$

Ce qui donne, pour $x \neq y$:

$$\left|\frac{f(x)-f(y)}{x-y}-g(x)\right| \leq 2\|g-f_n'\|_\infty + \left|\frac{f_n(x)-f_n(y)}{x-y}-f_n'(x)\right|$$

Pour $\varepsilon > 0$ donné, on peut trouver un entier n tel que $\|g-f_n'\|_\infty < \varepsilon$ et pour cet entier, par définition du nombre dérivé, on peut trouver un réel $\eta > 0$ tel que pour $x \neq y$ dans $[a,b]$ vérifiant $|x-y| < \eta$ on ait $\left|\frac{f_n(x)-f_n(y)}{x-y}-f_n'(x)\right| < \varepsilon$. On a donc $\left|\frac{f(x)-f(y)}{x-y}-g(x)\right| \leq 3\varepsilon$ pour $x \neq y$ dans $[a,b]$ tels que $|x-y| < \eta$, ce qui signifie que f est dérivable en x avec $f'(x) = g(x)$.
■

3.4.8 Dérivées partielles

On sait que si f est une fonction différentiable sur un ouvert $\mathcal{O}$ de $\mathbb{R}^n$, alors elle est continue et admet des dérivées partielles en tout point de $\mathcal{O}$. Mais réciproquement une fonction peut très bien admettre des dérivées partielles en tout point de $\mathcal{O}$ sans être différentiable. Par exemple la fonction f définie sur $\mathbb{R}^2$ par $f(0,0) = 0$ et $f(x,y) = \frac{xy}{x^2+y^2}$ pour $(x,y) \neq (0,0)$ admet des dérivées partielles en tout point de $\mathbb{R}^2$ (pour $(x,y) \neq (0,0)$ c'est clair avec les opérations classiques, en $(0,0)$ on a $\frac{f(x,0)-f(0,0)}{x} = 0$ pour tout $x \neq 0$ qui entraîne $\frac{\partial f}{\partial x}(0,0) = 0$ et pour la dérivée en y on conclut par symétrie) et n'est pas différentiable en $(0,0)$ car non continue en ce point ($f(x,x) = \frac{1}{2}$ ne tend pas vers 0 quand x tend vers 0).

On a quand même le résultat important suivant.

Théorème 3.28 *Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^n$ et f une fonction définie sur $\mathcal{O}$ à valeurs réelles (ou dans un espace normé) admettant des dérivées partielles par rapport à toutes les variables en tout point de $\mathcal{O}$. Si ces dérivées partielles sont continues en un point a de $\mathcal{O}$ alors f est différentiable en a.*

Démonstration. On se place dans le cas d'une fonction de deux variables à valeurs réelles et on note (x,y) les variables. En supposant que $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ existent en tout point de $\mathcal{O}$ et qu'elles sont continues en (a,b), il s'agit de montrer, en notant :

$$h = x-a,\ k = y-b,\ p = \frac{\partial f}{\partial x}(a,b),\ q = \frac{\partial f}{\partial y}(a,b)$$

que :

$$f(x,y) = f(a,b) + ph + qk + o(|h|+|k|)$$

Pour ce faire on écrit que :

$$f(x,y) - f(a,b) = f(x,y) - f(x,b) + f(x,b) - f(a,b)$$

et en utilisant le théorème des accroissements finis, on a :

$$\begin{cases} f(x,y) - f(x,b) = (y-b)\dfrac{\partial f}{\partial y}(x,c_{x,y}) \\ f(x,b) - f(a,b) = (x-a)\dfrac{\partial f}{\partial x}(d_x,b) \end{cases}$$

avec $c_{x,y}$ compris entre y et b et d_x compris entre a et x, de sorte que :

$$\begin{aligned} f(x,y) - f(a,b) - ph - qk = h\left(\frac{\partial f}{\partial x}(d_x, b) - \frac{\partial f}{\partial x}(a,b)\right) \\ + k\left(\frac{\partial f}{\partial y}(x, c_{x,y}) - \frac{\partial f}{\partial y}(a,b)\right) \end{aligned}$$

Avec la continuité des dérivées partielles en (a,b), on déduit que pour tout réel $\varepsilon > 0$ il existe un réel $\eta > 0$ tel que les conditions $|u-a| < \eta$ et $|v-b| < \eta$ entraînent $\left|\frac{\partial f}{\partial x}(u,v) - \frac{\partial f}{\partial x}(a,b)\right| < \varepsilon$, $\left|\frac{\partial f}{\partial y}(u,v) - \frac{\partial f}{\partial y}(a,b)\right| < \varepsilon$. On déduit alors que pour $|x-a| < \eta$, $|y-b| < \eta$ on a :

$$|f(x,y) - f(a,b) - ph - qk| < (|h| + |k|)\,\varepsilon$$

c'est-à-dire le résultat souhaité. ■

La démonstration du théorème de Schwarz qui suit sur la symétrie des dérivées partielles d'ordre 2 en cas de continuité utilise également le théorème des accroissements finis.

Théorème 3.29 (Schwarz) *Soient $\mathcal{O}$ un ouvert de $\mathbb{R}^2$ et f une fonction définie sur $\mathcal{O}$ à valeurs réelles admettant sur $\mathcal{O}$ des dérivées partielles $\frac{\partial^2 f}{\partial x \partial y}$ et $\frac{\partial^2 f}{\partial y \partial x}$ continues en un point (a,b) de $\mathcal{O}$. On a alors :*

$$\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$$

Démonstration. Pour h, k strictement positifs tels que $[a, a+h] \times [b, b+k]$ soit contenu dans $\mathcal{O}$, on note :

$$g(h,k) = f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a,b)$$

En notant $\varphi(x) = f(x, b+k) - f(x,b)$, on a $g(h,k) = \varphi(a+h) - \varphi(a)$ qui avec le théorème des accroissements finis s'écrit $g(h,k) = h\varphi'(a+\theta h)$, soit :

$$g(h,k) = h\left(\frac{\partial f}{\partial x}(a+\theta h, b+k) - \frac{\partial f}{\partial x}(a+\theta h, b)\right)$$

ce qui peut s'écrire en utilisant à nouveau le théorème des accroissements finis :

$$g(h,k) = hk\frac{\partial^2 f}{\partial y \partial x}(a+\theta h, b+\theta' k)$$

En procédant de manière analogue avec la fonction ψ définie par $\psi(y) = f(a+h, y) - f(a,y)$, on a :

$$\begin{aligned} g(h,k) = \psi(b+k) - \psi(b) &= k\left(\frac{\partial f}{\partial y}(a+h, b+\eta k) - \frac{\partial f}{\partial y}(a, b+\eta k)\right) \\ &= hk\frac{\partial^2 f}{\partial x \partial y}(a+\eta' h, b+\eta k) \end{aligned}$$

Ce qui donne $\frac{\partial^2 f}{\partial y \partial x}(a+\theta h, b+\theta' k) = \frac{\partial^2 f}{\partial x \partial y}(a+\eta' h, b+\eta k)$ où les réels $\theta, \theta', \eta, \eta'$ sont dans $]0,1[$. En faisant tendre (h,k) vers $(0,0)$ et en utilisant la continuité des dérivées partielles d'ordre 2, en on déduit que $\frac{\partial^2 f}{\partial x \partial y}(a,b) = \frac{\partial^2 f}{\partial y \partial x}(a,b)$. ■

L'exemple de la fonction f définie sur $\mathbb{R}^2$ par $f(0,0) = 0$ et $f(x,y) = \dfrac{xy(x^2 - y^2)}{x^2 + y^2}$ pour $(x,y) \neq (0,0)$ nous montre que le résultat précédent est faux si on en enlève l'hypothèse de continuité des dérivées partielles d'ordre 2.

3.4.9 Le théorème de Darboux

On peut donner une démonstration du théorème de Darboux qui utilise le théorème des accroissements finis et le théorème des valeurs intermédiaires.

Théorème 3.30 (Darboux) *Si f est une fonction à valeurs réelles définie et dérivable sur un intervalle I, alors sa fonction dérivée f' vérifie la propriété des valeurs intermédiaires.*

Démonstration. Soient $a < b$ dans I. Si $f'(a) = f'(b)$ il n'y a alors rien à montrer.

On suppose donc que $f'(a) < f'(b)$ et on se donne $\lambda \in \left]f'(a), f'(b)\right[$. On définit les fonctions τ_a et τ_b sur $[a,b]$ par :

$$\forall x \in [a,b],\ \tau_a(x) = \begin{cases} f'(a) \text{ si } x = a \\ \dfrac{f(x) - f(a)}{x - a} \text{ si } x \neq a \end{cases}$$

et :

$$\forall x \in [a,b],\ \tau_b(x) = \begin{cases} f'(b) \text{ si } x = b \\ \dfrac{f(b) - f(x)}{b - x} \text{ si } x \neq b \end{cases}$$

Ces fonctions sont continues sur $[a,b]$ puisque f est dérivable sur I et on a :

$$\tau_a(a) = f'(a) < \lambda < f'(b) = \tau_b(b)$$

On a alors deux possibilités :

- soit $\tau_a(a) = f'(a) < \lambda \leq \tau_a(b)$ et le théorème des valeurs intermédiaires nous dit qu'il existe un réel c dans $]a,b]$ tel que
$$\lambda = \tau_a(c) = \frac{f(c) - f(a)}{c - a} = f'(d)$$
avec d entre a et c ;
- soit $\tau_a(b) < \lambda < f'(b) = \tau_b(b)$ et en remarquant que :
$$\tau_a(b) = \frac{f(b) - f(a)}{b - a} = \tau_b(a)$$
on a $\tau_b(a) < \lambda < \tau_b(b)$ et le théorème des valeurs intermédiaires nous dit qu'il existe un réel c dans $]a,b[$ tel que
$$\lambda = \tau_b(c) = \frac{f(b) - f(c)}{b - c} = f'(d)$$

■

On peut aussi donner une démonstration du théorème de Darboux qui utilise le fait que l'image d'un connexe de $\mathbb{R}^2$ par une application continue à valeurs réelles est un connexe de $\mathbb{R}$, donc un intervalle (caractérisation des connexes de $\mathbb{R}$), et le théorème des accroissements finis.

Théorème 3.31 (Darboux) *Si f est une fonction à valeurs réelles définie et dérivable sur un intervalle I, alors sa fonction dérivée f' vérifie la propriété des valeurs intermédiaires.*

Démonstration. Il s'agit de montrer que $f'(I)$ est connexe dans $\mathbb{R}$, ce qui revient à dire que c'est un intervalle.

L'ensemble :

$$C = \left\{ \frac{f(x) - f(y)}{x - y} \mid (x, y) \in I^2,\ x < y \right\}$$

est un connexe de $\mathbb{R}$ comme image du connexe de $\mathbb{R}^2$, $E = \{(x, y) \in I^2,\ x < y\}$ (cet ensemble est convexe donc connexe), par l'application continue φ définie sur I^2 par

$$\varphi(x, y) = \begin{cases} \dfrac{f(x) - f(y)}{x - y} & \text{si } y \neq x \\ f'(x) \ \text{ si } y = x \end{cases}$$

Le théorème des accroissements finis nous dit que tout $z \in C$ s'écrit $z = f'(t)$ avec $t \in \overset{\circ}{I}$ et en écrivant que $f'(t) = \lim\limits_{h \to 0^+} \dfrac{f(t+h) - f(t)}{h}$, on déduit que z est aussi dans $\overline{C}$. On a donc $C \subset f'(I) \subset \overline{C}$ avec C connexe, ce qui entraîne que $f'(I)$ est connexe. ■

3.4.10 Nombres de Liouville

On dit qu'un réel (ou complexe) α est algébrique s'il existe un polynôme P non nul à coefficients entiers relatifs tel que $P(\alpha) = 0$. Parmi tous ces polynômes il en existe un de degré minimal et en le divisant par son coefficient dominant on dispose d'un polynôme P_α unitaire à coefficients rationnels de degré minimal qui annule α. Il est facile de vérifier que ce polynôme est unique. Le polynôme P_α est appelé le polynôme minimal de α et le degré de P_α est le degré du nombre algébrique α. On vérifie facilement que P_α est irréductible dans $\mathbb{Q}[X]$, c'est-à-dire qu'il ne peut pas s'écrire comme produit de deux polynômes non constants.

On dit qu'un réel est transcendant s'il n'est pas algébrique.

Le théorème qui suit nous permet de construire une infinité de nombres transcendants.

Théorème 3.32 (Liouville) *Soit α un nombre algébrique de degré $d \geq 1$.*
Si $d = 1$ alors α est rationnel et il existe une constante $C_\alpha > 0$ telle que pour tout nombre rationnel $r = \dfrac{p}{q}$ ($p \in \mathbb{Z}$, $q \in \mathbb{N}^*$) distinct de α on a $\left|\alpha - \dfrac{p}{q}\right| \geq \dfrac{C_\alpha}{q}$.
Si $d \geq 2$ alors α est irrationnel et il existe une constante $C_\alpha > 0$ telle que pour tout nombre rationnel $r = \dfrac{p}{q}$ on a $\left|\alpha - \dfrac{p}{q}\right| \geq \dfrac{C_\alpha}{q^d}$.

Démonstration. Si $d = 1$ alors le polynôme minimal de α s'écrit $P_\alpha(X) = X - s$ avec $s \in \mathbb{Q}$. De $P_\alpha(\alpha) = 0$ on déduit alors que $\alpha = s = \dfrac{a}{b}$ avec $a \in \mathbb{Z}$, $b \in \mathbb{N}^*$. De plus pour $r = \dfrac{p}{q} \in \mathbb{Q}$ distinct de α on a :

$$0 < \left|\alpha - \frac{p}{q}\right| = \left|\frac{aq - bp}{qb}\right|$$

donc $|aq - bp| \neq 0$ dans $\mathbb{N}$, ce qui équivaut à $|aq - bp| \geq 1$. On a donc :

$$\left|\alpha - \frac{p}{q}\right| = \left|\frac{aq - bp}{qb}\right| \geq \frac{1}{b}\frac{1}{q}$$

Supposons maintenant que $d \geq 2$. Le nombre α ne peut être rationnel, sans quoi il serait annulé par un polynôme unitaire de degré 1, ce qui contredit le caractère minimal de d. En notant P_α le polynôme minimal de α il existe un entier naturel non nul n tel que $Q_\alpha = nP_\alpha \in \mathbb{Z}[X]$ et pour tout rationnel $r = \dfrac{p}{q}$, on a $q^d Q_\alpha(r) \in \mathbb{Z}$. De plus cet entier est non nul. En effet si $q^d Q_\alpha(r) = 0$ alors r est racine de Q_α et $Q_\alpha(X) = (X - r) R(X)$ avec $X - r$ et R dans $\mathbb{Q}[X]$ non constants ($d \geq 1$), en contradiction avec P_α irréductible dans $\mathbb{Q}[X]$. On a donc :

$$\left|q^d \left(Q_\alpha(\alpha) - Q_\alpha(r)\right)\right| = \left|q^d Q_\alpha(r)\right| \geq 1$$

D'autre part, le théorème des accroissements finis permet d'écrire :

$$Q_\alpha(\alpha) - Q_\alpha(r) = (\alpha - r) Q'_\alpha(s)$$

avec s compris entre α et r. On distingue alors deux cas :
soit $|\alpha - r| \leq 1$ et dans ce cas $s \in [\alpha - 1, \alpha + 1]$ de sorte que :

$$1 \leq \left|q^d Q_\alpha(r)\right| = q^d |\alpha - r| \left|Q'_\alpha(s)\right| \leq \left(\sup_{s\in[\alpha-1,\alpha+1]} \left|Q'_\alpha(s)\right|\right) q^d |\alpha - r|$$

donc $C_1 = \sup\limits_{s\in[\alpha-1,\alpha+1]} \left|Q'_\alpha(s)\right| > 0$ et $|\alpha - r| \geq \dfrac{1}{C_1}\dfrac{1}{q^d}$;
soit $|\alpha - r| > 1$ et dans ce cas $|\alpha - r| > \dfrac{1}{q^d}$.

En posant $C_\alpha = \min\left(1, \dfrac{1}{C_1}\right)$, on a alors $\left|\alpha - \dfrac{p}{q}\right| \geq \dfrac{C_\alpha}{q^d}$. ■

Les nombres de Liouville sont définis par :

$$\xi = \sum_{n\geq 1} \frac{a_n}{10^{n!}}$$

où $(a_n)_{n\geq 1}$ est une suite d'entiers compris entre 0 et 9 avec $a_n \geq 1$ pour $n \geq n_0$. En notant pour $k \geq n_0$ donné :

$$p = 10^{k!} \sum_{n=1}^{k} \frac{a_n}{10^{n!}},\ q = 10^{k!}$$

on a :

$$0 < \left|\xi - \frac{p}{q}\right| = \sum_{n=k+1}^{+\infty} \frac{a_n}{10^{n!}} \leq \frac{9}{10^{(k+1)!}} \sum_{n=0}^{+\infty} \frac{1}{10^n} = \frac{9}{10^{(k+1)!}} \frac{10}{9} \leq \frac{1}{q^k}$$

Si ξ est algébrique de degré $d \geq 1$, alors :

$$\frac{C_\xi}{q^d} \leq \left|\xi - \frac{p}{q}\right| \leq \frac{1}{q^k}$$

pour tout $k \geq n_0$ et faisant tendre k vers l'infini on aboutit à une absurdité. En conclusion ξ est transcendant.

On peut utiliser ce résultat pour montrer qu'il y a autant de nombres transcendants que de réels. Pour ce faire on utilise l'application qui associe à $\xi = \sum\limits_{n\geq 1} \dfrac{a_n}{10^{n!}}$ le réel $\sum\limits_{n\geq 1} \dfrac{a_n}{10^n} \in]0,1[$ et on vérifie que c'est une bijection (si $x \in]0,1[$ est décimal on utilise son écriture décimale impropre comportant une infinité de 9).

4

Formules de Taylor

4.1 La formule de Taylor-Lagrange

Du théorème de Rolle on déduit le résultat suivant qui généralise le théorème des accroissements finis.

$[a,b]$ désigne un intervalle compact non réduit à un point.

Théorème 4.1 *Soit $f : [a,b] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^n$ qui est $n+1$ fois dérivable sur l'intervalle ouvert $]a,b[$.*
Il existe un point $c \in]a,b[$ tel que :

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{f^{(n+1)}(c)}{(n+1)!} (b-a)^{n+1}$$

Démonstration. On se ramène aux hypothèses du théorème de Rolle en considérant la fonction g définie sur $[a,b]$ par :

$$g(x) = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} (b-x)^k - \frac{\lambda}{(n+1)!} (b-x)^{n+1}$$

où la constante réelle λ est telle que $g(b) = g(a)$.

Le théorème de Rolle appliqué à la fonction g nous assure de l'existence d'un réel $c \in]a,b[$ tel que $g'(c) = 0$, ce qui équivaut à :

$$\sum_{k=1}^{n} \frac{f^{(k)}(c)}{(k-1)!} (b-c)^{k-1} - \sum_{k=0}^{n} \frac{f^{(k+1)}(c)}{k!} (b-c)^k + \frac{\lambda}{n!} (b-c)^n = 0$$

ou encore à $\lambda = f^{(n+1)}(c)$.

L'égalité $g(a) = g(b) = 0$ donne alors le résultat. ∎

Pour les fonctions à valeurs dans $\mathbb{R}^p$ ou dans un espace vectoriel normé E, on a le résultat suivant.

Théorème 4.2 (inégalité de Taylor-Lagrange) *Soit $f : [a,b] \to \mathbb{R}^p$ (ou plus généralement $f : [a,b] \to E$, où E est un espace vectoriel normé) de classe $\mathcal{C}^n$ qui est $n+1$ fois dérivable sur l'intervalle ouvert $]a,b[$, la fonction $f^{(n+1)}$ étant majorée par une constante M.*
On a :

$$\left\| f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} (b-a)^k \right\| \leq \frac{M}{(n+1)!} (b-a)^{n+1}$$

Démonstration. Dans le cas où $p = 1$ le résultat est une conséquence immédiate du théorème précédent.

En désignant par g, h les fonctions définies sur $[a, b]$ respectivement par :

$$g(x) = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(b-x)^k\,,\ h(x) = -\frac{M}{(n+1)!}(b-x)^{n+1}$$

on a :

$$\|g'(x)\| = \left\|\frac{f^{(n+1)}(x)}{n!}(b-x)^n\right\| \leq h'(x) = \frac{M}{n!}(b-x)^n$$

et l'inégalité des accroissements finis généralisée nous donne :

$$\begin{aligned}\|g(b) - g(a)\| &= \left\|f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k\right\| \\ &\leq h(b) - h(a) = \frac{M}{(n+1)!}(b-a)^{n+1}\end{aligned}$$

■

Exercice 4.1 *Soient n un entier naturel non nul et $(\alpha_k)_{0\leq k\leq n}$ une suite de réels deux à deux distincts.*
Montrer que pour tout polynôme $P \in \mathbb{R}_n[x]$ de degré n, la famille de polynômes :

$$\{P_k \mid P_k(X) = P(X + \alpha_k)\,,\ 0 \leq k \leq n\}$$

est une base de $\mathbb{R}_n[X]$.

Solution 4.1 *Pour P de degré n, la famille $\left(P^{(j)}\right)_{0\leq j\leq n}$ est échelonnée en degré et il est facile de vérifier que c'est une base de $\mathbb{R}_n[x]$.*
En effet, si $P(X) = \sum\limits_{k=0}^{n} a_k X^k$ alors la matrice de ce système dans la base canonique de $\mathbb{R}_n[x]$ est triangulaire supérieure avec les $\dfrac{n!}{(n-j)!}a_n$ pour éléments diagonaux.

En utilisant la formule de Taylor-Lagrange, on peut écrire que :

$$P_k(x) = P(x + \alpha_k) = \sum_{j=0}^{n} \frac{\alpha_k^j}{j!} P^{(j)}(x)$$

C'est à dire que la matrice du système $(P_k)_{0\leq k\leq n}$ dans la base $\left(\dfrac{1}{j!}P^{(j)}\right)_{0\leq j\leq n}$ est la matrice de Vandermonde $A_n = \left(\left(\alpha_k^j\right)\right)_{0\leq j,k\leq n}$.
Dans le cas particulier où les α_i sont deux à deux distincts cette matrice est inversible ($\det(A_n) = \prod\limits_{0\leq i<k\leq n}(\alpha_k - \alpha_i) \neq 0$) et en conséquence $(P_k)_{0\leq k\leq n}$ est une base de $\mathbb{R}_n[x]$.

Exercice 4.2 *Soit f une fonction deux fois dérivable de $\mathbb{R}$ dans $\mathbb{R}$ telle que $\lim\limits_{|x|\to+\infty} \dfrac{f(x)}{x} = 0$.*
Montrer qu'il existe un réel c tel que $f''(c) = 0$.

Solution 4.2 *Si $f' = 0$, f est alors constante et le résultat est trivial.*
On suppose donc que $f' \neq 0$ et on se donne un réel a tel que $f'(a) \neq 0$.
Supposons $f'(a) > 0$.
Si $f''(a) = 0$ c'est alors terminé.
On suppose que $f''(a) \neq 0$.
En utilisant la formule de Taylor-Lagrange à l'ordre 2, on peut écrire pour tout réel $x \neq a$:

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(c_x)}{2}(x-a)^2$$

où c_x est un réel compris entre a et x.
De $\lim\limits_{|x| \to +\infty} \frac{f(x)}{x} = 0$, on déduit que :

$$\lim_{x \to +\infty} \frac{f''(c_x)}{2} \frac{(x-a)^2}{x} = -f'(a)$$

ou encore que $\lim\limits_{x \to +\infty} \frac{f''(c_x)}{2}(x-a) = -f'(a)$.
On peut donc trouver un réel c_1 tel que $f''(c_1) < 0$.
De même avec $\lim\limits_{x \to -\infty} \frac{f(x)}{x} = 0$, on déduit que $\lim\limits_{x \to -\infty} \frac{f''(c_x)}{2}(x-a) = -f'(a)$ et il existe un réel c_2 tel que $f''(c_2) > 0$.
Le théorème de Darboux nous assure alors l'existence d'un réel c compris entre c_1 et c_2 tel que $f''(c) = 0$.

Exercice 4.3 *Soit $f : \mathbb{R}^+ \to \mathbb{R}$ une fonction deux fois dérivable de dérivée seconde bornée.*
Montrer que si $\lim\limits_{x \to +\infty} f(x) = 0$, on a alors $\lim\limits_{x \to +\infty} f'(x) = 0$.

Solution 4.3 *On note $M = \sup\limits_{x \in \mathbb{R}^+} |f''(x)|$.*
La formule de Taylor-Lagrange nous permet d'écrire pour $x > 0$ et $h > 0$, on a :

$$f(x+h) = f(x) + hf'(x) + \frac{f''(x+\theta_h h)}{2} h^2$$

avec $0 < \theta_h < 1$ et on en déduit que :

$$|f'(x)| \leq \left| \frac{f(x+h) - f(x)}{h} \right| + M\frac{h}{2}$$

Comme $\lim\limits_{x \to +\infty} f(x) = 0$, on peut trouver, pour tout réel $\varepsilon > 0$, un réel $a > 0$ tel que $|f(x)| < \varepsilon$ pour tout $x > a$, donc :

$$\forall x > a,\ \forall h > 0,\ |f'(x)| \leq \varphi(h) = 2\frac{\varepsilon}{h} + M\frac{h}{2}$$

L'étude des variations de la fonction φ sur $]0, +\infty[$ nous montre que cette fonction atteint son minimum en $2\sqrt{\frac{\varepsilon}{M}}$ et que ce minimum vaut $2\sqrt{\varepsilon M}$.
On déduit alors de l'inégalité précédente que :

$$\forall x > a,\ |f'(x)| \leq 2\sqrt{M}\sqrt{\varepsilon}$$

On a donc ainsi montré que $\lim\limits_{x \to +\infty} f'(x) = 0$.

4.2 Le théorème de Taylor-Young

On rappelle qu'une fonction f définie dans un voisinage du réel a admet une dérivée d'ordre $n \geq 2$ en a si cette fonction admet des dérivées jusqu'a l'ordre $n-1$ dans un voisinage de a et si $f^{(n-1)}$ est dérivable en a.

On suppose ici que f est une fonction à valeurs réelles définie sur un intervalle réel I non réduit à un point et que a est un point intérieur à I.

Définition 4.1 *On dit que f admet un développement limité à l'ordre n en a s'il existe une fonction polynomiale P dans $\mathbb{R}_n[x]$ et une fonction ε définie sur I à valeurs réelles telles que :*

$$\begin{cases} \forall x \in I,\ f(x) = P(x-a) + (x-a)^n \varepsilon(x) \\ \lim\limits_{x \to a} \varepsilon(x) = 0 \end{cases}$$

En utilisant la notation de Landau, on peut écrire :

$$f(x) = P(x-a) + o((x-a)^n)$$

Théorème 4.3 *Si f admet un développement limité à l'ordre n en a alors ce dernier est unique.*

Démonstration. Il suffit de remarquer que si R dans $\mathbb{R}_n[x]$ est tel que $R(x-a) = o((x-a)^n)$, on a alors $R = 0$. ■

Théorème 4.4 (Taylor-Young) *Si f est dérivable à l'ordre $n \geq 1$ en a, elle admet alors, au voisinage de a, le développement limité d'ordre n :*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k + o((x-a)^n)$$

Démonstration. En définissant la fonction g sur I par :

$$g(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

il s'agit de montrer que $\lim\limits_{x \to a} \dfrac{g(x)}{(x-a)^n} = 0$.

Pour ce faire on procède par récurrence sur $n \geq 1$.

Pour $n = 1$ le résultat découle de la définition du nombre dérivé $f'(a)$.

Supposons le résultat acquis pour $n - 1 \geq 1$.

La fonction g est dérivable dans un voisinage J de a avec :

$$g'(x) = f'(x) - \sum_{k=0}^{n-1} \frac{(f')^{(k)}(a)}{k!}(x-a)^k$$

et l'hypothèse de récurrence appliquée à f' nous dit que $\lim\limits_{x \to a} \dfrac{g'(x)}{(x-a)^{n-1}} = 0$, c'est-à-dire que pour tout réel $\varepsilon > 0$ il existe un réel $\eta > 0$ tel que $[a-\eta, a+\eta]$ soit contenu dans J et :

$$|x-a| < \eta \ \Rightarrow\ |g'(x)| \leq \varepsilon |x-a|^{n-1}$$

On a alors :

$$\begin{cases} \forall x \in [a-\eta, a], \ -\varepsilon (a-x)^{n-1} \leq g'(x) \leq \varepsilon (a-x)^{n-1} \\ \forall x \in [a, a+\eta], \ -\varepsilon (x-a)^{n-1} \leq g'(x) \leq \varepsilon (x-a)^{n-1} \end{cases}$$

et il en résulte que :

$$\begin{cases} \forall x \in [a-\eta, a], \ -\dfrac{\varepsilon}{n} (a-x)^{n} \leq g(a) - g(x) \leq \dfrac{\varepsilon}{n} (a-x)^{n} \\ \forall x \in [a, a+\eta], \ -\dfrac{\varepsilon}{n} (x-a)^{n} \leq g(x) - g(a) \leq \dfrac{\varepsilon}{n} (x-a)^{n} \end{cases}$$

(on utilise le fait que si $u' \leq v'$ sur $[\alpha, \beta]$, alors $u(\beta) - u(\alpha) \leq v(\alpha) - v(\beta)$), c'est-à-dire en définitive que :

$$\forall x \in [a-\eta, a+\eta], \ |g(x)| \leq \frac{\varepsilon}{n} |x-a|^n$$

ce qui signifie bien que $\lim\limits_{x \to a} \dfrac{g(x)}{(x-a)^n} = 0.$ ∎

Ce résultat permet d'obtenir les développements limités classiques en 0 :

$$\begin{aligned} & e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + o(x^n) \\ & \cos(x) = \sum_{k=0}^{n} \frac{(-1)^k x^{2k}}{(2k)!} + o(x^{2n+1}) \\ & \sin(x) = \sum_{k=0}^{n} \frac{(-1)^k x^{2k+1}}{(2k+1)!} + o(x^{2n+2}) \\ & (1+x)^\alpha = 1 + \sum_{k=1}^{n} \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!} x^k + o(x^n) \ \ (\alpha \in \mathbb{R} \setminus \mathbb{N}) \end{aligned}$$

Exercice 4.4 *Soient f une fonction de classe $\mathcal{C}^{n+1}$ à valeurs réelles définie sur un intervalle ouvert I et a est un point de I.*
Pour tout réel $h \in]0, b-a[$, on désigne par θ_h un réel dans $]0,1[$ tel que :

$$f(a+h) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} h^k + \frac{h^{n+1}}{(n+1)!} f^{(n+1)}(a + \theta_h h)$$

Montrer que si f est dérivable à l'ordre $n+2$ en a avec $f^{(n+2)}(a)$ non nul, on a alors :

$$\lim_{h \to 0} \theta_h = \frac{1}{n+2}$$

Solution 4.4 *En utilisant la formule de Taylor-Young, on a :*

$$f(a+h) = \sum_{k=0}^{n+2} \frac{f^{(k)}(a)}{k!} h^k + o\left(h^{n+2}\right)$$

et avec la formule de Taylor-Lagrange, on déduit par soustraction que :

$$\left(f^{(n+1)}(a+\theta_h h) - f^{(n+1)}(a)\right) \frac{h^{n+1}}{(n+1)!} = \frac{f^{(n+2)}(a)}{(n+2)!} h^{n+2} + o\left(h^{n+2}\right)$$

soit :

$$\theta_h \frac{f^{(n+1)}(a+\theta_h h) - f^{(n+1)}(a)}{\theta_h h} = \frac{f^{(n+2)}(a)}{n+2} + o(1)$$

et avec $f^{(n+2)}(a) \neq 0$, on déduit que pour h voisin de 0, on a :

$$\theta_h = \frac{\frac{f^{(n+2)}(a)}{n+2} + o(1)}{\frac{f^{(n+1)}(a+\theta_h h) - f^{(n+1)}(a)}{\theta_h h}} \underset{h \to 0}{\to} \frac{1}{n+2}$$

4.3 Formule de Taylor avec reste intégral

$[a, b]$ désigne un intervalle compact non réduit à un point et n est un entier naturel.

La formule de Taylor avec reste intégral qui suit permet d'obtenir des résultats plus fins que la formule de Taylor-Lagrange.

Cette formule nécessite une hypothèse supplémentaire de continuité de la dernière dérivée et est valable pour les fonctions à valeurs dans un espace de Banach.

Théorème 4.5 *Si f est une fonction à valeurs réelles (ou dans un espace de Banach) définie et de classe $\mathcal{C}^{n+1}$ sur $[a, b]$, on a alors :*

$$f(b) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k + \int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n \, dt.$$

Démonstration. La fonction g définie sur $[a, b]$ par :

$$g(x) = f(b) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(b-x)^k$$

étant de classe $\mathcal{C}^1$ sur $[a, b]$ on peut écrire :

$$g(b) - g(a) = \int_a^b g'(t) \, dt$$

soit :

$$-f(b) + \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(b-a)^k = -\int_a^b \frac{f^{(n+1)}(t)}{n!}(b-t)^n \, dt$$

■

Cette formule permet de retrouver l'inégalité de Taylor-Lagrange pour f de classe $\mathcal{C}^{n+1}$ sur $[a, b]$ et à valeurs dans un espace de Banach.

Exercice 4.5

1. *Soient $f : [-1, 1] \to \mathbb{R}$ une fonction paire de classe $\mathcal{C}^2$ et φ la fonction définie sur $[0, 1]$ par :*

$$\forall x \in [0, 1], \ \varphi(x) = \int_0^x f(t) \, dt - x f(x)$$

   (a) *Montrer que :*

$$\forall x \in [0, 1], \ \varphi(x) = \int_0^x \frac{t^2 - x^2}{2} f''(t) \, dt$$

   (b) *Montrer que pour tout x dans $]0, 1]$ il existe un réel c_x dans $[0, 1]$ tel que :*

$$\varphi(x) = -\frac{f''(c_x)}{3} x^3$$

2. *Soit $f : [a, b] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$.*
   Déduire de ce qui précède une expression de l'erreur de quadrature dans la méthode du trapèze.

Solution 4.5

1.

(a) *La fonction f étant paire, sa dérivée f' est impaire et la formule de Taylor avec reste intégrale à l'ordre 1 nous permet d'écrire :*

$$f(x) = f(0) + f'(0)x + \int_0^x (x-t) f''(t)\,dt = f(0) + \int_0^x (x-t) f''(t)\,dt$$

De même, la fonction $F : x \mapsto \int_0^x f(t)\,dt$ étant impaire, on a $F(0) = 0$, $F''(0) = f'(0) = 0$ et la formule de Taylor avec reste intégral à l'ordre 2 nous donne :

$$F(x) = F'(0)x + \int_0^x \frac{(x-t)^2}{2} F'''(t)\,dt = f(0)x + \int_0^x \frac{(x-t)^2}{2} f''(t)\,dt$$

On a donc, pour tout $x \in [0,1]$:

$$\varphi(x) = F(x) - xf(x) = \int_0^x \frac{t^2 - x^2}{2} f''(t)\,dt$$

(b) *Pour tout $x \in [0,1]$, la fonction $t \mapsto t^2 - x^2$ étant à valeurs négatives sur $[0,x]$, le théorème de la moyenne nous dit qu'il existe $c_x \in [0,x]$ tel que :*

$$\varphi(x) = f''(c_x) \int_0^x \frac{t^2 - x^2}{2}\,dt = -\frac{f''(c_x)}{3} x^3$$

2. Le changement de variable $x = \frac{a+b}{2} + \frac{b-a}{2} t$ avec $t \in [-1,1]$ nous ramène à l'intervalle $[-1,1]$.
En définissant la fonction g sur $[-1,1]$ par :

$$g(t) = f\left(\frac{a+b}{2} + \frac{b-a}{2} t\right)$$

l'erreur dans la méthode du trapèze sur $[a,b]$ s'écrit :

$$\begin{aligned} E(f) &= \int_a^b f(x)\,dx - \frac{b-a}{2}(f(a) + f(b)) \\ &= \frac{b-a}{2}\left(\int_{-1}^1 g(t)\,dt - (g(-1) + g(1))\right) \end{aligned}$$

En désignant par h la fonction définie sur $[-1,1]$ par $h(x) = g(-x) + g(x)$, on a :

$$E(f) = \frac{b-a}{2}\left(\int_0^1 h(x)\,dx - h(1)\right)$$

la fonction h étant paire et de classe $\mathcal{C}^2$ sur $[-1,1]$.
Le résultat de la question précédente nous dit alors qu'il existe un réel c dans $[0,1]$ tel que :

$$E(f) = -\frac{b-a}{2}\frac{h''(c)}{3}$$

En désignant par m [resp. M] la borne inférieure [resp. supérieure] de g'' sur $[-1,1]$, on a :

$$2m \leq h''(c) = g''(-c) + g''(c) \leq 2M$$

et le théorème des valeurs intermédiaires nous dit alors qu'il existe d dans $[0,1]$ tel que $h''(c) = 2g''(d)$. On a donc :

$$E(f) = -\frac{b-a}{3} g''(d) = -\frac{(b-a)^3}{12} f''(d)$$

Exercice 4.6

1. *Soient $f : [-1,1] \to \mathbb{R}$ une fonction paire de classe $\mathcal{C}^4$ et φ la fonction définie sur $[0,1]$ par :*

$$\forall x \in [0,1], \ \varphi(x) = \int_0^x f(t)\,dt - \frac{x}{3}(f(x) + 2f(0))$$

   (a) *Montrer que :*

$$\forall x \in [0,1], \ \varphi(x) = -\int_0^x \frac{(x+3t)(x-t)^3}{72} f^{(4)}(t)\,dt$$

   (b) *Montrer que pour tout x dans $]0,1]$ il existe un réel c_x dans $[0,1]$ tel que :*

$$\varphi(x) = -\frac{f^{(4)}(c_x)}{180} x^5$$

2. *Soit $f : [a,b] \to \mathbb{R}$ une fonction de classe $\mathcal{C}^4$.*
Déduire de ce qui précède une expression de l'erreur de quadrature dans la méthode de Simpson.

Solution 4.6

1.

   (a) *En tenant compte de la parité de la fonction f, la formule de Taylor avec reste intégrale à l'ordre 3 s'écrit :*

$$f(x) = f(0) + \frac{x^2}{2} f''(0)\,x + \int_0^x \frac{(x-t)^3}{6} f^{(4)}(t)\,dt$$

   De même, la fonction $F : x \mapsto \displaystyle\int_0^x f(t)\,dt$ étant impaire, la formule de Taylor avec reste intégrale à l'ordre 4 nous donne :

$$\begin{aligned} F(x) &= F'(0)\,x + \frac{F^{(3)}(0)}{6} x^3 + \int_0^x \frac{(x-t)^4}{24} F^{(5)}(t)\,dt \\ &= f(0)\,x + \frac{f''(0)}{6} x^3 + \int_0^x \frac{(x-t)^4}{24} f^{(4)}(t)\,dt \end{aligned}$$

   On a donc, pour tout $x \in [0,1]$:

$$\begin{aligned} \varphi(x) &= F(x) - \frac{x}{3}(f(x) + 2f(0)) \\ &= -\int_0^x \frac{(x+3t)(x-t)^3}{72} f^{(4)}(t)\,dt \end{aligned}$$

(b) *Pour tout* $x \in [0,1]$, *la fonction* $t \mapsto (x+3t)(x-t)^3$ *étant à valeurs positives sur* $[0,x]$, *le théorème de la moyenne nous dit qu'il existe* $c_x \in [0,x]$ *tel que :*

$$\varphi(x) = -f^{(4)}(c_x) \int_0^x \frac{(x+3t)(x-t)^3}{72} dt = -\frac{f^{(4)}(c_x)}{180} x^5$$

2. *Le changement de variable* $x = \frac{a+b}{2} + \frac{b-a}{2} t$ *avec* $t \in [-1,1]$ *nous ramène à l'intervalle* $[-1,1]$.
En définissant la fonction g *sur* $[-1,1]$ *par :*

$$g(t) = f\left(\frac{a+b}{2} + \frac{b-a}{2} t\right)$$

l'erreur dans la méthode de Simpson sur $[a,b]$ *s'écrit :*

$$\begin{aligned} E(f) &= \int_a^b f(x)\,dx - \frac{b-a}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right) \\ &= \frac{b-a}{2}\left(\int_{-1}^1 g(t)\,dt - \frac{1}{3}(g(-1) + 4g(0) + g(1))\right) \end{aligned}$$

En désignant par h *la fonction définie sur* $[-1,1]$ *par* $h(x) = g(-x) + g(x)$, *on a :*

$$E(f) = \frac{b-a}{2}\left(\int_0^1 h(x)\,dx - \frac{1}{3}(h(1) + 2h(0))\right)$$

la fonction h *étant paire et de classe* $\mathcal{C}^4$ *sur* $[-1,1]$. *Le résultat de la question précédente nous dit alors qu'il existe un réel* c *dans* $[0,1]$ *tel que :*

$$E(f) = -\frac{b-a}{2} \frac{h^{(4)}(c)}{180}$$

En désignant par m *[resp.* M*] la borne inférieure [resp. supérieure] de* $g^{(4)}$ *sur* $[-1,1]$, *on a :*

$$2m \leq h^{(4)}(c) = g^{(4)}(-c) + g^{(4)}(c) \leq 2M$$

et le théorème des valeurs intermédiaires nous dit alors qu'il existe d *dans* $[0,1]$ *tel que* $h^{(4)}(c) = 2g^{(4)}(d)$. *On a donc :*

$$E(f) = -(b-a) \frac{g^{(4)}(d)}{180} = -\frac{(b-a)^5}{2880} f^{(4)}(d)$$

4.4 Quelques applications de la formule de Taylor-Lagrange

4.4.1 Problèmes d'extrémums

Si une fonction dérivable $f : I \to \mathbb{R}$, où I est un intervalle ouvert, admet un extrémum local en un point a intérieur au domaine de définition, on a alors $f'(a) = 0$, la réciproque étant fausse comme le montre l'exemple de la fonction x^3 au voisinage de 0.

L'utilisation de la formule de Taylor-Lagrange permet de donner une condition nécessaire et suffisante d'extrémum.

I est un intervalle ouvert non vide.

Théorème 4.6 *Soient $f : I \to \mathbb{R}$ de classe $\mathcal{C}^n$ avec $n \geq 2$ et a dans I tel que $f^{(k)}(a) = 0$ pour tout k compris entre 1 et $n-1$ et $f^{(n)}(a) \neq 0$.*
La fonction f admet un maximum [resp. minimum] local en a si, et seulement si, n est pair et $f^{(n)}(a) < 0$ [resp. $f^{(n)}(a) > 0$].

Démonstration. Soit $\eta > 0$ tel que $]a-\eta, a+\eta[\subset I$.
Pour tout réel h tel que $|h| < \eta$ il existe un réel $\theta \in]0,1[$ tel que :

$$\begin{aligned} f(a+h) &= f(a) + \sum_{k=1}^{n-1} \frac{f^{(k)}(a)}{k!} h^k + \frac{f^{(n)}(a+\theta h)}{n!} h^n \\ &= f(a) + \frac{f^{(n)}(a+\theta h)}{n!} h^n \end{aligned}$$

Si $f^{(n)}(a) < 0$, du fait de la continuité de $f^{(n)}$, on peut choisir η assez petit de sorte que $f^{(n)}(x) < 0$ pour tout $x \in]a-\eta, a+\eta[$ et $f(a+h) - f(a)$ et de signe contraire à celui de h^n.
Il en résulte que pour n pair f admet un maximum local en a et pour n impair on a $f(a+h) > f(a)$ si $h < 0$, $f(a+h) < f(a)$ si $h > 0$, ce qui signifie qu'il n'y a pas d'extrémum local en a.
Pour $f^{(n)}(a) > 0$, le raisonnement est analogue. ■

Dans le cas $n = 2$, si $f'(a) = 0$ et $f''(a) \neq 0$, on a donc :

$$\begin{aligned} f''(a) < 0 &\Rightarrow \text{ maximum local en } a \\ f''(a) > 0 &\Rightarrow \text{ minimum local en } a \end{aligned}$$

4.4.2 Inégalités

En utilisant l'inégalité de Taylor-Lagrange, on obtient facilement les inégalités classiques suivantes :

$$\begin{cases} \forall x \in \mathbb{R}, \ |\sin(x) - x| \leq \dfrac{|x|^3}{3!} \\ \forall x \in \mathbb{R}, \ \left|\sin(x) - x + \dfrac{x^3}{3!}\right| \leq \dfrac{|x|^5}{5!} \\ \forall x \in \mathbb{R}, \ |\cos(x) - 1| \leq \dfrac{x^2}{2} \\ \forall x \in \mathbb{R}^+, \ |e^x - 1 - x| \leq \dfrac{x^2}{2} e^x \\ \forall x \in \mathbb{R}^-, \ |e^x - 1 - x| \leq \dfrac{x^2}{2} \end{cases}$$

4.4.3 Développements en série entières

Pour simplifier, on se place au voisinage de 0.

Si f est une fonction de classe $\mathcal{C}^\infty$ au voisinage de 0, la série entière $\sum\limits_{n=0}^{+\infty} \dfrac{f^{(n)}(0)}{n!} x^n$ peut être de rayon de convergence nul ou converger vers une autre fonction que la fonction f.

Par exemple, la fonction f définie par $f(0) = 0$ et $f(x) = e^{-\frac{1}{x^2}}$ pour $x \neq 0$ est indéfiniment dérivable sur $\mathbb{R}$ avec toutes ses dérivées en 0 qui sont nulles et donc la série entière $\sum\limits_{n=0}^{+\infty} \dfrac{f^{(n)}(0)}{n!} x^n$ converge vers $g = 0 \neq f$.

On a toutefois le résultat classique suivant, où I est un intervalle ouvert centré en 0.

Théorème 4.7 *Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^\infty$.*
La fonction f est développable en série entière au voisinage de 0 si, et seulement si, il existe un réel $r > 0$ tel que $]-r, r[\subset I$ et la suite de fonctions $(R_n)_{n\in\mathbb{N}}$ définie par :

$$R_n(x) = f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

converge simplement vers 0 sur $]-r, r[$.
Dans ce cas, on a $f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$ pour tout $x \in]-r, r[$ et le rayon de convergence de cette série entière est supérieur ou égal à r.

Pour montrer que la suite $(R_n)_{n\in\mathbb{N}}$ des restes converge simplement vers 0, on peut utiliser l'expression de Lagrange du reste $R_n(x) = \frac{f^{(n+1)}(\theta x)}{(n+1)!} x^{n+1}$ avec $0 < \theta < 1$ ou sa représentation intégrale :

$$R_n(x) = \int_0^x \frac{f^{(n+1)}(t)}{n!} (x-t)^n \, dt = \frac{x^{n+1}}{n!} \int_0^1 \frac{f^{(n+1)}(\theta x)}{n!} (1-\theta)^n \, d\theta$$

Par exemple dans le cas de la fonction exponentielle réelle, on a pour tout réel x :

$$|R_n(x)| = \frac{e^{\theta x}}{(n+1)!} |x|^{n+1} \leq \frac{e^{|x|}}{(n+1)!} |x|^{n+1} \underset{n\to+\infty}{\to} 0$$

Il en résulte que $e^x = \sum_{n=0}^{+\infty} \frac{1}{n!} x^n$ pour tout $x \in \mathbb{R}$ et le rayon de convergence de cette série est infini.

À partir de ce résultat on est amené à définir la fonction exponentielle complexe par $e^z = \sum_{n=0}^{+\infty} \frac{1}{n!} z^n$ pour tout $z \in \mathbb{C}$.

Ce exemple est un cas particulier du résultat suivant.

Corollaire 4.1 *Soit $f : I \to \mathbb{R}$ une fonction de classe $\mathcal{C}^\infty$.*
S'il existe un réel $r > 0$ tel que $]-r, r[\subset I$ et pour tout x dans $]-r, r[$ on peut trouver une constante M_x avec :

$$\forall n \in \mathbb{N}, \ \left| f^{(n)}(x) \right| \leq M_x$$

alors f est développable en série entière dans $]-r, r[$ avec $f(x) = \sum_{n=0}^{+\infty} \frac{f^{(n)}(0)}{n!} x^n$.

4.4.4 Majoration de l'erreur dans les méthodes de Lagrange et de Newton

Voir les paragraphe 8.4 et 8.5.

4.4.5 Inégalités de Kolmogorov

Le résultat qui suit utilise la formule de Taylor-Lagrange et le fait que sur l'espace vectoriel $\mathbb{R}_n[x]$ de dimension finie toutes les normes sont équivalentes.

Théorème 4.8 *Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^{n+1}$, où n est un entier naturel non nul, telle que les fonctions f et $f^{(n+1)}$ soient bornées, alors toutes les dérivées $f^{(k)}$, pour k compris entre 1 et n, sont également bornées sur $\mathbb{R}$.*

Démonstration. Pour $x \in \mathbb{R}$ on désigne par P_x le polynôme réel de degré au plus égal à n défini par :

$$P_x(t) = \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!} t^k$$

En utilisant la formule de Taylor-Lagrange on peut écrire :

$$P_x(t) = f(x+t) - \frac{f^{(n+1)}(x+\theta t)}{(n+1)!} t^{n+1}$$

où le réel θ est dans $]0,1[$.

Si les fonctions f et $f^{(n+1)}$ sont bornées sur $\mathbb{R}$, on en déduit alors qu'il existe une constante $M > 0$ telle que :

$$\forall (x,t) \in \mathbb{R} \times [0,1], \ |P_x(t)| \leq M \tag{4.1}$$

Sur $\mathbb{R}_n[t]$ on définit des normes en posant pour tout $P = \sum_{k=0}^{n} a_k t^k \in \mathbb{R}_n[t]$:

$$\begin{cases} \|P\|_\infty = \sup_{t \in [0,1]} |P(t)| \\ \|P\| = \max_{0 \leq k \leq n} |a_k| \end{cases}$$

L'espace vectoriel $\mathbb{R}_n[t]$ étant de dimension finie, toutes les normes y sont équivalentes, il existe donc une constante $\alpha > 0$ telle que $\|P\| \leq \alpha \|P\|_\infty$ pour tout $P \in \mathbb{R}_n[t]$.

De l'inégalité (4.1) on déduit alors que :

$$\forall x \in \mathbb{R}, \ \|P_x\| = \max_{0 \leq k \leq n} \left| \frac{f^{(k)}(x)}{k!} \right| \leq \alpha \|P_x\|_\infty \leq \alpha M$$

et :

$$\forall k \in \{1, \cdots, n\}, \ \forall x \in \mathbb{R}, \ \left| f^{(k)}(x) \right| \leq \alpha k! M,$$

c'est-à-dire que toutes les dérivées $f^{(k)}$, pour k compris entre 1 et n, sont bornées sur $\mathbb{R}$. ■

On peut donner une autre démonstration de ce résultat en utilisant des matrices de Vandermonde.

Pour tout réel x et tout entier p compris entre 1 et n, la formule de Taylor à l'ordre n sur l'intervalle $[x, x+p]$ s'écrit :

$$f(x+p) - f(x) - \frac{f^{(n+1)}(x+p\theta_p)}{(n+1)!} p^{n+1} = \sum_{k=1}^{n} \frac{f^{(k)}(x)}{k!} p^k$$

avec $0 < \theta_p < 1$, ce qui peut s'écrire matriciellement $V(x) = AU(x)$, où on a noté :

$$V(x) = \begin{pmatrix} f(x+1) - f(x) - \dfrac{f^{(n+1)}(x+\theta_1)}{(n+1)!} \\ \vdots \\ f(x+n) - f(x) - \dfrac{f^{(n+1)}(x+n\theta_n)}{(n+1)!} n^{n+1} \end{pmatrix}, \ U(x) = \begin{pmatrix} \dfrac{f^{(1)}(x)}{1!} \\ \vdots \\ \dfrac{f^{(n)}(x)}{n!} \end{pmatrix}$$

et :

$$A = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2^2 & \cdots & 2^n \\ \vdots & \vdots & \ddots & \vdots \\ n & n^2 & \cdots & n^n \end{pmatrix}$$

On a alors :

$$\det(A) = n! \begin{vmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n & \cdots & n^{n-1} \end{vmatrix} = n! \prod_{1 \leq 1 < j \leq n} (i-j) \neq 0$$

(déterminant de Vandermonde), c'est-à-dire que la matrice A est inversible.

On peut donc écrire que $U(x) = A^{-1}V(x)$ et $\|U(x)\|_\infty \leq \|A^{-1}\|_\infty \|V(x)\|_\infty$, soit :

$$\begin{aligned} \max_{1\leq k\leq n} \left| \frac{f^{(k)}(x)}{k!} \right| &\leq \left\|A^{-1}\right\|_\infty \max_{1\leq k\leq n} \left| f(x+p) - f(x) - \frac{f^{(n+1)}(x+p\theta_p)}{(n+1)!} p^{n+1} \right| \\ &\leq \left\|A^{-1}\right\|_\infty \left(2\|f\|_\infty + \frac{n^{n+1}}{(n+1)!} \left\|f^{(n+1)}\right\|_\infty \right) \end{aligned}$$

le réel x étant quelconque.

On a donc pour tout entier k compris entre 1 et n :

$$\left\|f^{(k)}\right\|_\infty \leq k! \left\|A^{-1}\right\|_\infty \left(2\|f\|_\infty + \frac{n^{n+1}}{(n+1)!} \left\|f^{(n+1)}\right\|_\infty \right)$$

Les inégalités de Kolmogorov qui suivent nous donnent une majoration plus précise.

Dans le cas $n = 1$ on a le résultat suivant.

Lemme 4.1 *Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^2$ telle que les fonctions f et f'' soient bornées sur $\mathbb{R}$, la fonction f' est alors bornée sur $\mathbb{R}$ et on a :*

$$\|f'\|_\infty \leq \sqrt{2\|f\|_\infty \|f''\|_\infty}$$

Démonstration. Si $\|f''\|_\infty = 0$, la fonction f est alors affine, mais étant bornée elle est nécessairement constante, donc $f' = 0$ et l'inégalité est triviale.

On suppose donc que $\|f''\|_\infty > 0$.

La formule de Taylor-Lagrange nous permet d'écrire pour $x \in \mathbb{R}$ et $h > 0$:

$$\begin{cases} f(x+h) = f(x) + hf'(x) + \dfrac{f''(x+\theta_1 h)}{2} h^2 \\ f(x-h) = f(x) - hf'(x) + \dfrac{f''(x+\theta_2 h)}{2} h^2 \end{cases}$$

et en soustrayant ces deux égalités :

$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} + \frac{h}{4}\left(f''(x+\theta_2 h) - f''(x+\theta_1 h)\right)$$

Il en résulte que :

$$|f'(x)| \leq \frac{\|f\|_\infty}{h} + \frac{h\|f''\|_\infty}{2}$$

et :

$$\forall h > 0,\ \|f'\|_\infty \leq \varphi(h) = \frac{\|f\|_\infty}{h} + \frac{h\,\|f''\|_\infty}{2}$$

L'étude des variations de la fonction φ nous montre que cette fonction atteint son minimum en $\sqrt{\dfrac{2\,\|f\|_\infty}{\|f''\|_\infty}}$ et que ce minimum vaut $\sqrt{2\,\|f\|_\infty\,\|f''\|_\infty}$.

D'où le résultat. ■

Théorème 4.9 (Kolmogorov) *Si f est une fonction de classe $\mathcal{C}^{n+1}$ de $\mathbb{R}$ dans $\mathbb{R}$, où n est un entier naturel non nul, telle que f et $f^{(n+1)}$ soient bornées sur $\mathbb{R}$, toutes les dérivées $f^{(k)}$, pour k compris entre 1 et n, sont alors bornées sur $\mathbb{R}$ et on a :*

$$\left\|f^{(k)}\right\|_\infty \leq 2^{\frac{k(n+1-k)}{2}}\,\|f\|_\infty^{1-\frac{k}{n+1}}\,\left\|f^{(n+1)}\right\|_\infty^{\frac{k}{n+1}} \tag{4.2}$$

Démonstration. On procède par récurrence sur $n \geq 1$.

Pour $n = 1$ c'est le lemme précédent.

On supposant le résultat acquis pour $n - 1 \geq 1$ et en appliquant le lemme à $f^{(n-1)}$, on a :

$$\left\|f^{(n)}\right\|_\infty \leq \sqrt{2\,\|f^{(n-1)}\|_\infty\,\|f^{(n+1)}\|_\infty}$$

En utilisant l'hypothèse de récurrence, on a :

$$\left\|f^{(n-1)}\right\|_\infty \leq 2^{\frac{n-1}{2}}\,\|f\|_\infty^{\frac{1}{n}}\,\left\|f^{(n)}\right\|_\infty^{1-\frac{1}{n}}$$

ce qui donne avec l'inégalité précédente :

$$\left\|f^{(n)}\right\|_\infty^2 \leq 2\left\|f^{(n-1)}\right\|_\infty\left\|f^{(n+1)}\right\|_\infty \leq 2^{\frac{n+1}{2}}\,\|f\|_\infty^{\frac{1}{n}}\,\left\|f^{(n)}\right\|_\infty^{1-\frac{1}{n}}\,\left\|f^{(n+1)}\right\|_\infty$$

Si $\left\|f^{(n)}\right\|_\infty = 0$, f est alors un polynôme de degré au plus égal à n borné sur $\mathbb{R}$, c'est donc une constante et les inégalités de Kolmogorov sont trivialement vérifiées.

On suppose donc que $\left\|f^{(n)}\right\|_\infty > 0$ et on déduit de l'inégalité précédente que :

$$\left\|f^{(n)}\right\|_\infty^{1+\frac{1}{n}} \leq 2^{\frac{n+1}{2}}\,\|f\|_\infty^{\frac{1}{n}}\,\left\|f^{(n+1)}\right\|_\infty$$

soit :

$$\left\|f^{(n)}\right\|_\infty \leq 2^{\frac{n}{2}}\,\|f\|_\infty^{\frac{1}{n+1}}\,\left\|f^{(n+1)}\right\|_\infty^{\frac{n}{n+1}}$$

c'est-à-dire l'inégalité (4.2) pour $k = n$.

En utilisant l'hypothèse de récurrence, pour k compris entre 1 et $n - 1$, on a :

$$\begin{aligned}\left\|f^{(k)}\right\|_\infty &\leq 2^{\frac{k(n-k)}{2}}\,\|f\|_\infty^{1-\frac{k}{n}}\,\left\|f^{(n)}\right\|_\infty^{\frac{k}{n}} \\ &\leq 2^{\frac{k(n-k)}{2}}\,\|f\|_\infty^{1-\frac{k}{n}}\,2^{\frac{k}{2}}\,\|f\|_\infty^{\frac{k}{n(n+1)}}\,\left\|f^{(n+1)}\right\|_\infty^{\frac{k}{n+1}}\end{aligned}$$

soit :

$$\left\|f^{(k)}\right\|_\infty \leq 2^{\frac{k(n+1-k)}{2}}\,\|f\|_\infty^{1-\frac{k}{n+1}}\,\left\|f^{(n+1)}\right\|_\infty^{\frac{k}{n+1}}$$

c'est-à-dire les inégalités de Kolmogorov. ■

4.4.6 Estimation de l'erreur dans la méthode des rectangles

Pour une fonction f suffisamment dérivable, on peut donner, en utilisant la formule de Taylor-Lagrange, un développement asymptotique de l'erreur d'approximation dans la méthode des rectangles à gauche.

Par exemple, pour $f \in \mathcal{C}^3([a,b], \mathbb{R})$ on a le résultat suivant, où on a noté :

$$R_n(f) = \frac{b-a}{n} \sum_{k=0}^{n-1} f(x_{n,k})$$

Lemme 4.2 *Pour toute fonction $f \in \mathcal{C}^3([a,b], \mathbb{R})$, en désignant par $F : x \mapsto \int_a^x f(t)\,dt$ la primitive de f nulle en a, il existe des réel c_1, c_2, c_3 dans $[a,b]$ tels que :*

$$\begin{aligned} F(b) &= f(a)(b-a) + \frac{b-a}{2}(f(b)-f(a)) - \frac{(b-a)^2}{12}(f'(b)-f'(a)) \\ &+ \frac{(b-a)^4}{24}\left(f^{(3)}(c_1) - f^{(3)}(c_2) + f^{(3)}(c_3)\right) \end{aligned}$$

Démonstration. La fonction $F : x \mapsto \int_a^x f(t)\,dt$ est de classe $\mathcal{C}^4$ sur $[a,b]$ et le théorème de Taylor-Lagrange nous dit qu'il existe un réel $c_1 \in]a,b[$ tel que :

$$\begin{aligned} F(b) &= \int_a^b f(x)\,dx \quad (4.3) \\ &= F(a) + F'(a)(b-a) + \frac{F''(a)}{2}(b-a)^2 + \frac{F^{(3)}(a)}{6}(b-a)^3 + \frac{F^{(4)}(c)}{24}(b-a)^4 \quad (4.4) \\ &= f(a)(b-a) + \frac{f'(a)}{2}(b-a)^2 + \frac{f''(a)}{6}(b-a)^3 + \frac{f^{(3)}(c_1)}{24}(b-a)^4 \end{aligned}$$

On a aussi, avec la formule de Taylor-Lagrange pour f :

$$f(b) = f(a) + f'(a)(b-a) + \frac{f''(a)}{2}(b-a)^2 + \frac{f^{(3)}(c_2)}{6}(b-a)^3$$

ce qui donne :

$$\frac{f'(a)}{2}(b-a)^2 = \frac{b-a}{2}(f(b)-f(a)) - \frac{f''(a)}{4}(b-a)^3 - \frac{f^{(3)}(c_2)}{12}(b-a)^4$$

et en reportant dans (4.3), on obtient :

$$\begin{aligned} F(b) &= f(a)(b-a) + \frac{b-a}{2}(f(b)-f(a)) - \frac{f''(a)}{4}(b-a)^3 - \frac{f^{(3)}(c_2)}{12}(b-a)^4 \\ &+ \frac{f''(a)}{6}(b-a)^3 + \frac{f^{(3)}(c_1)}{24}(b-a)^4 \\ &= f(a)(b-a) + \frac{b-a}{2}(f(b)-f(a)) - \frac{f''(a)}{12}(b-a)^3 + \frac{(b-a)^4}{24}\left(f^{(3)}(c_1) - f^{(3)}(c_2)\right) \end{aligned}$$

De même, avec :

$$f'(b) = f'(a) + f''(a)(b-a) + \frac{f^{(3)}(c_3)}{2}(b-a)^2$$

on obtient :

$$\frac{f''(a)}{12}(b-a)^3 = \frac{(b-a)^2}{12}\left(f'(b)-f'(a)\right) - \frac{f^{(3)}(c_3)}{24}(b-a)^4$$

et :

$$\begin{aligned} F(b) &= f(a)(b-a) + \frac{b-a}{2}\left(f(b)-f(a)\right) - \frac{(b-a)^2}{12}\left(f'(b)-f'(a)\right) \\ &+ \frac{(b-a)^4}{24}\left(f^{(3)}(c_1) - f^{(3)}(c_2) + f^{(3)}(c_3)\right) \end{aligned}$$

■

Théorème 4.10 *Pour toute fonction $f \in \mathcal{C}^3([a,b], \mathbb{R})$ on a le développement asymptotique :*

$$R_n(f) = \int_a^b f(x)\,dx - \frac{b-a}{2n}\left(f(b)-f(a)\right) + \frac{(b-a)^2}{12n^2}\left(f'(b)-f'(a)\right) + O\left(\frac{1}{n^3}\right)$$

Démonstration. On a, pour tout k compris entre 0 et $n-1$:

$$\begin{aligned} \int_{x_{n,k}}^{x_{n,k+1}} f(x)\,dx &= f(x_{n,k})(x_{n,k+1}-x_{n,k}) + \frac{x_{n,k+1}-x_{n,k}}{2}\left(f(x_{n,k+1})-f(x_{n,k})\right) \\ &- \frac{(x_{n,k+1}-x_{n,k})^2}{12}\left(f'(x_{n,k+1})-f'(x_{n,k})\right) \\ &+ \frac{(x_{n,k+1}-x_{n,k})^4}{24}\left(f^{(3)}(c_{1,k}) - f^{(3)}(c_{2,k}) + f^{(3)}(c_{3,k})\right) \\ &= \frac{b-a}{n} f(x_{n,k}) + \frac{b-a}{2n}\left(f(x_{n,k+1})-f(x_{n,k})\right) \\ &- \frac{(b-a)^2}{12n^2}\left(f'(x_{n,k+1})-f'(x_{n,k})\right) \\ &+ \frac{(b-a)^4}{24n^4}\left(f^{(3)}(c_{1,k}) - f^{(3)}(c_{2,k}) + f^{(3)}(c_{3,k})\right) \end{aligned}$$

et :

$$\begin{aligned} \int_a^b f(x)\,dx &= \sum_{k=0}^{n-1} \int_{x_{n,k}}^{x_{n,k+1}} f(x)\,dx \\ &= \frac{b-a}{n}\sum_{k=0}^{n-1} f(x_{n,k}) + \frac{b-a}{2n}\sum_{k=0}^{n-1}\left(f(x_{n,k+1})-f(x_{n,k})\right) \\ &- \frac{(b-a)^2}{12n^2}\sum_{k=0}^{n-1}\left(f'(x_{n,k+1})-f'(x_{n,k})\right) \\ &+ \frac{(b-a)^4}{24n^4}\sum_{k=0}^{n-1}\left(f^{(3)}(c_{1,k}) - f^{(3)}(c_{2,k}) + f^{(3)}(c_{3,k})\right) \end{aligned}$$

soit en désignant par c_n, d_n et e_n des réels dans $[a,b]$ tels que $\sum_{k=0}^{n-1} f^{(3)}(c_{1,k}) = nf^{(3)}(c_n)$, $\sum_{k=0}^{n-1} f^{(3)}(c_{2,k}) = nf^{(3)}(d_n)$, et $\sum_{k=0}^{n-1} f^{(3)}(c_{3,k}) = nf^{(3)}(e_n)$:

$$\begin{aligned} \int_a^b f(x)\,dx &= R_n(f) + \frac{b-a}{2n}\left(f(b)-f(a)\right) - \frac{(b-a)^2}{12n^2}\left(f'(b)-f'(a)\right) \\ &+ \frac{(b-a)^4}{24n^3}\left(f^{(3)}(c_n) - f^{(3)}(d_n) + f^{(3)}(e_n)\right) \end{aligned}$$

avec :

$$\left|f^{(3)}(c_n) - f^{(3)}(d_n) + f^{(3)}(e_n)\right| \leq 3\left\|f^{(3)}\right\|_\infty$$

ce qui donne bien le résultat annoncé. ■

Exemple 4.1 *Pour* $f(t) = \dfrac{1}{1+t}$ *sur* $[0,1]$, *on obtient :*

$$\sum_{k=n}^{2n-1} \frac{1}{k} = \ln(2) + \frac{1}{4n} + \frac{1}{16n^2} + O\left(\frac{1}{n^3}\right)$$

Exemple 4.2 *Pour* $f(t) = \dfrac{1}{(1+t)^\alpha}$ *sur* $[0,1]$, *avec* $\alpha > 0$, $\alpha \neq 1$ *:*

$$n^{\alpha-1} \sum_{k=n}^{2n-1} \frac{1}{k^\alpha} = \frac{1}{1-\alpha}\left(\frac{1}{2^{\alpha-1}} - 1\right) - \frac{1}{2n}\left(\frac{1}{2^\alpha} - 1\right) + \frac{\alpha}{12n^2}\left(1 - \frac{1}{2^{\alpha+1}}\right) + O\left(\frac{1}{n^3}\right).$$

4.5 Une application de la formule de Taylor avec reste intégral. Un théorème de Bernstein

L'exemple de la fonction f définie par $f(0) = 0$ et $f(x) = e^{-\frac{1}{x^2}}$ pour $x \neq 0$ nous montre qu'une fonction indéfiniment dérivable sur $\mathbb{R}$ n'est pas nécessairement développable en série entière.

Le théorème de Bernstein qui suit nous dit qu'avec l'hypothèse supplémentaire de positivité des dérivées d'ordres pairs de la fonction f, on est assuré du développement en série entière.

Lemme 4.3 *Soit f une fonction à valeurs réelles de classe $\mathcal{C}^\infty$ sur $]-a,a[$ avec $a > 0$.*
Si f est paire et $f^{(2k)}(x) \geq 0$ pour tout entier naturel k et tout $x \in \,]-a,a[$, la fonction f est alors développable en série entière sur $]-a,a[$.

Démonstration. Pour tout entier naturel n et tout $x \in \,]-a,a[$ on note :

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k, \quad R_n(x) = f(x) - P_n(x)$$

Il s'agit alors de montrer que $\lim\limits_{n\to+\infty} R_n(x) = 0$ pour tout $x \in \,]-a,a[$.

Avec la parité de la fonction f on déduit que $f^{(2k+1)}(0) = 0$ pour tout $k \geq 0$ et avec la positivité des $f^{(2k)}$ que :

$$\forall x \in \,]-a,a[, \quad P_{2n}(x) = \sum_{k=0}^{n} \frac{f^{(2k)}(0)}{(2k)!} x^{2k} \geq 0, \quad R_{2n}(x) \leq f(x)$$

D'autre part, avec la formule de Taylor avec reste intégral on peut écrire pour tout $x \in \,]0,a[$:

$$R_{2n}(x) = \int_0^x \frac{(x-t)^{2n}}{(2n)!} f^{(2n+1)}(t)\,dt = \frac{x^{2n+1}}{(2n)!} \int_0^1 (1-u)^{2n} f^{(2n+1)}(xu)\,du$$

et avec la croissance de $f^{(2n+1)}$ (du fait que $f^{(2n)} \geq 0$) on a pour tous $r \in \,]0,a[$, $x \in \,]0,r[$, $u \in [0,1]$:

$$0 = f^{(2n+1)}(0) \leq f^{(2n+1)}(xu) \leq f^{(2n+1)}(ru)$$

ce qui entraîne :

$$0 \leq R_{2n}(x) \leq \frac{x^{2n+1}}{(2n)!} \int_0^1 (1-u)^{2n} f^{(2n+1)}(ru)\,du = \frac{x^{2n+1}}{r^{2n+1}} R_{2n}(r) \leq \frac{x^{2n+1}}{r^{2n+1}} f(r)$$

Il en résulte que $\lim\limits_{n\to+\infty} R_{2n}(x) = 0$ pour tout $x \in]0,a[$ et par parité le résultat est également vrai sur $]-a,a[$.

Enfin avec $P_{2n}(x) = P_{2n+1}(x)$, on déduit que $R_{2n}(x) = R_{2n+1}(x)$ et $\lim\limits_{n\to+\infty} R_n(x) = 0$ pour tout $x \in]-a,a[$. ■

Théorème 4.11 (Bernstein) *Soit f une fonction à valeurs réelles de classe $\mathcal{C}^\infty$ sur $]-a,a[$ avec $a>0$.*
Si $f^{(2k)}(x) \geq 0$ pour tout entier naturel k et tout $x \in]-a,a[$, la fonction f est alors développable en série entière sur $]-a,a[$.

Démonstration. On se ramène aux hypothèses du lemme précédent en introduisant la fonction g définie sur $]-a,a[$ par $g(x) = f(x) + f(-x)$.

Cette fonction est de classe $\mathcal{C}^\infty$, paire avec $g^{(2k)}(x) = f^{(2k)}(x) + f^{(2k)}(-x) \geq 0$ sur $]-a,a[$.

En notant $R_{n,f}$ [resp. $R_{n,g}$] le reste intégral dans la formule de Taylor à l'ordre n pour f [resp. g], on a pour tous $x \in]-a,a[$, $u \in [0,1]$:

$$0 \leq f^{(2n)}(xu) \leq g^{(2n)}(xu) = g^{(2n)}(|x|\,u)$$

et :

$$0 \leq |R_{2n-1,f}(x)| = \frac{x^{2n}}{(2n-1)!} \int_0^1 (1-u)^{2n-1} f^{(2n)}(xu)\,du \leq R_{2n-1,g}(x)$$

En utilisant le lemme précédent, on déduit donc que $\lim\limits_{n\to+\infty} R_{2n-1,f}(x) = 0$ pour tout $x \in]-a,a[$, puis avec :

$$f(x) = P_{2n,f}(x) + R_{2n,f}(x) = P_{2n-1,f}(x) + R_{2n-1,f}(x)$$

on déduit que :

$$R_{2n,f}(x) = -\frac{f^{(2n)}(0)}{(2n)!} x^{2n} + R_{2n-1,f}(x) = -\frac{1}{2} \frac{g^{(2n)}(0)}{(2n)!} x^{2n} + R_{2n-1,f}(x)$$

et avec $\lim\limits_{n\to+\infty} \dfrac{g^{(2n)}(0)}{(2n)!} x^{2n} = 0$ (convergence de la série de Taylor de la fonction paire g), on a $\lim\limits_{n\to+\infty} R_{2n,f}(x) = 0$ pour tout $x \in]-a,a[$.

D'où le résultat. ■

Dans le cas particulier où toutes les dérivées de f sont positives, on peut donner une démonstration plus simple du théorème de Bernstein.

Théorème 4.12 (Bernstein) *Soit f une fonction à valeurs réelles de classe $\mathcal{C}^\infty$ sur $]-a,a[$ avec $a>0$.*
Si $f^{(k)}(x) \geq 0$ pour tout entier naturel k et tout $x \in]-a,a[$, la fonction f est alors développable en série entière sur $]-a,a[$.

Démonstration. Pour $r \in]0, a[$ fixé et $x \in]0, r[$ on a pour tout entier naturel n, en gardant les notations qui précèdent :

$$R_n(x) = \int_0^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt = \int_0^x \frac{(r-t)^n}{n!} f^{(n+1)}(t) \left(\frac{x-t}{r-t}\right)^n dt$$

La fonction $\varphi : t \mapsto \dfrac{x-t}{r-t}$ étant décroissante sur $[0, x]$, on en déduit que :

$$\forall t \in [0, x], \quad 0 = \varphi(x) \leq \varphi(t) \leq \varphi(0) = \frac{x}{r}$$

et :

$$0 \leq R_n(x) \leq \left(\frac{x}{r}\right)^n \int_0^r \frac{(r-t)^n}{n!} f^{(n+1)}(t)\, dt = \left(\frac{x}{r}\right)^n R_n(r)$$

avec $R_n(r) = f(r) - P_n(r) \leq f(r)$. Avec $0 < \dfrac{x}{r} < 1$, on en déduit alors que $\lim\limits_{n \to +\infty} R_n(x) = 0$.

Pour $x \in]-a, 0[$, en écrivant $x = -x'$ avec $x \in]0, a[$, le changement de variable $u = -t$ donne, en tenant compte de $0 \leq f^{(n+1)}(-u) \leq f^{(n+1)}(u)$ ($f^{(n+2)} \geq 0$ entraîne $f^{(n+1)}$ croissante) :

$$\begin{aligned} |R_n(x)| &= \int_0^{x'} \frac{(x'-u)^n}{n!} f^{(n+1)}(-u)\, du \\ &\leq \int_0^{x'} \frac{(x'-u)^n}{n!} f^{(n+1)}(u)\, du = R_n(x') \end{aligned}$$

et $\lim\limits_{n \to +\infty} R_n(x) = 0$.

Pour $x = 0$, on a $R_n(0) = 0$. ■

Exemple 4.3 *On peut utiliser le théorème de Bernstein pour montrer que la fonction* tan *est développable en série entière sur* $\left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$.
Pour ce faire, on vérifie que, pour tout entier naturel n, il existe une fonction polynomiale P telle :

$$\tan^{(2n+1)}(x) = P_n\left(\tan^2(x)\right) \geq 0$$

donc la dérivée $\tan'$ *est développable en série entière sur* $\left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$ *et il en est de même de* tan *par primitivation.*

5

Fonctions convexes d'une variable réelle

5.1 Fonctions convexes

I désigne un intervalle réel non réduit à un point et f est une application définie sur I à valeurs réelles.

On rappelle que l'épigraphe de f est la partie de $\mathbb{R}^2$ définie par :

$$\mathcal{E}(f) = \left\{(x, y) \in \mathbb{R}^2 \mid y \geq f(x)\right\}$$

Définition 5.1 *On dit que la fonction f est :*

- *convexe si :*

$$\forall (x, y) \in I^2,\ \forall \lambda \in [0, 1],\ f((1-\lambda)x + \lambda y) \leq (1-\lambda) f(x) + \lambda f(y) \tag{5.1}$$

- *strictement convexe si pour tout couple (x, y) de points distincts de I et tout réel $\lambda \in \left]0, 1\right[$, on a :*

$$f((1-\lambda)x + \lambda y) < (1-\lambda) f(x) + \lambda f(y)$$

- *concave [resp. strictement concave] si la fonction $-f$ est convexe [resp. strictement convexe].*

Pour $\lambda = 0$ ou $\lambda = 1$, (5.1) est toujours vérifié (c'est une égalité). On peut donc se limiter à $\lambda \in \left]0, 1\right[$ pour la définition d'une fonction convexe.

Théorème 5.1 *L'application f est convexe si, et seulement si, son épigraphe est convexe dans $\mathbb{R}^2$.*

Démonstration. Supposons que la fonction f soit convexe.

Si $(x, y), (x', y')$ sont dans l'épigraphe $\mathcal{E}(f)$ et λ dans $[0, 1]$, on a alors :

$$(1-\lambda) y + \lambda y' \geq (1-\lambda) f(x) + \lambda f(x') \geq f((1-\lambda)x + \lambda x')$$

c'est-à-dire que le point $(1-\lambda)(x, y) + \lambda(x', y')$ est dans $\mathcal{E}(f)$.

Réciproquement, supposons que l'épigraphe $\mathcal{E}(f)$ soit convexe.

Pour x, x' dans I et λ dans $[0, 1]$, les points $(x, f(x))$ et $(x', f(x'))$ sont dans $\mathcal{E}(f)$, donc aussi le point $(1-\lambda)(x, f(x)) + \lambda(x', f(x'))$, ce qui signifie que :

$$(1-\lambda) f(x) + \lambda f(x') \geq f((1-\lambda)x + \lambda x')$$

La fonction f est donc convexe. ∎

Une autre interprétation graphique de la convexité est donnée par le résultat suivant.

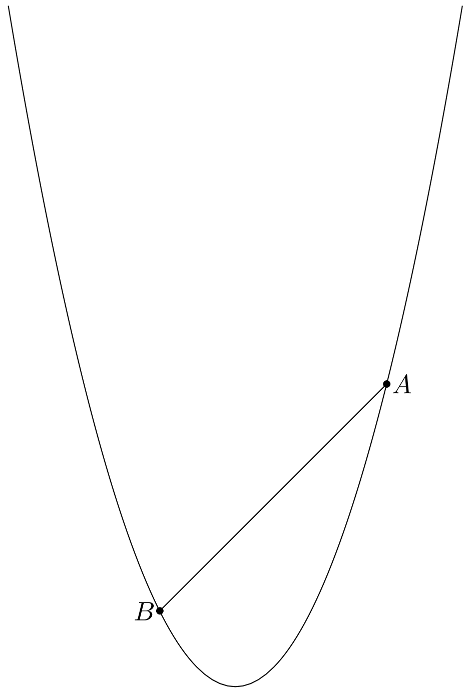

Figure 5.1 – Fonction convexe

Théorème 5.2 *La fonction f est convexe si, et seulement si, pour tout couple (A, B) de points du graphe de f avec $A = (a, f(a))$, $B = (b, f(b))$, $a < b$, la courbe représentative de la restriction de f à l'intervalle $[a, b]$ est au dessous de la corde $[AB]$ (figure 5.1*

Démonstration. On suppose que $a < b$.
Soit h l'application affine qui coïncide avec f en a et b :

$$\forall x \in \mathbb{R},\ h(x) = \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b)$$

(équation de la corde $[A, B]$).

Tout point x de $[a, b]$ s'écrit $x = (1-\lambda) a + \lambda b$ où $\lambda = \dfrac{x-a}{b-a}$, $1 - \lambda = \dfrac{b-x}{b-a}$ et avec la convexité de f, on a :

$$f(x) \leq (1-\lambda) f(a) + \lambda f(b) = \frac{b-x}{b-a} f(a) + \frac{x-a}{b-a} f(b) = h(x)$$

Réciproquement si, pour tous $a < b$ dans I, le graphe de la restriction de f à l'intervalle $[a, b]$ est sous la corde $[A, B]$, l'inégalité $f(x) \leq h(x)$ pour tout $x \in [a, b]$ se traduit par :

$$f(x) = f((1-\lambda) a + \lambda b) \leq h(x) = (1-\lambda) f(a) + \lambda f(b)$$

en posant $\lambda = \dfrac{x-a}{b-a}$. ∎

Remarque 5.1 *Dans le cas où f est strictement convexe, toute droite coupe le graphe de f en au plus deux points (pour $A \neq B$ le graphe de la restriction de f à $]a, b[$ est strictement au dessous de la corde $[A, B]$).*

Exemple 5.1 *Une fonction affine est à la fois convexe et concave sur $\mathbb{R}$. Graphiquement on peut se convaincre que la réciproque est vérifiée, c'est ce qu'on montrera un peu plus loin.*

Exemple 5.2 *Une limite simple de fonctions convexes est convexe puisque les inégalités de convexité sont conservées par passage à la limite.*

Exemple 5.3 *De l'exemple précédent, on déduit que si $(f_n)_{n\in\mathbb{N}}$ est une suite de fonctions convexes de I dans $\mathbb{R}$ telle que la série $\sum f_n$ converge simplement sur I vers une fonction f, alors cette fonction f est convexe.*

Exemple 5.4 *Soient J un intervalle réel non réduit à un point et $\varphi : I \times J \to \mathbb{R}$ une fonction telle que, pour tout $x \in I$ la fonction $t \mapsto \varphi(x,t)$ soit intégrable sur J.*
Si, pour tout $t \in J$, la fonction $x \mapsto \varphi(x,t)$ est convexe sur I, la fonction $f : x \mapsto \int_J \varphi(x,t)\,dt$ est alors convexe puisque les inégalités de convexité sont conservées par intégration.

Exemple 5.5 *Soient $a < b$ deux réels, $\varphi : [a,b] \to \mathbb{R}$ une fonction croissante. La fonction f définie sur $[a,b]$ par :*

$$f(x) = \int_a^x \varphi(t)\,dt$$

est convexe.
En effet, pour $x < y$ dans $[a,b]$ et λ dans $]0,1[$, on a :

$$\begin{aligned}
\theta(\lambda) &= f((1-\lambda)x + \lambda y) - (1-\lambda) f(x) - \lambda f(y) \\
&= \int_a^{x+\lambda(y-x)} \varphi(t)\,dt - \int_a^x \varphi(t)\,dt - \lambda\left(\int_a^y \varphi(t)\,dt - \int_a^x \varphi(t)\,dt\right) \\
&= \int_x^{x+\lambda(y-x)} \varphi(t)\,dt - \lambda \int_x^y \varphi(t)\,dt \\
&= \int_0^{\lambda(y-x)} \varphi(x+z)\,dz - \lambda \int_0^{y-x} \varphi(x+z)\,dz \\
&= \lambda \int_0^{y-x} \varphi(x+\lambda t)\,dt - \lambda \int_0^{y-x} \varphi(x+z)\,dz \\
&= \lambda \int_0^{y-x} (\varphi(x+\lambda t) - \varphi(x+t))\,dt \leq 0
\end{aligned}$$

pour φ croissante.
Dans le cas où φ est continue, la fonction f est dérivable de dérivée $f' = \varphi$ croissante. Nous verrons un peu plus loin que cela entraîne la convexité de f.
Pour f deux fois dérivable sur I avec $f''(x) \geq 0$ pour tout $x \in I$, la fonction f' est croissante, donc en écrivant que :

$$f(x) = \int_a^x f'(t)\,dt + f(a)$$

on en déduit que f est convexe.

Exemple 5.6 *De l'exemple précédent, on déduit que la fonction* $\exp$ *est convexe sur $\mathbb{R}$ et la fonction* $\ln$ *est concave sur $\mathbb{R}^{+,*}$.*

Théorème 5.3 *Si une fonction f est convexe sur I, elle est alors bornée sur tout segment $[a,b] \subset I$.*

Démonstration. Avec les notations de la démonstration précédente, on a pour tout $x \in [a,b] \subset I$:

$$f(x) \leq h(x) \leq M = \max\{h(a), h(b)\} = \max\{f(a), f(b)\}$$

donc la restriction de f à $[a,b]$ est majorée.

En écrivant tout point x de $[a,b] \subset I$ sous la forme $x = \dfrac{a+b}{2} + t$, où $t \in \left[-\dfrac{b-a}{2}, \dfrac{b-a}{2}\right]$, on a :

$$\begin{aligned} f\left(\frac{a+b}{2}\right) &= f\left(\frac{1}{2}\left(\frac{a+b}{2}+t\right)+\frac{1}{2}\left(\frac{a+b}{2}-t\right)\right) \\ &\leq \frac{1}{2}f\left(\frac{a+b}{2}+t\right)+\frac{1}{2}f\left(\frac{a+b}{2}-t\right) \end{aligned}$$

soit :

$$f(x) = f\left(\frac{a+b}{2}+t\right) \geq 2f\left(\frac{a+b}{2}\right) - f\left(\frac{a+b}{2}-t\right)$$

avec $f\left(\dfrac{a+b}{2}-t\right) \leq M$, ce qui nous donne :

$$f(x) \geq m = 2f\left(\frac{a+b}{2}\right) - M$$

donc la restriction de f à $[a,b]$ est minorée. ∎

Théorème 5.4 *Si une fonction f est convexe sur I et admet un minimum local en un point α de I, ce minimum est alors global.*

Démonstration. Supposons que ce minimum soit non global. Il existe alors un point x_0 différent de α dans I tel que $f(x_0) < f(\alpha)$.

Avec la convexité de f, on a pour tout $x = (1-\lambda)x_0 + \lambda\alpha \neq \alpha$ dans l'intervalle d'extrémités x_0 et α avec $\lambda \in [0,1[$:

$$f(x) \leq (1-\lambda)f(x_0) + \lambda f(\alpha) < (1-\lambda)f(\alpha) + \lambda f(\alpha) = f(\alpha)$$

ce qui contredit le caractère minimum local de α. ∎

Théorème 5.5 *Une combinaison linéaire à coefficients positifs de fonctions convexes de I dans $\mathbb{R}$ est convexe.*

Démonstration. Soient $(f_k)_{1\leq k\leq n}$ une suite de fonctions convexes de I dans $\mathbb{R}$, $(\lambda_k)_{1\leq k\leq n}$ une suite de réels positifs et $f = \sum\limits_{k=1}^{n} \lambda_k f_k$.

Pour $(x,y) \in I^2$ et $\lambda \in [0,1]$, on a :

$$\begin{aligned} f((1-\lambda)x+\lambda y) &= \sum_{k=1}^{n} \lambda_k f_k((1-\lambda)x+\lambda y) \\ &\leq \sum_{k=1}^{n} \lambda_k((1-\lambda)f_k(x)+\lambda f_k(y)) = (1-\lambda)f(x)+\lambda f(y) \end{aligned}$$

∎

Théorème 5.6 *Si $(f_k)_{1\leq k\leq n}$ est une suite finie de fonctions convexes de I dans $\mathbb{R}$, la fonction $f = \max\limits_{1\leq k\leq n} (f_k)$ est alors convexe.*

Démonstration. Pour $(x,y) \in I^2$, $\lambda \in [0,1]$ et $k \in \{1, \cdots, n\}$ on a :

$$\begin{aligned} f_k\left((1-\lambda)\,x + \lambda y\right) &\leq (1-\lambda)\,f_k(x) + \lambda f_k(y) \\ &\leq (1-\lambda)\,f(x) + \lambda f(y) \end{aligned}$$

donc $f\left((1-\lambda)\,x + \lambda y\right) \leq (1-\lambda)\,f(x) + \lambda f(y)$. ∎

Remarque 5.2 *Le produit de deux fonctions convexes n'est pas nécessairement convexe comme le montre l'exemple de la fonction $x \mapsto x^3$ (le graphe de la restriction à $[-1,1]$ n'est pas sous la corde).*

Remarque 5.3 *La composée de deux fonctions convexes n'est pas nécessairement convexe comme le montre l'exemple de la composée d'une fonction convexe f non affine avec la fonction $-x$, ce qui donne la fonction $-f$ concave qui est non convexe (car non affine).*
Par contre si f est une fonction convexe et φ une fonction convexe croissante alors la fonction composée $\varphi \circ f$ est convexe.
En effet, pour $(x,y) \in I^2$ et $\lambda \in [0,1]$, on a $f\left((1-\lambda)\,x + \lambda y\right) \leq (1-\lambda)\,f(x) + \lambda f(y)$ pour f convexe et avec la croissance et la convexité de φ, on a :

$$\begin{aligned} \varphi\left(f\left((1-\lambda)\,x + \lambda y\right)\right) &\leq \varphi\left((1-\lambda)\,f(x) + \lambda f(y)\right) \\ &\leq (1-\lambda)\,\varphi\left(f(x)\right) + \lambda\varphi\left(f(y)\right) \end{aligned}$$

La fonction $\varphi \circ f$ est donc convexe sur I.

5.2 Fonctions logarithmiquement convexes

Définition 5.2 *On dit qu'une fonction f définie sur I et à valeurs strictement positives est logarithmiquement convexe (ou plus simplement log-convexe) si la fonction $\ln(f)$ est convexe sur I.*

Théorème 5.7 *Une fonction log-convexe sur I est convexe sur I.*

Démonstration. Si f est log-convexe, la fonction $f = e^{\ln(f)}$ est alors convexe comme composée d'une fonction convexe avec une fonction convexe croissante. ∎

La réciproque du théorème précédant est fausse comme le montre l'exemple de la fonction $x \mapsto x$ qui est convexe alors que la fonction ln ne l'est pas (on verra plus loin qu'elle est concave).

S'il est facile de vérifier que le produit de deux fonctions log-convexes est log-convexe, le résultat analogue pour la somme n'est pas si simple à montrer. Il est conséquence du résultat suivant qui donne une définition équivalente de la notion de fonction log-convexe.

Théorème 5.8 *Une fonction f définie sur I et à valeurs strictement positives est logarithmiquement convexe si, et seulement si, pour tout réel $\alpha > 0$ la fonction $g_\alpha : x \mapsto f(x)\,\alpha^x$ est convexe sur I.*

Démonstration. Si f est log-convexe sur I, pour tout réel $\alpha > 0$ la fonction $x \mapsto \ln(g_\alpha(x)) = \ln(f(x)) + \ln(\alpha)x$ est convexe comme somme d'une fonction convexe et d'une fonction affine.

Réciproquement supposons que pour tout réel $\alpha > 0$ la fonction g_α soit convexe.

Pour $x \neq y$ dans I et $\lambda \in]0,1[$, on a alors en notant $z = (1-\lambda)x + \lambda y$:

$$f(z)\alpha^z \leq (1-\lambda) f(x)\alpha^x + \lambda f(y)\alpha^y$$

soit :

$$f(z) \leq (1-\lambda) f(x)\alpha^{x-z} + \lambda f(y)\alpha^{y-z}$$

avec $x - z = -\lambda(y-x)$ et $y - z = (1-\lambda)(y-x)$.

On a donc pour x, y, λ fixés :

$$\forall \alpha > 0,\ f(z) \leq \varphi(\alpha) = (1-\lambda) f(x)\alpha^{-\lambda(y-x)} + \lambda f(y)\alpha^{(1-\lambda)(y-x)}$$

La fonction φ est dérivable sur $]0,+\infty[$ avec :

$$\varphi'(\alpha) = (1-\lambda)\lambda(y-x)\alpha^{-\lambda(y-x)-1}\left(-f(x) + f(y)\alpha^{y-x}\right)$$

L'étude de ses variations nous dit alors que φ atteint son minimum en α défini par $\alpha^{y-x} = \dfrac{f(x)}{f(y)}$, ce minimum valant $f(x)^{1-\lambda} f(y)^\lambda$.

On a donc $f(z) \leq f(x)^{1-\lambda} f(y)^\lambda$ et $\ln(f(z)) \leq (1-\lambda)\ln(f(x) + \lambda\ln(f(y)))$, ce qui traduit la log-convexité de f. ■

Corollaire 5.1 *La somme de deux fonctions log-convexes de I dans $\mathbb{R}^{+,*}$ est une fonction log-convexe.*

Démonstration. Si f, g sont log-convexes, alors pour tout réel $\alpha > 0$ les fonctions $f\alpha^x$ et $g\alpha^x$ sont convexes, ce qui entraîne la convexité de la somme $(f+g)\alpha^x$ et la log-convexité de $f+g$. ■

5.3 Régularité des fonctions convexes

Ici encore, I est un intervalle réel non réduit à un point et f est une application définie sur I à valeurs réelles.

On note $\overset{\circ}{I}$ l'intérieur de l'intervalle I.

Pour $x \neq y$ dans I, on note :

$$p(x,y) = \frac{f(y) - f(x)}{y - x}$$

la pente de la droite (MN) où $M = (x, f(x))$ et $N = (y, f(y))$ sont deux points du graphe de f.

Comme x et y jouent des rôles symétriques, on a $p(x,y) = p(y,x)$.

Théorème 5.9 *Si f est convexe sur I, sa restriction à tout intervalle $[a,b] \subset \overset{\circ}{I}$ est alors lipschitzienne. En conséquence, f est continue sur $\overset{\circ}{I}$.*

Démonstration. On se donne, une fonction f convexe sur I, un intervalle $[a, b] \subset \overset{\circ}{I}$ et un réel $\varepsilon > 0$ tel que $[a - \varepsilon, b + \varepsilon] \subset I$.

On se donne $x \neq y$ dans $[a, b]$ et on note $z = y + \dfrac{y - x}{|y - x|}\varepsilon$ ($z = y \pm \varepsilon$, où $\pm$ est le signe $y - x$).

Pour $x < y$, on a $z = y + \varepsilon$ et $y \in]x, z[$ et pour $x > y$, on a $z = y - \varepsilon$ et $y \in]z, x[$, donc dans tous les cas, on peut écrire que $y = (1 - \lambda) x + \lambda z$ avec $\lambda \in]0, 1[$ est donné par :

$$\begin{aligned}\lambda &= \frac{y - x}{z - x} = \frac{y - x}{y + \frac{y-x}{|y-x|}\varepsilon - x} = \frac{(y - x)|y - x|}{|y - x|(y - x) + (y - x)\varepsilon} \\ &= \frac{|y - x|}{|y - x| + \varepsilon}\end{aligned}$$

On a donc :

$$f(y) = f((1 - \lambda) x + \lambda z) \leq (1 - \lambda) f(x) + \lambda f(z)$$

soit :

$$\begin{aligned}f(y) - f(x) &\leq \lambda (f(z) - f(x)) = \frac{|y - x|}{|y - x| + \varepsilon}(f(z) - f(x)) \\ &\leq \frac{(f(z) - f(x))}{\varepsilon}|y - x|\end{aligned}$$

Comme la fonction convexe f est bornée sur $[a - \varepsilon, b + \varepsilon]$, il existe des constantes $m \leq M$ telles que :

$$\forall t \in [a - \varepsilon, b + \varepsilon], \ m \leq f(t) \leq M$$

ce qui nous donne :

$$f(y) - f(x) \leq \frac{M - m}{\varepsilon}|y - x|$$

pour tous réels $x \neq y$ dans $[a, b]$, ce qui équivaut à $|f(y) - f(x)| \leq \dfrac{M - m}{\varepsilon}|y - x|$ (pour $f(x) > f(y)$, on applique l'inégalité précédente au couple (y, x)).

La fonction f est donc lipschitzienne sur tout $[a, b] \subset \overset{\circ}{I}$ et en conséquence elle est uniformément continue sur ces intervalles.

Il en résulte que f est continue sur $\overset{\circ}{I}$. ∎

Remarque 5.4 *Une fonction convexe sur I n'est pas nécessairement continue sur tout l'intervalle I comme le montre l'exemple de la fonction f définie sur $[0, 1]$ par $f(x) = 0$ si $0 < x < 1$ et $f(0) = f(1) = 1$.*

Théorème 5.10 *Si f est continue sur I et convexe sur $\overset{\circ}{I}$, elle est alors convexe sur I.*

Démonstration. Pour $x < y$ dans I il existe un entier $n_0 \geq 1$ tel que $x_n = x + \dfrac{1}{n}$ et $y_n = y - \dfrac{1}{n}$ soient dans $\overset{\circ}{I}$ pour tout $n \geq n_0$ avec $x_n < y_n$.

On a alors avec la convexité de f sur $\overset{\circ}{I}$ et la continuité sur I, pour tout réel $\lambda \in [0, 1]$:

$$\begin{aligned}f((1 - \lambda) x + \lambda y) &= \lim_{n \to +\infty} f((1 - \lambda) x_n + \lambda y_n) \\ &\leq \lim_{n \to +\infty} ((1 - \lambda) f(x_n) + \lambda f(y_n)) \\ &\leq (1 - \lambda) f(x) + \lambda f(y)\end{aligned}$$

∎

Théorème 5.11 *Pour $f : I \to \mathbb{R}$ les trois propriétés suivantes sont équivalentes :*

1. *f est convexe ;*
2. *pour tous x, y, z dans I tels que $x < y < z$, on a :*

$$p(x, y) \leq p(x, z) \leq p(y, z)$$

(inégalité des trois pentes, figure 5.2) ;

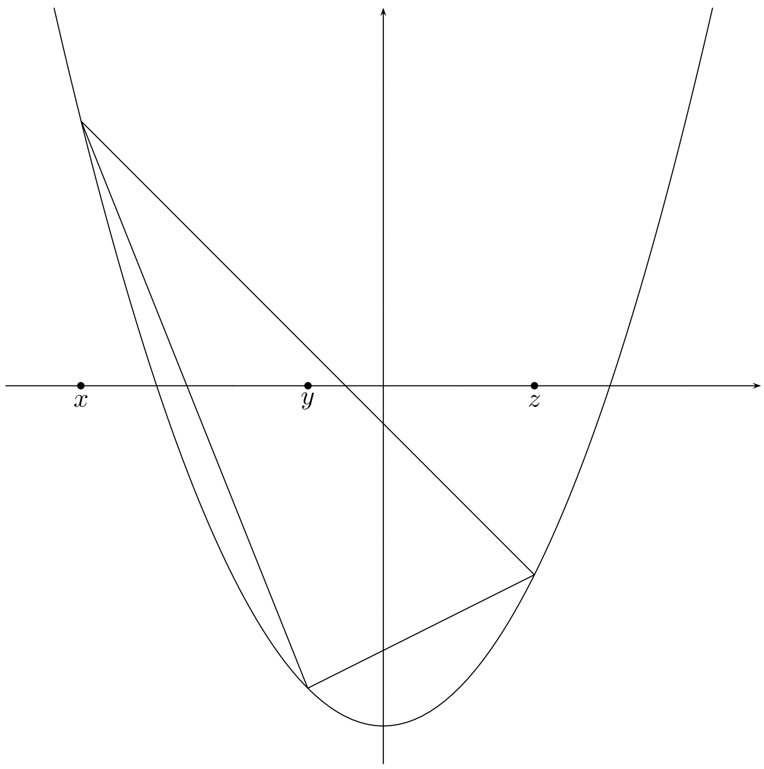

FIGURE 5.2 – Inégalité des trois pentes

3. *pour tout $a \in I$ la fonction :*

$$\tau_a : x \mapsto \frac{f(x) - f(a)}{x - a}$$

est croissante sur $I - \{a\}$.

Démonstration. Tout est basé sur le fait que tout point $y \in]x, z[$ s'écrit $y = (1 - \lambda) x + \lambda z$ avec $\lambda = \dfrac{y - x}{z - x}$ et $1 - \lambda = \dfrac{z - y}{z - x}$ dans $]0, 1[$.

(1) $\Rightarrow$ (2) Si f est convexe, on a alors pour $y = (1 - \lambda) x + \lambda z \in]x, z[$:

$$f(y) \leq (1 - \lambda) f(x) + \lambda f(z)$$

soit :

$$f(y) - f(x) \leq \lambda (f(z) - f(x)) = \frac{y - x}{z - x} (f(z) - f(x))$$

et :

$$f(y) - f(z) \leq (1 - \lambda)(f(x) - f(z)) = \frac{z - y}{z - x} (f(x) - f(z))$$

ce qui nous donne :

$$p(x,y) = \frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(x)}{z - x} = p(x,z)$$

et :

$$-p(y,z) = \frac{f(y) - f(z)}{z - y} \leq -p(x,z) = \frac{f(x) - f(z)}{z - x}$$

ou encore $p(x,z) \leq p(y,z)$.

(2) $\Rightarrow$ (3) Pour $x < y < a$, on a :

$$\tau_a(x) = p(x,a) \leq p(y,a) = \tau_a(y)$$

Pour $x < a < y$, on a :

$$\tau_a(x) = p(x,a) \leq p(a,y) = \tau_a(y)$$

Pour $a < x < y$, on a :

$$\tau_a(x) = p(a,x) \leq p(a,y) = \tau_a(y)$$

Ou alors, on peut remarquer que la fonction :

$$(a,x,y) \mapsto \frac{\tau_a(y) - \tau_a(x)}{y - x} = \frac{p(a,y) - p(a,x)}{y - x}$$

est invariante par permutation de (a,x,y) et positive pour $x < y < a$.

(3) $\Rightarrow$ (1) Pour $x = y$, l'inégalité de convexité est une égalité trivialement vérifiée. On se donne donc $x < y$ dans I, λ dans $]0,1[$ et on note $t = (1-\lambda)x + \lambda y \in]x,y[$. Avec la croissance de τ_y sur $I - \{y\}$ on a :

$$\tau_y(x) = \frac{f(x) - f(y)}{x - y} \leq \tau_y(t) = \frac{f(t) - f(y)}{t - y}$$

ce qui s'écrit :

$$(t-y)\frac{f(x) - f(y)}{x - y} + f(y) \geq f(t)$$

ou encore :

$$f(t) \leq \frac{y-t}{y-x} f(x) + \frac{t-x}{y-x} f(y) = (1-\lambda) f(x) + \lambda f(y)$$

ce qui traduit la convexité de f. ■

De ce théorème, on déduit que la fonction f est concave si, et seulement si, pour tout $a \in I$ la fonction τ_a est décroissante sur $I - \{a\}$, ce qui se déduit immédiatement de $\tau_{-f,a} = -\tau_{f,a}$.

On démontre de manière analogue que la fonction f est strictement convexe [resp. strictement concave] si, et seulement si, pour tout $a \in I$ la fonction τ_a est strictement croissante [resp. strictement décroissante] sur $I - \{a\}$.

C'est principalement la propriété (3) qui va nous permettre de déduire quelques propriétés importantes de régularité des fonctions convexes.

Exercice 5.1 *En utilisant la caractérisation* (3) *du théorème précédent, retrouver le fait que si f est convexe sur I, sa restriction à tout intervalle $[a,b] \subset \overset{\circ}{I}$ est alors lipschitzienne (et f est continue sur $\overset{\circ}{I}$).*

Solution 5.1 *On utilise la croissance des fonctions τ_x, pour tout $x \in [a,b]$.*
Soient $[a,b] \subset \overset{\circ}{I}$ et $\varepsilon > 0$ tel que $[a-\varepsilon, b+\varepsilon] \subset I$.
Pour $a \leq x < y \leq b$ on a $\tau_x(a-\varepsilon) \leq \tau_x(y) \leq \tau_x(b+\varepsilon)$ ce qui se traduit par :

$$\tau_x(a-\varepsilon)(y-x) \leq f(y) - f(x) \leq \tau_x(b+\varepsilon)(y-x)$$

avec :

$$\tau_x(a-\varepsilon) = \tau_{a-\varepsilon}(x) \geq \tau_{a-\varepsilon}(a) = m \text{ et } \tau_x(b+\varepsilon) = \tau_{b+\varepsilon}(x) \leq \tau_{b+\varepsilon}(b) = M$$

On a donc :

$$m(y-x) \leq f(y) - f(x) \leq M(y-x)$$

soit $|f(y) - f(x)| \leq \max(|m|, |M|)|y-x|$ pour tous réels x, y dans $[a,b]$ (pour $f(y) \geq f(x)$, on a $|f(y) - f(x)| = f(y) - f(x) \leq M(y-x) = |M||y-x|$ et pour $f(y) \leq f(x)$, on a $|f(y) - f(x)| = f(x) - f(y) \leq m(x-y) = |m||y-x|$).
La fonction f est donc lipschitzienne sur tout $[a,b] \subset \overset{\circ}{I}$.

Théorème 5.12 *Une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ est affine si, et seulement si, elle est à la fois convexe et concave.*

Démonstration. Il est clair qu'une fonction affine est convexe et concave.

Si $f : \mathbb{R} \to \mathbb{R}$ est convexe et concave, la fonction τ_0 est alors constante sur $\mathbb{R}^*$ puisque qu'elle est croissante et décroissante.

On a donc $\tau_0(x) = \tau_0(1)$ pour tout $x \in \mathbb{R}^*$, ce qui équivaut à $f(x) = ax + b$ pour tout $x \in \mathbb{R}^*$ avec $a = \tau_0(1) = f(1) - f(0)$ et $b = f(0)$.

Cette égalité étant encore vérifiée pour $x = 0$, la fonction f est affine. ∎

Exercice 5.2 *Montrer que si f, g sont deux fonctions convexes de $\mathbb{R}$ dans $\mathbb{R}$ telles que la fonction $f + g$ soit affine, alors les fonctions f et g sont également affines.*

Solution 5.2 *On a $f(x) + g(x) = h(x) = ax + b$ pour tout réel x.*
La fonction convexe $f = h - g$ est alors également concave comme somme de deux fonctions concaves, elle est donc affine et aussi $g = h - f$.

Exercice 5.3 *Montrer qu'une fonction f de $\mathbb{R}$ dans $\mathbb{R}$ est constante si, et seulement si, elle est convexe et majorée.*

Solution 5.3 *Soit $f : \mathbb{R} \to \mathbb{R}$ convexe et majorée.*
Si f n'est pas constante il existe $a < b$ tels que $f(a) \neq f(b)$.
Supposons que $f(a) < f(b)$. Avec la croissance de τ_b on a :

$$\forall x > b,\ \tau_b(x) = \frac{f(x) - f(b)}{x - b} \geq \tau_b(a) = \frac{f(b) - f(a)}{b - a} = \lambda > 0$$

soit :

$$\forall x > b,\ f(x) \geq \lambda x + \mu$$

avec $\lim\limits_{x \to +\infty} (\lambda x + \mu) = +\infty$, ce qui est en contradiction avec le fait que f est majorée.
Le cas où $f(a) > f(b)$ se traite de manière analogue en considérant τ_a et la limite en $-\infty$.
Avec la croissance de τ_a on a :

$$\forall x < a,\ \tau_a(x) = \frac{f(x) - f(a)}{x - a} \leq \tau_a(b) = \frac{f(b) - f(a)}{b - a} = \lambda < 0$$

soit :

$$\forall x < a,\ f(x) \geq \lambda x + \mu$$

avec $\lim\limits_{x \to -\infty} (\lambda x + \mu) = +\infty$, *ce qui est en contradiction avec le fait que* f *est majorée.*

Remarque 5.5 *Le résultat de l'exercice précédent n'est pas valable pour une fonction convexe majorée sur un intervalle réel distinct de* $\mathbb{R}$. *Par exemple la fonction* f *définie sur* $\mathbb{R}^+$ *par* $f(x) = \dfrac{1}{1+x}$ *est convexe et majorée par* 1 *sans être constante.*

Exercice 5.4 *Montrer que* f *est convexe si, et seulement si, pour tous* $x < y < z$ *dans* I *on a :*

$$d(x,y,z) = \begin{vmatrix} 1 & 1 & 1 \\ x & y & z \\ f(x) & f(y) & f(z) \end{vmatrix} \geq 0$$

Solution 5.4 *Pour* $x < y < z$ *dans* I *on a :*

$$\begin{aligned} d(x,y,z) &= \begin{vmatrix} 1 & 0 & 0 \\ x & y-x & z-x \\ f(x) & f(y)-f(x) & f(z)-f(x) \end{vmatrix} \\ &= (y-x)(z-x) \begin{vmatrix} 1 & 1 \\ \dfrac{f(y)-f(x)}{y-x} & \dfrac{f(z)-f(x)}{z-x} \end{vmatrix} \\ &= (y-x)(z-x) \begin{vmatrix} 1 & 1 \\ p(x,y) & p(x,z) \end{vmatrix} \\ &= (y-x)(z-x)(p(x,z)-p(x,y)) \end{aligned}$$

Si f *est convexe, on a alors* $p(x,y) \leq p(x,z)$, *ce qui équivaut à* $d(x,y,z) \geq 0$.
Réciproquement si $d(x,y,z) \geq 0$ *pour tous* $x < y < z$ *dans* I, *on a alors* $p(x,y) \leq p(x,z)$ *et :*

$$(y-x)(f(z)-f(x)) - (z-x)(f(y)-f(x)) \geq 0$$

ce qui équivaut à :

$$(y-x) f(z) + (z-y) f(x) - (z-x) f(y) \geq 0$$

encore équivalent à :

$$(z-y)(f(x)-f(z)) + (z-x)(f(z)-f(y)) \geq 0$$

(en écrivant $y - x = (y-z) + (z-x)$*) et divisant par* $(z-y)(z-x)$ *on aboutit à* $-p(x,z) + p(y,z) \geq 0$. *On a donc* $p(x,y) \leq p(x,z) \leq p(y,z)$ *pour tous* $x < y < z$, *ce qui équivaut à la convexité de* f.

Avec l'exercice qui suit, on s'intéresse aux branches infinies des fonctions convexes.

Exercice 5.5 *Soit* f *une fonction convexe de* $\mathbb{R}^+$ *dans* $\mathbb{R}$.

1. *Montrer que la fonction* $g : x \mapsto \dfrac{f(x)}{x}$ *admet une limite finie ou égale à* $+\infty$ *quand* x *tend vers* $+\infty$.
2. *Montrer que si* $\lim\limits_{x \to +\infty} \dfrac{f(x)}{x} = \alpha \in \mathbb{R}$ *alors la fonction* $h : x \mapsto f(x) - \alpha x$ *admet une limite finie ou égale à* $-\infty$ *quand* x *tend vers* $+\infty$.

Solution 5.5

1. *Si f est convexe sur $\mathbb{R}^+$ la fonction $\tau_0 : x \mapsto \dfrac{f(x) - f(0)}{x}$ est alors croissante de $\mathbb{R}^{+,*}$ dans $\mathbb{R}$, donc elle admet une limite $\alpha \in \mathbb{R} \cup \{+\infty\}$ et avec $\lim\limits_{x \to +\infty} \dfrac{f(0)}{x} = 0$ on déduit que $\lim\limits_{x \to +\infty} \dfrac{f(x)}{x} = \alpha.$*
Pour $\alpha = \infty$, on a une branche parabolique.
Par exemple pour $f : x \mapsto x^2$ on a $\lim\limits_{x \to +\infty} \dfrac{f(x)}{x} = +\infty.$
2. *On montre tout d'abord que la fonction h est décroissante de $\mathbb{R}^+$ dans $\mathbb{R}$.*
Cette fonction est convexe comme somme de deux fonctions convexes, donc pour tout réel $a \in \mathbb{R}^+$ la fonction $\tau_a : x \mapsto \dfrac{h(x) - h(a)}{x - a}$ est croissante sur $\mathbb{R}^+ \setminus \{a\}$ et avec :
$$\lim_{x \to +\infty} \tau_a(x) = \lim_{x \to +\infty} \left(\frac{f(x) - f(a)}{x - a} - \alpha \right) = \lim_{x \to +\infty} \left(\frac{f(x)}{x} - \alpha \right) = 0$$
on déduit que $\tau_a(x) \leq 0$ pour tout $x \in \mathbb{R}^+ \setminus \{a\}$ et en particulier pour $x > a$ on en déduit que $h(x) \leq h(a)$.
La fonction h est donc décroissante sur $\mathbb{R}^+$ et en conséquence elle admet une limite $\beta \in \mathbb{R} \cup \{-\infty\}$.
Pour $\beta \in \mathbb{R}$, on a une droite asymptote et pour $\beta = -\infty$, on a une direction asymptotique.
Par exemple pour $f : x \mapsto -\ln(1 + x)$ on a $\lim\limits_{x \to +\infty} \dfrac{f(x)}{x} = 0$ et $\lim\limits_{x \to +\infty} (f(x) - \alpha x) = -\infty.$

Exercice 5.6 *Soit f une fonction de $\mathbb{R}_+^*$ dans $\mathbb{R}$ et g la fonction définie sur $\mathbb{R}_+^*$ par $g(x) = xf\left(\dfrac{1}{x}\right)$.*
Montrer que les fonctions f et g sont simultanément convexes.

Solution 5.6 *Pour $a \in \mathbb{R}_+^*$ et $x \in \mathbb{R}_+^* \setminus \{a\}$, on a :*
$$\begin{aligned}
p_g(x, a) &= \frac{g(a) - g(x)}{a - x} = \frac{af\left(\frac{1}{a}\right) - xf\left(\frac{1}{x}\right)}{a - x} \\
&= \frac{(a - x) f\left(\frac{1}{a}\right) + x\left(f\left(\frac{1}{a}\right) - f\left(\frac{1}{x}\right)\right)}{a - x} \\
&= f\left(\frac{1}{a}\right) + x\frac{f\left(\frac{1}{a}\right) - f\left(\frac{1}{x}\right)}{a - x} \\
&= f\left(\frac{1}{a}\right) + x\frac{f\left(\frac{1}{a}\right) - f\left(\frac{1}{x}\right)}{\frac{1}{a} - \frac{1}{x}} \frac{\frac{1}{a} - \frac{1}{x}}{a - x} \\
&= f\left(\frac{1}{a}\right) - \frac{1}{a} p_f\left(\frac{1}{x}, \frac{1}{a}\right)
\end{aligned}$$

Si f est convexe, la fonction $t \mapsto p_f\left(t, \dfrac{1}{a}\right)$ est croissante, donc la fonction $x \mapsto p_f\left(\dfrac{1}{x}, \dfrac{1}{a}\right)$ est décroissante et $x \mapsto p_g(x, a)$ est croissante, ce qui signifie que g est convexe.
En remarquant que $f(x) = xg\left(\dfrac{1}{x}\right)$, on a la réciproque.

On peut aussi procéder directement comme suit, mais c'est un peu plus compliqué.
Pour $0 < x < y$ *et* $0 < \lambda < 1$*, on a :*

$$\begin{aligned} g\left(\left(1-\lambda\right)x+\lambda y\right) &= \left(\left(1-\lambda\right)x+\lambda y\right) f\left(\frac{1}{\left(1-\lambda\right)x+\lambda y}\right) \\ &= \left(\left(1-\lambda\right)x+\lambda y\right) f\left(\left(1-\mu\right)\frac{1}{x}+\mu\frac{1}{y}\right), \end{aligned}$$

où μ *est défini par :*

$$\frac{1}{\left(1-\lambda\right)x+\lambda y} = \left(1-\mu\right)\frac{1}{x}+\mu\frac{1}{y} = \mu\frac{x-y}{xy}+\frac{1}{x},$$

soit :

$$\mu = \frac{\lambda y}{\left(1-\lambda\right)x+\lambda y} \in \left]0,1\right[.$$

Si f *est convexe, alors :*

$$\begin{aligned} g\left(\left(1-\lambda\right)x+\lambda y\right) &= \frac{\lambda y}{\mu} f\left(\left(1-\mu\right)\frac{1}{x}+\mu\frac{1}{y}\right) \\ &\leq \frac{\lambda y}{\mu}\left(\left(1-\mu\right)f\left(\frac{1}{x}\right)+\mu f\left(\frac{1}{y}\right)\right), \end{aligned}$$

ce qui donne, en écrivant que $1-\mu = \left(1-\lambda\right)x\dfrac{\mu}{\lambda y}$ *:*

$$g\left(\left(1-\lambda\right)x+\lambda y\right) \leq \frac{\lambda y}{\mu}\left(1-\lambda\right)x\frac{\mu}{\lambda y}f\left(\frac{1}{x}\right)+\frac{\lambda y}{\mu}\mu f\left(\frac{1}{y}\right) = \left(1-\lambda\right)g\left(x\right)+\lambda g\left(y\right),$$

ce qui signifie que g *convexe.*

On déjà vu qu'une fonction convexe sur I est continue sur l'intérieur $\overset{\circ}{I}$, mais elle n'est pas nécessairement dérivable (prendre par exemple $|x|$ sur $[-1,1]$).

Théorème 5.13 *Si* f *est convexe sur* I*, elle admet alors une dérivée à droite et à gauche en tout point de* $\overset{\circ}{I}$*, les fonctions dérivées à droite et à gauche étant croissantes sur* $\overset{\circ}{I}$ *et pour* $a < b$ *dans* $\overset{\circ}{I}$ *on a :*

$$f'_g\left(a\right) \leq f'_d\left(a\right) \leq \frac{f\left(b\right)-f\left(a\right)}{b-a} \leq f'_g\left(b\right) \leq f'_d\left(b\right)$$

Démonstration. Pour $c \in \overset{\circ}{I}$ et a, b dans I tels que $a < c < b$, avec la croissance de τ_c sur $I - \{c\}$, on a :

$$\forall x \in [a,b]-\{c\},\ \tau_c\left(a\right) \leq \tau_c\left(x\right) \leq \tau_c\left(b\right)$$

donc τ_c est croissante bornée sur $[a,b] \setminus \{c\}$ et les limites suivantes existent :

$$f'_g\left(c\right) = \lim_{\substack{x\to c\\ x<c}} \tau_c\left(x\right) = \sup_{a\leq x<c} \tau_c\left(x\right) \leq f'_d\left(c\right) = \lim_{\substack{x\to c\\ x>c}} \tau_c\left(x\right) = \inf_{c<x\leq b} \tau_c\left(x\right)$$

Pour $a < b$ dans $\overset{\circ}{I}$, on a :

$$f'_d\left(a\right) = \inf_{x\in]a,+\infty[\cap I} \tau_a\left(x\right) \leq \tau_a\left(b\right) = \tau_b\left(a\right) \leq \sup_{x\in]-\infty,b[\cap I} \tau_b\left(x\right) = f'_g\left(b\right)$$

■

On retrouve le fait qu'une fonction convexe sur I est continue sur $\overset{\circ}{I}$.

La version suivante du théorème de Rolle nous sera utile. Pour sa démonstration, il suffit de reprendre la démonstration de la version classique du théorème.

Théorème 5.14 (Rolle) *Si φ est une fonction continue sur un segment $[a,b]$ et dérivable à droite et à gauche en tout point de $]a,b[$ telle que $\varphi(a)=\varphi(b)$, il existe alors un point $c\in\,]a,b[$ tel que $\varphi'_g(c)\,\varphi'_d(c)\leq 0$.*

Démonstration. Si φ est constante, elle est alors dérivable sur $[a,b]$ de dérivée nulle et tout point c de $]a,b[$ convient.

On suppose φ non constante.

Comme φ est continue sur le compact $[a,b]$, elle est bornée et atteint ses bornes, c'est-à-dire qu'il existe α,β dans $[a,b]$ tels que $\varphi(\alpha)=\inf\limits_{x\in[a,b]}\varphi(x)$ et $\varphi(\beta)=\sup\limits_{x\in[a,b]}\varphi(x)$.

Si α et β sont dans $\{a,b\}$, on a alors $\varphi(\alpha)\leq\varphi(x)\leq\varphi(\beta)=\varphi(\alpha)$ pour tout $x\in[a,b]$ et φ est constante contrairement à l'hypothèse de départ, on a donc $\alpha\in\,]a,b[$ ou $\beta\in\,]a,b[$.

Si $\alpha\in\,]a,b[$, on a alors $\dfrac{\varphi(x)-\varphi(\alpha)}{x-\alpha}\geq 0$ pour $\alpha<x\leq b$, donc $\varphi'_d(\alpha)\geq 0$ et $\dfrac{\varphi(x)-\varphi(\alpha)}{x-\alpha}\leq 0$ pour $a\leq x<\alpha$, donc $\varphi'_g(\alpha)\leq 0$.

Si $\beta\in\,]a,b[$, on a alors $\dfrac{\varphi(x)-\varphi(\beta)}{x-\beta}\leq 0$ pour $\beta<x\leq b$, donc $\varphi'_d(\beta)\leq 0$ et $\dfrac{\varphi(x)-\varphi(\beta)}{x-\beta}\geq 0$ pour $a\leq x<\beta$, donc $\varphi'_g(\beta)\geq 0$. ∎

Dans le cas où φ est dérivable sur $]a,b[$, on retrouve le théorème de Rolle usuel ($\varphi'_g(c)\,\varphi'_d(c)=(\varphi'(c))^2\leq 0$, donc $\varphi'(c)=0$).

En utilisant le théorème de Rolle, on déduit la caractérisation suivante de convexité pour les fonctions deux fois dérivable.

Théorème 5.15 *Si f est deux fois dérivable sur I, elle est alors convexe sur I si, et seulement si, $f''(x)\geq 0$ pour tout $x\in I$.*

Démonstration. Pour $x<y$ fixés dans I et $\lambda\in[0,1]$ on pose :

$$\varphi(\lambda)=f(\lambda x+(1-\lambda)y)-\lambda f(x)-(1-\lambda)f(y)$$

Cette fonction est deux fois dérivable sur $[0,1]$ avec :

$$\left\{\begin{array}{l}\varphi'(\lambda)=(x-y)\,f'(\lambda x+(1-\lambda)y)+f(y)-f(x)\\ \varphi''(\lambda)=(x-y)^2 f''(\lambda x+(1-\lambda)y)\geq 0\end{array}\right.$$

La fonction φ' est donc croissante sur $[0,1]$.

D'autre part, on a $\varphi(0)=\varphi(1)=0$, le théorème de Rolle nous dit alors qu'il existe $c\in\,]0,1[$ tel que $\varphi'(c)=0$.

Avec la croissance de φ' on a alors $\varphi'(\lambda)\leq\varphi'(c)=0$ pour tout $\lambda\in[0,c]$ et $\varphi'(\lambda)\geq\varphi'(c)=0$ pour tout $\lambda\in[c,1]$, c'est-à-dire que φ est décroissante sur $[0,c]$ et croissante sur $[c,1]$, il en résulte que $\varphi(\lambda)\leq 0$ pour tout $\lambda\in[0,1]$, c'est-à-dire que f est convexe. ∎

Avec l'exemple 5.5 on a une autre démonstration du théorème précédent.

On a aussi la version suivante du théorème des accroissements finis.

Théorème 5.16 *Si φ est une fonction continue sur un segment $[a,b]$ et dérivable à droite et à gauche en tout point de $]a,b[$, il existe alors un point $c\in\,]a,b[$ tel que $\dfrac{\varphi(b)-\varphi(a)}{b-a}$ soit dans le segment d'extrémités $\varphi'_g(c)$ et $\varphi'_d(c)$.*

Démonstration. On se ramène aux hypothèses du théorème de Rolle en considérant la fonction g définie sur $[a,b]$ par :

$$g(x)=\varphi(x)-\varphi(a)-\lambda(x-a)$$

où la constante réelle λ est telle que $g(b) = g(a) = 0$ (soit $\lambda = \dfrac{\varphi(b) - \varphi(a)}{b-a}$).

Le théorème de Rolle appliqué à la fonction g (qui est bien continue sur $[a, b]$ et dérivable à droite et gauche sur $]a, b[$) nous assure de l'existence d'un réel $c \in]a, b[$ tel que $g'_g(c)\, g'_d(c) \leq 0$, ce qui équivaut à $\left(\varphi'_g(c) - \lambda\right)\left(\varphi'_d(c) - \lambda\right) \leq 0$ et revient à dire que $\lambda = \dfrac{\varphi(b) - \varphi(a)}{b-a}$ est dans le segment d'extrémités $\varphi'_g(c)$ et $\varphi'_d(c)$. ■

Dans le cas où f est dérivable sur $]a, b[$, on retrouve le théorème des accroissements finis usuel.

Corollaire 5.2 *Soit φ une fonction continue sur un segment $[a, b]$ et dérivable à droite et à gauche en tout point de $]a, b[$. Si $\varphi'_d(x) \geq 0$ et $\varphi'_g(x) \geq 0$ pour tout $x \in]a, b[$, la fonction φ est alors croissante sur $[a, b]$.*

Démonstration. Pour $x < y$ dans $[a, b]$, le quotient $\dfrac{\varphi(y) - \varphi(x)}{y - x}$ qui est compris entre $\varphi'_d(c_{x,y})$ et $\varphi'_g(c_{x,y})$ (où $c_{x,y}$ est entre x et y) est positif, donc $\varphi(y) \geq \varphi(x)$ et φ est croissante. ■

Corollaire 5.3 *Soient φ une fonction dérivable sur I et α un point de I. Si $(x - \alpha)\varphi'(x) \geq 0$ [resp. $(x - \alpha)\varphi'(x) \leq 0$] pour tout $x \in I$, la fonction φ admet alors un minimum [resp. maximum] global en α.*

Démonstration. Pour $x \in I - \{\alpha\}$ le théorème des accroissements finis nous permet d'écrire :

$$\varphi(x) - \varphi(\alpha) = (x - \alpha)\varphi'(c_x) = \frac{x - \alpha}{c_x - \alpha}(c_x - \alpha)\varphi'(c_x)$$

avec c_x strictement compris entre α et x. Avec $\dfrac{x - \alpha}{c_x - \alpha} > 0$ et $(c_x - \alpha)\varphi'(c_x) \geq 0$ [resp. $(c_x - \alpha)\varphi'(c_x) \leq 0$] on en déduit que $\varphi(x) \geq \varphi(\alpha)$ [resp. $\varphi(x) \leq \varphi(\alpha)$].

La fonction φ admet donc un minimum [resp. maximum] global en α. ■

Théorème 5.17 *Dans le cas où I est un intervalle ouvert, une fonction f est convexe sur I si, et seulement si, elle est dérivable à droite et à gauche sur I de dérivées f'_d et f'_g croissantes.*

Démonstration. On sait déjà que la condition est nécessaire.

Réciproquement supposons f dérivable à droite et à gauche sur I de dérivées f'_d et f'_g croissantes. Pour $x < y$ dans I et $\lambda \in [0, 1]$ la fonction :

$$\varphi : t \mapsto f((1 - \lambda)t + \lambda y) - (1 - \lambda)f(t) - \lambda f(y)$$

est continue sur $[x, y]$ et dérivable à droite et à gauche en tout point de $]x, y[$ avec :

$$\varphi'_d(t) = (1 - \lambda)\left(f'_d((1 - \lambda)t + \lambda y) - f'_d(t)\right) \geq 0$$

et :

$$\varphi'_g(t) = (1 - \lambda)\left(f'_g((1 - \lambda)t + \lambda y) - f'_g(t)\right) \geq 0$$

($(1 - \lambda)t + \lambda y \in [t, y]$), elle est donc croissante, ce qui nous donne $\varphi(x) \leq \varphi(y) = 0$, c'est-à-dire l'inégalité de convexité. ■

La caractérisation suivante de la convexité des fonctions dérivables est souvent utile.

Théorème 5.18 *Soit f une fonction dérivable sur I. Les propriétés suivantes sont équivalentes :*

1. *f est convexe sur I ;*
2. *la fonction dérivée f' est croissante sur I ;*
3. *la courbe représentative de f est située au dessus de sa tangente en tout point a de I.*

Démonstration. (1) $\Rightarrow$ (2) Si f est convexe et dérivable, pour $x < y$ dans I il existe u, v dans $\overset{\circ}{I}$ tels que $x < u < v < y$ et avec la croissance de τ_u et τ_v on a :

$$\tau_u(x) = \frac{f(x) - f(u)}{x - u} \leq \tau_u(v) = \frac{f(v) - f(u)}{v - u} = \tau_v(u) \leq \tau_v(y) = \frac{f(y) - f(v)}{y - v}$$

et faisant tendre respectivement u vers x et v vers y, on aboutit à $f'(x) \leq f'(y)$, ce qui prouve la croissance de f' sur I (on savait déjà que $f' = f'_g = f'_d$ est croissante sur $\overset{\circ}{I}$).

(2) $\Rightarrow$ (1) Si f' est croissante, la fonction f est continue sur I et convexe sur $\overset{\circ}{I}$, donc convexe sur I.

(2) $\Rightarrow$ (3) Si f' est croissante sur I, pour tout $a \in I$ l'application φ définie sur I par :

$$\varphi(x) = f(x) - f(a) - f'(a)(x - a)$$

est telle que $(x - a)\varphi'(x) = (x - a)(f'(x) - f'(a)) \geq 0$ pour tout $x \in I$, donc φ admet un minimum global en a et $\varphi(x) \geq \varphi(a) = 0$ pour tout $x \in I$, ce qui signifie que le graphe de f est au dessus de la tangente en a.

(3) $\Rightarrow$ (2) Pour $x < y$ dans I on a :

$$\begin{cases} f(x) \geq f(y) + f'(y)(x - y) \\ f(y) \geq f(x) + f'(x)(y - x) \end{cases}$$

qui par addition donne $(f'(y) - f'(x))(y - x) \geq 0$ équivalent à la croissance de f' sur I. ■

On retrouve le fait que, si f est deux fois dérivable sur I, elle est alors convexe sur I si et seulement $f''(x) \geq 0$ pour tout $x \in I$.

Exercice 5.7 *Soit q une fonction continue de $\mathbb{R}$ dans $\mathbb{R}^+$ non identiquement nulle. Montrer que l'unique solution à valeurs réelles bornée sur $\mathbb{R}$ de l'équation différentielle $y'' - qy = 0$ est la fonction nulle.*

Solution 5.7 *Le théorème de Cauchy-Lipschitz linéaire nous dit que l'ensemble des solution de l'équation différentielle $y'' - qy = 0$ est un espace vectoriel de dimension 2.*
On désigne par y une solution de cette équation.
Si q est la fonction nulle alors $y(x) = ax + b$ et les fonctions constantes sont les seules solutions bornées. On suppose donc que la fonction q n'est pas identiquement nulle.
La fonction $z = y^2$ est deux fois dérivable avec :

$$z'' = 2\left((y')^2 + yy''\right) = 2\left((y')^2 + qy^2\right) \geq 0$$

Il en résulte que la fonction z est convexe sur $\mathbb{R}$ et cette fonction est bornée si y l'est, elle est alors constante et avec y continue on déduit que la solution y est également constante, ce qui entraîne $qy = y'' = 0$ et $y = 0$ puisque q n'est pas la fonction nulle.

Exercice 5.8 *Soient q une fonction continue de $\mathbb{R}$ dans $\mathbb{R}^+$ et y une solution à valeurs réelles de l'équation différentielle $y'' - qy = 0$. Montrer que si y s'annule en deux points distincts alors elle est identiquement nulle.*

Solution 5.8 *La fonction $z = y^2$ est deux fois dérivable sur $\mathbb{R}$ avec :*

$$z'' = 2\left((y')^2 + yy''\right) = 2\left((y')^2 + qy^2\right) \geq 0$$

Il en résulte que cette fonction est convexe sur $\mathbb{R}$.
Si $y(x_1) = y(x_2)$ avec $x_1 < x_2$ alors $z(x_1) = z(x_2)$ et avec la convexité de la fonction z qui est à valeurs positives on déduit que z est nulle sur l'intervalle $[x_1, x_2]$. En effet, tout $x \in [x_1, x_2]$ s'écrit $x = (1-\lambda)x_1 + \lambda x_2$ avec $\lambda \in [0,1]$ et on a :

$$0 \leq f(x) \leq (1-\lambda) f(x_1) + \lambda f(x_2) = 0$$

On a donc $y(x) = 0$ pour tout $x \in [x_1, x_2]$. En particulier, pour $x \in]x_1, x_2[$ on a $y(x) = y'(x) = 0$ et le théorème de Cauchy-Lipschitz nous dit alors que $y = 0$.

5.4 Fonctions mid-convexes

Ici encore, I est un intervalle réel non réduit à un point et f est une application définie sur I à valeurs réelles.

Définition 5.3 *On dit que $f : I \to \mathbb{R}$ est mid-convexe si :*

$$\forall (x,y) \in I^2,\ f\left(\frac{x+y}{2}\right) \leq \frac{f(x) + f(y)}{2}$$

Exemple 5.7 *Une fonction convexe sur I est mid-convexe (prendre $\lambda = \frac{1}{2}$ dans l'inégalité de convexité).*

Lemme 5.1 *Si $f : I \to \mathbb{R}$ est mid-convexe, on a alors pour tout entier $p \geq 1$ et toute suite $(x_k)_{1 \leq k \leq 2^p}$ de points de I, on a :*

$$f\left(\frac{1}{2^p}\sum_{k=1}^{2^p} x_k\right) \leq \frac{1}{2^p}\sum_{k=1}^{2^p} f(x_k)$$

Démonstration. Le résultat est acquis pour $p = 1$.
En supposant acquis le résultat pour $p \geq 1$, on pour toute suite $(x_k)_{1 \leq k \leq 2^{p+1}}$ de points de I :

$$\begin{aligned}
f\left(\frac{1}{2^{p+1}}\sum_{k=1}^{2^{p+1}} x_k\right) &= f\left(\frac{1}{2}\left(\frac{1}{2^p}\sum_{k=1}^{2^p} x_k + \frac{1}{2^p}\sum_{k=1}^{2^p} x_{2^p+k}\right)\right) \\
&\leq \frac{1}{2}\left(f\left(\frac{1}{2^p}\sum_{k=1}^{2^p} x_k\right) + f\left(\frac{1}{2^p}\sum_{k=1}^{2^p} x_{2^p+k}\right)\right) \\
&\leq \frac{1}{2}\left(\frac{1}{2^p}\sum_{k=1}^{2^p} f(x_k) + \frac{1}{2^p}\sum_{k=1}^{2^p} f(x_{2^p+k})\right) = \frac{1}{2^{p+1}}\sum_{k=1}^{2^{p+1}} f(x_k)
\end{aligned}$$

■

Théorème 5.19 *La fonction f est mid-convexe si, et seulement si, pour tout entier $n \geq 2$ et toute suite $(x_k)_{1\leq k\leq n}$ de points de I, on a :*

$$f\left(\frac{x_1+\cdots+x_n}{n}\right) \leq \frac{f(x_1)+\cdots+f(x_n)}{n} \tag{5.2}$$

Démonstration. Il est clair que la condition est suffisante.

Pour les entiers de la forme $n = 2^p$ avec $p \geq 1$, le résultat est acquis.

Si $n \geq 2$ est un entier qui n'est pas une puissance de 2, il existe alors un entier $p \geq 1$ tel que $2^{p-1} < n < 2^p$.

Si $(x_k)_{1\leq k\leq n}$ est une suite de points de I, en notant :

$$x_{n+1} = \cdots = x_{2^p} = \frac{1}{n}\sum_{k=1}^{n} x_k$$

on a :

$$\begin{aligned} f\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) &= f\left(\frac{1}{2^p}\left(\sum_{k=1}^{n} x_k + \frac{2^p-n}{n}\sum_{k=1}^{n} x_k\right)\right) \\ &= f\left(\frac{1}{2^p}\sum_{k=1}^{2^p} x_k\right) \leq \frac{1}{2^p}\sum_{k=1}^{2^p} f(x_k) = \frac{1}{2^p}\left(\sum_{k=1}^{n} f(x_k) + \frac{2^p-n}{2^p} f\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)\right) \end{aligned}$$

soit :

$$\left(1-\frac{2^p-n}{2^p}\right) f\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) = \frac{n}{2^p} f\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) \leq \frac{1}{2^p}\sum_{k=1}^{n} f(x_k)$$

ou encore :

$$f\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right) \leq \frac{1}{n}\sum_{k=1}^{n} f(x_k)$$

■

Corollaire 5.4 *La fonction f est mid-convexe si, et seulement si :*

$$\forall (x,y) \in I^2,\ \forall \lambda \in [0,1] \cap \mathbb{Q},\ f((1-\lambda)x + \lambda y) \leq (1-\lambda) f(x) + \lambda f(y)$$

Démonstration. Il est clair que la condition est suffisante (prendre $\lambda = \frac{1}{2}$).

Réciproquement, soient $f : I \to \mathbb{R}$ mid-convexe et $(x,y) \in I^2$.

Pour $\lambda = 0$ ou $\lambda = 1$, on a toujours une égalité.

Si $\lambda \in]0,1[\cap \mathbb{Q}$, il s'écrit $\lambda = \frac{p}{q}$, où p, q sont des entiers tels que $1 \leq p < q$ et on a :

$$f((1-\lambda)x + \lambda y) = f\left(\frac{q-p}{q}x + \frac{p}{q}y\right) = f\left(\frac{1}{q}\sum_{k=1}^{q} x_k\right)$$

en notant :

$$x_1 = \cdots = x_{q-p} = x \text{ et } x_{q-p+1} = \cdots = x_q = y$$

ce qui nous donne :

$$f((1-\lambda)x + \lambda y) \leq \frac{1}{q}\sum_{k=1}^{q} f(x_k) = \frac{q-p}{q} f(x) + \frac{p}{q} f(y) = (1-\lambda) f(x) + \lambda f(y)$$

■

Théorème 5.20 *Une fonction $f : I \to \mathbb{R}$ continue est convexe si, et seulement si, elle est mid-convexe.*

Démonstration. Si λ est un réel compris entre 0 et 1, on a $\lambda = \lim\limits_{n\to+\infty} \lambda_n$, où $\lambda_n \in [0,1] \cap \mathbb{Q}$ (densité de $\mathbb{Q}$ dans $\mathbb{R}$) et avec la continuité de f on déduit que, pour $x < y$ dans I, on a :

$$\begin{aligned} f\left((1-\lambda)x + \lambda y\right) &= \lim_{n\to+\infty} f\left((1-\lambda_n)x + \lambda_n y\right) \\ &\leq \lim_{n\to+\infty} \left((1-\lambda_n) f(x) + \lambda_n f(y)\right) \\ &\leq (1-\lambda) f(x) + \lambda f(y) \end{aligned}$$

ce qui prouve la convexité de f. ∎

La caractérisation précédente est encore valable pour les fonctions monotones.

Théorème 5.21 *Une fonction $f : I \to \mathbb{R}$ monotone est convexe si, et seulement si, elle est mid-convexe.*

Démonstration. Supposons que f soit croissante.

En utilisant les approximation décimales par excès tout réel $\lambda \in]0,1[$ peut s'écrire comme limite d'une suite décroissante $(\lambda_n)_{n\in\mathbb{N}}$ de rationnels dans $]0,1[$ et pour tous $x < y$ dans I, on a :

$$f\left((1-\lambda)x + \lambda y\right) = f\left(x + \lambda(y-x)\right) \leq f\left(x + \lambda_n(y-x)\right) \leq (1-\lambda_n) f(x) + \lambda_n f(y)$$

ce qui nous donne en faisant n vers l'infini, $f\left((1-\lambda)x + \lambda y\right) \leq (1-\lambda) f(x) + \lambda f(y)$.

Pour f décroissante, on utilise les approximation décimales par défaut qui nous donnent une suite croissante $(\lambda_n)_{n\in\mathbb{N}}$ de rationnels dans $]0,1[$ qui converge vers λ et on a :

$$f\left((1-\lambda)x + \lambda y\right) = f\left(x + \lambda(y-x)\right) \leq f\left(x + \lambda_n(y-x)\right) \leq (1-\lambda_n) f(x) + \lambda_n f(y)$$

∎

5.5 Inégalités de convexité

5.5.1 Quelques inégalités classiques

Avec la stricte convexité de la fonction exponentielle, on déduit que :

$$\forall x \in \mathbb{R},\ e^x \geq x + 1$$

l'égalité étant réalisée uniquement pour $x = 0$. Cette inégalité traduit le fait que le graphe de la fonction est strictement au dessus de la tangente en 0 d'équation $y = 1 + x$.

On peut également déduire l'inégalité :

$$e^{\frac{a+b}{2}} \leq \frac{e^a + e^b}{2}$$

l'égalité étant réalisée si, et seulement si, $a = b$.

Avec la stricte concavité de la fonction sin sur $\left[0, \frac{\pi}{2}\right]$ on déduit que :

$$\forall x \in \left]0, \frac{\pi}{2}\right[,\ \frac{2}{\pi}x < \sin(x) < x$$

L'inégalité de gauche (inégalité de Jordan) traduit le fait que le graphe de la fonction est strictement au dessus de la corde définie par les points $A = (0, 0)$ et $B = \left(\frac{\pi}{2}, 1\right)$, cette corde ayant pour équation $y = \frac{2}{\pi}x$ et l'inégalité de droite traduit le fait que le graphe de la fonction est strictement sous la tangente en 0, cette tangente ayant pour équation $y = x$.

Avec la stricte concavité de la fonction logarithme on déduit que :

$$\forall x \in \mathbb{R}^{+,*},\ \ln(x) \leq x - 1$$

l'égalité étant uniquement réalisée pour $x = 1$. Cette inégalité traduit le fait que le graphe de la fonction est strictement au dessous de la tangente en 1, cette tangente ayant pour équation $y = x - 1$.

On en déduit également le résultat suivant qui peut être utilisé pour montrer l'inégalité de Hölder.

Lemme 5.2 *Si p, q sont deux réels strictement positifs tels que $\frac{1}{p} + \frac{1}{q} = 1$, alors :*

$$\forall u \in \mathbb{R}^{+,*},\ \forall v \in \mathbb{R}^{+,*},\ u^{\frac{1}{p}} v^{\frac{1}{q}} \leq \frac{1}{p}u + \frac{1}{q}v$$

Démonstration. La concavité de la fonction ln sur $\mathbb{R}^{+,*}$ nous dit que pour tout réel $\lambda \in [0, 1]$ et tous réels u, v strictement positifs on a :

$$\ln(\lambda u + (1 - \lambda) v) \geq \lambda \ln(u) + (1 - \lambda) \ln(v) = \ln\left(u^{\lambda} v^{1-\lambda}\right)$$

et avec la croissance de la fonction exp sur $\mathbb{R}^{+,*}$ on en déduit que :

$$u^{\lambda} v^{1-\lambda} \leq \lambda u + (1 - \lambda) v$$

Cette inégalité étant également vérifiée pour $u = 0$ ou $v = 0$.

En prenant $\lambda = \frac{1}{p}$ dans l'inégalité précédente, on a $1 - \lambda = \frac{1}{q}$ et :

$$u^{\frac{1}{p}} v^{\frac{1}{q}} \leq \frac{1}{p}u + \frac{1}{q}v$$

■

Cette inégalité généralise l'inégalité classique $\sqrt{uv} \leq \frac{1}{2}(u + v)$ qui est une conséquence de l'inégalité $\left(\sqrt{u} - \sqrt{v}\right)^2 \geq 0$.

Cette inégalité de convexité peut également s'écrire :

$$\forall (x, y) \in \left(\mathbb{R}^{+,*}\right)^2,\ xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q$$

5.5.2 Les inégalités de Jensen

Définition 5.4 *Étant donnée une suite finie $(x_i)_{1 \leq i \leq n}$ de points d'un intervalle I, on dit qu'un réel x est combinaison linéaire convexe des x_i, pour i compris entre 1 et n, s'il existe une suite $(\lambda_i)_{1 \leq i \leq n}$ de réels positifs ou nuls telle que :*

$$\sum_{i=1}^{p} \lambda_i = 1 \text{ et } x = \sum_{i=1}^{p} \lambda_i x_i$$

Une combinaison linéaire convexe de points de I est dans I, puisque I est convexe.

Théorème 5.22 (Jensen) *Une fonction $f : I \to \mathbb{R}$ est convexe si, et seulement si, pour toute combinaison linéaire convexe $\sum_{i=1}^{n} \lambda_i x_i$ de points de I, on a :*

$$f\left(\sum_{i=1}^{p} \lambda_i x_i\right) \leq \sum_{i=1}^{p} \lambda_i f(x_i)$$

Démonstration. Il est clair que la condition est suffisante (prendre $p = 2$).

Pour la réciproque, on procède par récurrence sur $n \geq 1$.

Pour $n = 1$ le résultat est trivial et pour $n = 2$ il s'agit de la définition d'une fonction convexe.

Supposons le résultat acquis pour $n \geq 2$ et soit $\sum_{i=1}^{n+1} \lambda_i x_i$ une combinaison linéaire convexe de points de I. On note $\lambda = \sum_{i=1}^{n} \lambda_i$.

Si $\lambda = 0$, alors tous les λ_i, pour $1 \leq i \leq n$, sont nuls, $\lambda_{n+1} = 1$ et l'inégalité est trivialement une égalité.

Si $\lambda \neq 0$, en notant $x' = \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} x_i$, on a $\sum_{i=1}^{n} \frac{\lambda_i}{\lambda} = 1$, $x' \in I$ et :

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) = f(\lambda x' + \lambda_{n+1} x_{n+1}) \leq \lambda f(x') + \lambda_{n+1} f(x_{n+1})$$

puisque $\lambda \geq 0$, $\lambda_{n+1} \geq 0$ et $\lambda + \lambda_{n+1} = 1$.

Avec l'hypothèse de récurrence on déduit alors que :

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leq \lambda \sum_{i=1}^{n} \frac{\lambda_i}{\lambda} f(x_i) + \lambda_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i f(x_i)$$

■

Cette inégalité de convexité peut aussi être utile sous la forme suivante.

Corollaire 5.5 *Si f est une fonction convexe définie sur un intervalle I, alors pour toute combinaison linéaire à coefficients réels positifs non tous nuls $\sum_{i=1}^{p} \lambda_i x_i$ d'éléments de I, on a en notant $\lambda = \sum_{i=1}^{p} \lambda_i$:*

$$f\left(\frac{1}{\lambda} \sum_{i=1}^{p} \lambda_i x_i\right) \leq \frac{1}{\lambda} \sum_{i=1}^{p} \lambda_i f(x_i)$$

En utilisant les sommes de Riemann on en déduit le résultat suivant qui est la version « continue » de l'inégalité précédente.

Théorème 5.23 (Jensen) *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction convexe. Pour toute fonction continue $\varphi : [a, b] \to \mathbb{R}$ (avec $a < b$), on a :*

$$f\left(\frac{1}{b-a} \int_a^b \varphi(t)\, dt\right) \leq \frac{1}{b-a} \int_a^b f \circ \varphi(t)\, dt$$

Démonstration. On remarque tout d'abord que f convexe sur $\mathbb{R}$ est continue, il en donc de même de $f \circ \varphi$ et cette fonction est bien intégrable sur $[a, b]$.

Avec la convexité de f, on peut écrire que pour tout entier $n \geq 1$ on a :

$$f\left(\frac{1}{n}\sum_{k=1}^{n}\varphi\left(a+k\frac{b-a}{n}\right)\right) \leq \frac{1}{n}\sum_{k=1}^{n} f\circ\varphi\left(a+k\frac{b-a}{n}\right)$$

qui par passage à la limite quand n tend vers l'infini donne, compte tenu de la continuité de f et de la définition de l'intégrale d'une fonction continue sur $[a, b]$ comme limite des sommes de Riemann :

$$f\left(\frac{1}{b-a}\int_a^b \varphi(t)\,dt\right) \leq \frac{1}{b-a}\int_a^b f\circ\varphi(t)\,dt$$

■

5.5.3 L'inégalité de Hölder

Théorème 5.24 (Hölder) *Soient p, q deux réels strictement positifs tels que $\frac{1}{p}+\frac{1}{q}=1$. Pour tous vecteurs x, y dans $\mathbb{C}^n$ on a :*

$$\left|\sum_{i=1}^{n} x_i\overline{y_i}\right| \leq \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}|y_i|^q\right)^{\frac{1}{q}}$$

Démonstration. En utilisant la convexité de la fonction $t \mapsto t^p$ pour $p > 1$, on déduit que pour toute suite $(\lambda_i)_{1\leq i\leq n}$ de réels positifs non tous nuls et toute suite $(t_i)_{1\leq i\leq n}$ de réels strictement positifs, on a en notant $\lambda = \sum\limits_{i=1}^{n}\lambda_i$:

$$\left(\frac{1}{\lambda}\sum_{i=1}^{n}\lambda_i t_i\right)^p \leq \frac{1}{\lambda}\sum_{i=1}^{n}\lambda_i t_i^p$$

ce qui peut s'écrire, compte tenu de $1-\frac{1}{p}=\frac{1}{q}$:

$$\sum_{i=1}^{n}\lambda_i t_i \leq \left(\sum_{i=1}^{n}\lambda_i\right)^{\frac{1}{q}}\left(\sum_{i=1}^{n}\lambda_i t_i^p\right)^{\frac{1}{p}}$$

cette inégalité étant encore vérifiée si tous les λ_i et tous les t_i sont nuls.

Si x et y sont deux vecteurs dans $\mathbb{C}^n$, en prenant pour tout entier i compris entre 1 et n :

$$\lambda_i = |y_i|^q,\ t_i = |x_i|\,|y_i|^{-\frac{q}{p}}$$

on a $\lambda_i t_i = |x_i|\,|y_i|^{q\left(1-\frac{1}{p}\right)} = |x_i|\,|y_i|$ et $\lambda_i t_i^p = |x_i|^p\,|y_i|^{q-q} = |x_i|^p$, de sorte que l'inégalité précédente donne :

$$\left|\sum_{i=1}^{n}x_i\overline{y_i}\right| \leq \sum_{i=1}^{n}|x_i|\,|y_i| \leq \left(\sum_{i=1}^{n}|x_i|^p\right)^{\frac{1}{p}}\left(\sum_{i=1}^{n}|y_i|^q\right)^{\frac{1}{q}}$$

■

L'inégalité de Hölder s'écrit, en notant $\langle x \mid y\rangle$ le produit scalaire hermitien canonique de $\mathbb{C}^n$:

$$|\langle x \mid y\rangle| \leq \|x\|_p \|y\|_q$$

et cette inégalité est encore valable pour $p = 1$ et $q = +\infty$.

Pour $p = q = 2$ on retrouve l'inégalité de Cauchy-Schwarz.

Corollaire 5.6 *Pour tout réel $p \geq 1$ l'application*

$$x \mapsto \|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

définit une norme sur $\mathbb{C}^n$.

5.5.4 Comparaison des moyennes harmonique, géométrique et arithmétique

Pour toute suite finie $x = (x_i)_{1\leq i\leq n}$ de réels strictement positifs, on note :

$$A(x) = \frac{\sum\limits_{i=1}^{n} x_i}{n},\ G(x) = \left(\prod_{i=1}^{n} x_i\right)^{\frac{1}{n}},\ H(x) = \frac{n}{\sum\limits_{i=1}^{n} \frac{1}{x_i}}$$

les moyennes arithmétique, géométrique et harmonique de x.

En notant $y = \left(\dfrac{1}{x_i}\right)_{1\leq i\leq n}$ on peut remarquer que :

$$H(x) = \frac{1}{A(y)},\ G(x) = \frac{1}{G(y)}$$

De la concavité de la fonction logarithme on déduit le résultat suivant.

Théorème 5.25 *Pour toute suite finie $x = (x_i)_{1\leq i\leq n}$ de réels strictement positifs, on a :*

$$H(x) \leq G(x) \leq A(x)$$

Démonstration. La concavité de la fonction ln nous permet d'écrire :

$$\ln(G(x)) = \frac{1}{n}\sum_{i=1}^{n} \ln(x_i) \leq \ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \ln(A(x))$$

ce qui équivaut, du fait de la croissance de la fonction exp, à $G(x) \leq A(x)$.

Il en résulte que :

$$H(x) = \frac{1}{A(y)} \leq \frac{1}{G(y)} = G(x) \leq A(x)$$

■

Pour $n = 2$ on retrouve l'inégalité $\sqrt{x_1 x_2} \leq \dfrac{x_1 + x_2}{2}$ conséquence de la positivité de $\left(\sqrt{x_1} - \sqrt{x_2}\right)^2$.

Exercice 5.9 *Soit* $x = (x_i)_{1 \leq i \leq n}$ *une suite finie de réels strictement positifs telle que* $\prod_{i=1}^{n} x_i = 1$. *Montrer que* $\prod_{i=1}^{n} (2 + x_i) \geq 3^n$ *en précisant dans quel cas l'égalité est réalisée.*

Solution 5.9 *On a, pour tout* i *:*

$$2 + x_i = 1 + 1 + x_i = 3A(1, 1, x_i) \geq 3G(1, 1, x_i) = 3x_i^{\frac{1}{3}}$$

et :

$$\prod_{i=1}^{n} (2 + x_i) \geq 3^n \prod_{i=1}^{n} x_i^{\frac{1}{3}} = 3^n \left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{3}} = 3^n$$

l'égalité étant réalisée si, et seulement si, tous les x_i *valent* 1 *(si l'un des* x_i *vaut* 1*, on a* $2 + x_i > 3x_i^{\frac{1}{3}}$*).*

Exercice 5.10 *Déterminer, parmi toutes les ellipses d'aire donnée* $S > 0$*, celles de périmètre minimal.*

Solution 5.10 *Il nous suffit de considérer les ellipses de représentation paramétrique :*

$$\gamma : t \in [0, 2\pi] \mapsto (a \cos(t), b \sin(t))$$

avec $0 < b \leq a$.
Le périmètre d'une telle ellipse est :

$$\begin{aligned} L(a, b) &= \int_0^{2\pi} \|\gamma'(t)\| \, dt = \int_0^{2\pi} \sqrt{a^2 \sin^2(t) + b^2 \cos^2(t)} dt \\ &= 4 \int_0^{\frac{\pi}{2}} \sqrt{a^2 \sin^2(t) + b^2 \cos^2(t)} dt \end{aligned}$$

et l'aire est :

$$\begin{aligned} S(a, b) &= \int \int_{\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1} dx dx = 4b \int_0^a \left(\int_0^{\sqrt{1 - \frac{x^2}{a^2}}} dy \right) dx \\ &= 4b \int_0^a \sqrt{1 - \frac{x^2}{a^2}} dx = 4ab \int_0^1 \sqrt{1 - t^2} dt \\ &= 4ab \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2(\theta)} \cos(\theta) \, d\theta = 4ab \int_0^{\frac{\pi}{2}} \cos^2(\theta) \, d\theta = \pi ab \end{aligned}$$

Il s'agit alors de trouver le minimum de la fonction L*, pour* $0 < b \leq a$ *tels que* $\pi ab = S$.
En utilisant la concavité de la fonction $x \mapsto \sqrt{x}$*, on a en notant* $\lambda = \sin^2(t)$ *:*

$$\begin{aligned} \sqrt{a^2 \sin^2(t) + b^2 \cos^2(t)} &= \sqrt{\lambda a^2 + (1 - \lambda) b^2} \\ &\geq \lambda \sqrt{a^2} + (1 - \lambda) \sqrt{b^2} \\ &\geq \lambda a + (1 - \lambda) b = a \sin^2(t) + b \cos^2(t) \end{aligned}$$

et :

$$\begin{aligned} L(a, b) &\geq 4 \int_0^{\frac{\pi}{2}} \left(a \sin^2(t) + b \cos^2(t) \right) dt = \pi (a + b) \\ &\geq 2\pi \sqrt{ab} = 2\sqrt{\pi S} \end{aligned}$$

Mais pour $a = b$ *et* $\pi ab = \pi a^2 = S$*, on a* $L(a, a) = 2a\pi = 2\sqrt{\pi S}$.
Le minimum est donc atteint pour le cercle de rayon a.

6

Fonctions exponentielle et logarithme

6.1 La méthode d'Euler pour l'équation différentielle $y' = y$

La méthode d'Euler est basée sur le théorème des accroissements finis qui permet d'écrire, pour une fonction f de classe $\mathcal{C}^1$ dans un voisinage du point x :

$$f(x+h) = f(x) + hf'(x+\theta h) \backsimeq f(x) + hf'(x)$$

où h est réel non nul destiné à tendre vers 0 et θ (qui dépend de x et h) est dans $]0,1[$.

On s'intéresse ici au problème de Cauchy qui consiste à déterminer une fonction y de classe $\mathcal{C}^1$ de $\mathbb{R}$ dans $\mathbb{R}$ et telle que :

$$\begin{cases} y(0) = 1 \\ \forall x \in \mathbb{R},\ y'(x) = y(x) \end{cases} \tag{6.1}$$

En supposant que ce problème admet une solution (ce que nous dit le théorème de Cauchy-Lipschitz linéaire) on se fixe un réel $x > 0$ et on utilise la méthode d'Euler pour obtenir une valeur approchée de $y(x)$. Pour ce faire on se donne un entier $n \geq 1$ auquel on associe la subdivision de $[0, x]$ définie par les points :

$$x_k = k\frac{x}{n}\ \ (0 \leq k \leq n)$$

et on définit une suite $(y_k)_{0 \leq k \leq n}$ d'approximations des éléments de la suite $(y(x_k))_{0 \leq k \leq n}$.

Pour $k = 0$, on pose $y_0 = y(x_0) = 1$.

En supposant construits $y_0, \ldots, y_{k-1}$ pour $1 \leq k \leq n$, on écrit que :

$$y(x_k) = y\left(x_{k-1} + \frac{x}{n}\right) \backsimeq y(x_{k-1}) + \frac{x}{n}y'(x_{k-1}) = y(x_{k-1})\left(1 + \frac{x}{n}\right)$$

et avec $y(x_{k-1}) \backsimeq y_{k-1}$, on obtient l'approximation :

$$y(x_k) \backsimeq y_k = y_{k-1}\left(1 + \frac{x}{n}\right)$$

La suite $(y_k)_{0 \leq k \leq n}$ est donc définie par la relation de récurrence :

$$\begin{cases} y_0 = 1 \\ y_k = y_{k-1}\left(1 + \dfrac{x}{n}\right)\ \ (1 \leq k \leq n) \end{cases}$$

ce qui nous donne $y_k = \left(1 + \dfrac{x}{n}\right)^k$ et en particulier on a l'approximation :

$$y(x) = y(x_n) \backsimeq y_n = \left(1 + \frac{x}{n}\right)^n$$

6.2 La suite de fonctions $\left(\left(1+\frac{x}{n}\right)^n\right)_{n\geq 1}$

Nous sommes donc amené à étudier la suite de fonctions $(u_n)_{n\geq 1}$ définie par :

$$\forall n \geq 1,\ \forall x \in \mathbb{R},\ u_n(x) = \left(1+\frac{x}{n}\right)^n$$

et nous allons montrer dans ce paragraphe que cette suite converge vers une fonction f qui est l'unique solution du problème (6.1).

Pour $x = 0$, cette suite est stationnaire sur 1.

Le lemme suivant nous sera utile.

Lemme 6.1 *Pour tout réel x, on a :*

$$\lim_{n\to+\infty} u_n(x)\, u_n(-x) = 1$$

Démonstration. Pour tout $n \geq 1$, on a :

$$1 - u_n(x)\, u_n(-x) = 1 - \left(1-\frac{x^2}{n^2}\right)^n = \frac{x^2}{n^2}\sum_{k=0}^{n-1}\left(1-\frac{x^2}{n^2}\right)^k$$

Pour tout $n > |x|$, on a $0 < 1 - \frac{x^2}{n^2} \leq 1$ et :

$$|1 - u_n(x)\, u_n(-x)| \leq \frac{x^2}{n^2} n = \frac{x^2}{n} \underset{n\to+\infty}{\to} 0$$

■

En notant E la fonction partie entière, on associe à tout réel x l'entier n_x défini par :

$$n_x = \begin{cases} 1 \text{ pour } x \geq 0 \\ E(|x|) + 1 \text{ pour } x < 0 \end{cases}$$

et on a $u_n(x) > 0$ pour tout $n \geq n_x$.

Nous allons montrer dans ce qui suit que pour tout réel x, la suite $(u_n(x))_{n\geq 1}$ converge vers un réel $f(x) > 0$.

En utilisant le lemme précédent on voit qu'il suffit de montrer ce résultat pour $x > 0$ ou pour $x < 0$.

Lemme 6.2 *Pour tout entier $n_0 \geq 1$ et tout entier $n \geq n_0$, la fonction $\frac{u_{n+1}}{u_n}$ est strictement croissante sur $\mathbb{R}^+$ et strictement décroissante sur $]-n_0, 0]$.*

Démonstration. On se fixe un entier $n_0 \geq 1$.

Pour tout $n \geq n_0$ et tout $x \in \left]-n_0, 0\right]$, on a $n \geq n_0 > -x$ et $u_n(x) > 0$, cette inégalité étant également vérifiée pour $x > 0$ et $n \geq 1$.

Pour $n \geq n_0$ la restriction de la fonction u_n à $]-n_0, +\infty[$ est dérivable à valeurs strictement positives avec $\frac{u_n'(x)}{u_n(x)} = \frac{n}{n+x}$ et :

$$\forall x \in \left]-n_0, +\infty\right[,\ \frac{u_{n+1}'(x)}{u_{n+1}(x)} - \frac{u_n'(x)}{u_n(x)} = \frac{x}{(n+x)(n+1+x)}$$

En utilisant $\left(\frac{u_{n+1}}{u_n}\right)' = \frac{u_{n+1}'u_n - u_{n+1}u_n'}{u_n^2}$, on en déduit que :

$$\begin{cases} \forall x > 0, \ \left(\frac{u_{n+1}}{u_n}\right)'(x) > 0 \\ \forall x \in \left]-n_0, 0\right[, \ \left(\frac{u_{n+1}}{u_n}\right)'(x) < 0 \end{cases}$$

c'est-à-dire que pour $n \geq n_0$, la fonction $\frac{u_{n+1}}{u_n}$ est strictement croissante sur $\mathbb{R}^+$ et strictement décroissante sur $\left]-n_0, 0\right]$. ∎

Lemme 6.3 *Pour tout réel x la suite $(u_n(x))_{n\geq n_x}$ est à valeurs strictement positives et pour tout entier $n_0 \geq 1$, tout réel non nul x dans $\left]-n_0, +\infty\right[$, la suite $(u_n(x))_{n\geq n_0}$ est strictement croissante.*

Démonstration. Par définition de n_x, on a $u_n(x) > 0$ pour tout réel x et tout entier $n \geq n_x$. Pour x non nul dans $\left]-n_0, +\infty\right[$ le lemme précédent nous dit que :

$$\forall n \geq n_0, \ \frac{u_{n+1}(x)}{u_n(x)} > \frac{u_{n+1}(0)}{u_n(0)} = 1$$

c'est-à-dire que la suite $(u_n(x))_{n\geq n_0}$ est strictement croissante. ∎

Remarque 6.1 *La croissance de la suite $(u_n(x))_{n\geq n_x}$ peut aussi se montrer directement en utilisant l'inégalité de Bernoulli à savoir : $(1-a)^n > 1 - na$ pour $a < 1$ non nul.*
Pour $n \geq n_x$ on a :

$$\begin{aligned} u_{n+1}(x) &= \left(1+\frac{x}{n+1}\right)^{n+1} = \left(1+\frac{x}{n}\right)^{n+1}\left(\frac{1+\frac{x}{n+1}}{1+\frac{x}{n}}\right)^{n+1} \\ &= \left(1+\frac{x}{n}\right)^{n+1}\left(\frac{1+x+n}{1+x+n+\frac{x}{n}}\right)^{n+1} = \left(1+\frac{x}{n}\right)^{n+1}\left(1-\frac{x}{n^2+(1+x)n+x}\right)^{n+1} \\ &= \left(1+\frac{x}{n}\right)^{n+1}\left(1-\frac{x}{(n+1)(n+x)}\right)^{n+1} \end{aligned}$$

Pour $x < 0$, on a $\frac{x}{(n+1)(n+x)} < 0$ et pour $x > 0$, $\frac{x}{(n+1)(n+x)} < 1$ (c'est équivalent à $n^2 + (1+x)n > 0$), on peut donc utiliser l'inégalité de Bernoulli pour écrire que :

$$\left(1-\frac{x}{(n+1)(n+x)}\right)^{n+1} > 1 - \frac{x}{n+x} = \frac{n}{n+x}$$

et donc :

$$\frac{u_{n+1}(x)}{u_n(x)} = \left(1+\frac{x}{n}\right)\left(1-\frac{x}{(n+1)(n+x)}\right)^{n+1} > \left(1+\frac{x}{n}\right)\frac{n}{n+x} = 1$$

Théorème 6.1 *Pour tout réel x, la suite $(u_n(x))_{n\geq 1}$ converge vers un réel $f(x)$.*
De plus on a $f(0) = 1$, $f(x) > 0$ et $f(-x) = \frac{1}{f(x)}$ pour tout réel x.

Démonstration. Pour $x = 0$ on a $u_n(0) = 1$ pour tout $n \geq 1$ et $\lim\limits_{n \to +\infty} u_n(0) = 1 = f(0)$.

Pour $x < 0$, on a $0 < 1 + \dfrac{x}{n} < 1$ et $0 < u_n(x) < 1$ pour tout $n \geq n_x$, c'est-à-dire que la suite $(u_n(x))_{n \geq n_x}$ est bornée et comme par ailleurs elle est croissante à partir d'un certain rang n_0, on en déduit qu'elle est convergente. On note $f(x) = \lim\limits_{n \to +\infty} u_n(x)$.

De la stricte croissance de $(u_n(x))_{n \geq n_0}$, on déduit que $f(x) > u_{n_0}(x) > 0$.

Enfin pour $x > 0$, on a :

$$u_n(x) = \frac{u_n(x)\, u_n(-x)}{u_n(-x)} \underset{n \to +\infty}{\to} \frac{1}{f(-x)}$$

(lemme 6.1), soit $\lim\limits_{n \to +\infty} u_n(x) = \dfrac{1}{f(-x)} = f(x)$. ■

Remarque 6.2 *De $f(x) f(-x) = 1$ pour tout réel x, on déduit que f n'est pas une fonction polynomiale.*
En effet si f est polynomiale de degré $n \geq 0$ (f n'est pas nulle), alors $f(x) f(-x)$ est polynomiale de degré $2n$ et constante, on a donc $n = 0$ et f est constante, ce qui est incompatible avec $f(x) > u_1(x) = 1 + x$ pour $x > 0$.

Remarque 6.3 *Avec :*

$$f(x) = \frac{1}{f(-x)} = \lim_{n \to +\infty} \frac{1}{u_n(-x)}$$

on déduit que $f(x)$ est aussi limite de la suite $(v_n(x))_{n \geq n_{-x}}$ définie par :

$$\forall n \geq n_{-x},\ v_n(x) = \frac{1}{u_n(-x)} = \left(1 - \frac{x}{n}\right)^{-n}$$

Pour x non nul, il existe un entier n_0 tel que $(u_n(x))_{n \geq n_0}$ est strictement croissante, $(v_n(x))_{n \geq n_0}$ est strictement décroissante et en conséquence :

$$\forall n \geq n_0,\ u_n(x) < f(x) < v_n(x)$$

Lemme 6.4 *Pour tout $x \in]-1, 1[$, on a :*

$$1 + x \leq f(x) \leq \frac{1}{1 - x}$$

Démonstration. Pour $n_0 = 1$ et $x \in]-1, +\infty[$ on a vu précédemment que la suite $(u_n(x))_{n \geq 1}$ est strictement croissante et donc $u_1(x) = 1 + x \leq f(x)$.

Pour $x \in]-1, 1[$, $-x$ est aussi dans $]-1, 1[$ et on a $u_1(-x) = 1 - x \leq f(-x)$, ce qui donne $f(x) = \dfrac{1}{f(-x)} \leq \dfrac{1}{1 - x}$. ■

Lemme 6.5 *La fonction f est continue en* 0.

Démonstration. Du lemme précédent on déduit immédiatement que $\lim\limits_{x \to 0} f(x) = 1 = f(0)$, ce qui signifie que f est continue en 0. ■

Lemme 6.6 *Pour tout suite réelle $(x_n)_{n \in \mathbb{N}}$ convergente vers un réel x, on a $\lim\limits_{n \to +\infty} u_n(x_n) = f(x)$.*

Démonstration. On a, pour tout $n \in \mathbb{N}$:

$$u_n(x_n) = \frac{u_n(x_n)\,u_n(-x)}{u_n(-x)}$$

avec :

$$u_n(x_n)\,u_n(-x) = \left(\left(1+\frac{x_n}{n}\right)\left(1-\frac{x}{n}\right)\right)^n = \left(1+\frac{\varepsilon_n}{n}\right)^n = u_n(\varepsilon_n)$$

où :

$$\varepsilon_n = x_n - x - \frac{xx_n}{n} \underset{n\to+\infty}{\to} 0$$

Pour n assez grand, on a $\varepsilon_n \in \left]-1,+\infty\right[$ de sorte que la suite $(u_m(\varepsilon_n))_{m\geq 1}$ est croissante et donc :

$$u_1(\varepsilon_n) = 1+\varepsilon_n \leq u_n(\varepsilon_n) \leq f(\varepsilon_n)$$

avec $\lim\limits_{n\to+\infty} u_1(\varepsilon_n) = 1$ et $\lim\limits_{n\to+\infty} f(\varepsilon_n) = 1$ (continuité de f en 0), ce qui implique $\lim\limits_{n\to+\infty} u_n(x_n) = 1$.

On a donc en définitive $\lim\limits_{n\to+\infty} u_n(x_n)\,u_n(-x) = 1$ et $\lim\limits_{n\to+\infty} u_n(x_n) = \dfrac{1}{f(-x)} = f(x)$. ■

Théorème 6.2 *Pour tous x, y dans $\mathbb{R}$ on a $f(x+y) = f(x)\,f(y)$.*

Démonstration. Pour x, y dans $\mathbb{R}$ et $n \geq 1$, on a :

$$u_n(x)\,u_n(y) = \left(\left(1+\frac{x}{n}\right)\left(1+\frac{y}{n}\right)\right)^n = \left(1+\frac{z_n}{n}\right)^n = u_n(z_n)$$

avec $z_n = x+y+\dfrac{xy}{n} \underset{n\to+\infty}{\to} x+y$. On déduit alors du lemme précédent que :

$$f(x)\,f(y) = \lim_{n\to+\infty} u_n(x)\,u_n(y) = \lim_{n\to+\infty} u_n(z_n) = f(x+y)$$

■

Corollaire 6.1 *La fonction f est continue sur $\mathbb{R}$.*

Démonstration. Pour x, y dans $\mathbb{R}$ on a :

$$f(y) - f(x) = f(x+y-x) - f(x) = f(x)\,(f(y-x)-1) \underset{y\to x}{\to} 0$$

en utilisant la continuité en 0 de f. ■

Remarque 6.4 *La continuité de f peut aussi se montrer en utilisant le fait que f est convexe. En effet pour tout $n_0 \geq 1$ et tout $n \geq n_0$ la fonction u_n est convexe sur $\left]-n_0,+\infty\right[$ (sa dérivée seconde est positive sur cet intervalle), donc $f = \lim\limits_{n\to+\infty} u_n$ est aussi convexe sur cet intervalle. Comme n_0 est quelconque, la fonction f est convexe sur $\mathbb{R}$. Sachant qu'une fonction convexe sur un intervalle est continue sur l'intérieur, on déduit que f est continue sur $\mathbb{R}$.*

En utilisant le lemme de Dini, on peut montrer que la suite $(u_n)_{n\geq 1}$ converge uniformément vers f sur tout intervalle compact $[a,b]$.

Théorème 6.3 *Si $I = [a,b]$ est un intervalle réel compact, alors la suite $(u_n)_{n\geq 1}$ converge uniformément vers f sur I.*

Démonstration. On peut trouver un entier n_0 tel que $I = [a, b] \subset]-n_0, +\infty[$ et pour tout x dans I la suite $(u_n(x))_{n \geq n_0}$ est croissante. En se restreignant à I, on a donc une suite croissante $(u_n)_{n \geq n_0}$ de fonctions continues sur I qui converge simplement sur I vers une fonction continue f, le théorème de Dini nous dit alors que la convergence est uniforme sur I. ∎

Remarque 6.5 *Sachant qu'une limite uniforme sur $\mathbb{R}$ d'une suite de fonctions polynomiales est nécessairement polynomiale (exercice classique), on en déduit que la suite $(u_n)_{n \geq 1}$ ne peut converger uniformément sur $\mathbb{R}$ vers f puisque cette dernière n'est pas polynomiale.*

Théorème 6.4 *La fonction f est l'unique solution sur $\mathbb{R}$ du problème de Cauchy $y' = y$ avec la condition initiale $y(0) = 1$.*

Démonstration. On a pour tout $x \in]0, 1[$:

$$1 \leq \frac{f(x) - 1}{x} \leq \frac{1}{1 - x}$$

et pour tout $x \in]-1, 0[$:

$$1 \geq \frac{f(x) - 1}{x} \geq \frac{1}{1 - x}$$

(lemme 6.4), donc $\lim\limits_{x \to 0} \frac{f(x) - 1}{x} = 1$, ce qui signifie que f est dérivable en 0 de dérivée égale à 1.

Puis avec :

$$\frac{f(x + h) - f(x)}{h} = f(x) \frac{f(h) - 1}{h}$$

pour tout $x \in \mathbb{R}$ et tout $h \in \mathbb{R}^*$, on déduit que $\lim\limits_{h \to 0} \frac{f(x + h) - f(x)}{h} = f(x)$, ce qui signifie que f est dérivable en x avec $f'(x) = f(x)$. Comme $f(0) = 1$, la fonction f est bien solution du problème de Cauchy.

Si y est une autre solution, la fonction z définie sur $\mathbb{R}$ par $z(x) = y(x) f(-x)$ est telle que $z' = 0$ avec $z(0) = 0$, c'est donc la fonction nulle et nécessairement $y(x) = \frac{1}{f(-x)} = f(x)$ pour tout réel x. ∎

De ce résultat on déduit que la fonction f est indéfiniment dérivable sur $\mathbb{R}$ avec $f^{(n)} = f$ pour tout entier naturel n.

De l'équation fonctionnelle vérifiée par la fonction f, on déduit facilement par récurrence que, pour tout réel non nul a, on a $f(n \cdot a) = (f(a))^n$ pour tout entier naturel n, puis avec $f(-x) = \frac{1}{f(x)}$, on déduit que cette relation est valable pour tout entier relatif n. Si $r = \frac{p}{q}$ est un entier relatif, on a alors $f(r) = f\left(p\frac{1}{q}\right) = \left(f\left(\frac{1}{q}\right)\right)^p$ et avec $f(1) = f\left(q\frac{1}{q}\right) = \left(f\left(\frac{1}{q}\right)\right)^q$, on déduit que $f\left(\frac{1}{q}\right) = (f(1))^{\frac{1}{q}}$ et $f(r) = (f(1))^{\frac{p}{q}} = (f(1))^r$.

En notant $e = f(1)$, on a donc $f(r) = e^r$ pour tout rationnel r, ce qui nous conduit à noter $f(x) = e^x$ pour tout réel x et la fonction ainsi définie est appelée fonction exponentielle réelle. On la note aussi $f(x) = \exp(x)$.

6.3 Propriétés de la fonction exponentielle réelle

On sait déjà que la fonction exp est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, à valeurs strictement positives et solution de l'équation fonctionnelle $f(x+y) = f(x) f(y)$.

De $\exp'(x) = \exp(x) > 0$ pour tout réel x, on déduit que exp est strictement croissante sur $\mathbb{R}$ et pour tout $x > 0$, on a $e^x > e^0 = 1$.

De $\exp''(x) = \exp(x) > 0$ pour tout réel x, on déduit que exp est convexe.

Avec $\exp(x) > u_1(x) = 1 + x$ pour $x > 0$, on déduit que $\lim\limits_{x \to +\infty} \exp(x) = +\infty$ et avec $\exp(-x) = \dfrac{1}{\exp(x)}$, on déduit que $\lim\limits_{x \to -\infty} \exp(x) = 0$.

La fonction exp réalise donc un homéomorphisme de $\mathbb{R}$ sur $]0, +\infty[$. Son inverse est appelé fonction logarithme et noté ln.

Cette fonction ln est dérivable sur $]0, +\infty[$ de dérivée égale à $\dfrac{1}{t}$ (dérivation de la fonction réciproque). La fonction ln est donc la primitive sur $]0, +\infty[$ de la fonction $\dfrac{1}{t}$ nulle en 1.

Avec :

$$\frac{\exp(x)}{x^n} > \frac{u_{n+1}(x)}{x^n} = \left(\frac{1}{x} + \frac{1}{n+1}\right)^n \left(1 + \frac{x}{n+1}\right)$$

pour $x > 0$ et $n \geq 1$, on déduit que $\lim\limits_{x \to +\infty} \dfrac{\exp(x)}{x^n} = +\infty$.

6.4 Extension aux espaces de Banach

La définition précédente de la fonction exponentielle peut être étendue aux algèbres de Banach.

On rappelle que si E est une algèbre unitaire sur $\mathbb{R}$, on dit que c'est une algèbre de Banach si elle est munie d'une norme multiplicative $x \mapsto \|x\|$ (c'est-à-dire que $\|\cdot\|$ est une norme sur E et $\|xy\| \leq \|x\| \|x\|$ pour tous x, y dans E) telle que $\|1\| = 1$ (où 1 est le neutre pour la multiplication) et si $(E, \|\cdot\|)$ est un espace de Banach. On peut considérer par exemple $(\mathbb{C}, |\cdot|)$ qui est une algèbre de Banach commutative ou $(\mathcal{M}_n(\mathbb{C}), \|\cdot\|)$ muni d'une quelconque norme multiplicative. Cette dernière algèbre n'est pas commutative.

Pour tout z dans E et tout n dans $\mathbb{N}^*$, on note $u_n(z) = \left(1 + \dfrac{1}{n} z\right)^n$.

Pour $z \in E$ et $m > n$, on a :

$$\begin{aligned}
\|u_m(z) - u_n(z)\| &= \left\| \sum_{k=0}^{n} \left(\binom{m}{k} \frac{1}{m^k} - \binom{n}{k} \frac{1}{n^k} \right) z^k - \sum_{k=n+1}^{m} \binom{m}{k} \frac{1}{m^k} z^k \right\| \\
&\leq \sum_{k=0}^{n} \left| \binom{m}{k} \frac{1}{m^k} - \binom{n}{k} \frac{1}{n^k} \right| \|z\|^k + \sum_{k=n+1}^{m} \binom{m}{k} \frac{1}{m^k} \|z\|^k
\end{aligned}$$

(comme z et 1 commutent, on peut utiliser la formule du binôme). Pour k compris entre 2 et n, on a :

$$\begin{aligned}
\binom{m}{k} \frac{1}{m^k} &= \frac{m(m-1)\dots(m-(k-1))}{k! m^k} = \frac{1}{k!} \left(1 - \frac{1}{m}\right) \cdots \left(1 - \frac{k-1}{m}\right) \\
&\geq \frac{1}{k!} \left(1 - \frac{1}{n}\right) \cdots \left(1 - \frac{k-1}{n}\right) = \binom{n}{k} \frac{1}{n^k}
\end{aligned}$$

et pour $k = 0$ ou 1, ces deux quantités valent 1. On a donc :

$$\begin{aligned}\|u_m(z) - u_n(z)\| &\leq \sum_{k=0}^{n} \left(\binom{m}{k} \frac{1}{m^k} - \binom{n}{k} \frac{1}{n^k} \right) \|z\|^k + \sum_{k=n+1}^{m} \binom{m}{k} \frac{1}{m^k} \|z\|^k \\ &\leq u_m(\|z\|) - u_n(\|z\|)\end{aligned}$$

et il en résulte que la suite $(u_n(z))_{n\geq 1}$ est de Cauchy dans $(E, \|\cdot\|)$. Cet espace étant complet, la suite y est convergente. On note e^z sa limite dans E.

Faisant tendre m vers l'infini dans l'inégalité précédente, on a :

$$\forall n \geq 1,\ \|e^z - u_n(z)\| \leq e^{\|z\|} - u_n(\|z\|)$$

et on en déduit que la convergence est uniforme sur tout compact de E.

Si $(z_n)_{n\in\mathbb{N}}$ est une suite d'éléments de E qui converge vers z, on en déduit que la suite $(u_n(z_n))_{n\in\mathbb{N}}$ converge vers e^z.

En écrivant que :

$$u_n(z)\, u_n(-z) = \left(1 - \frac{1}{n^2} z^2\right)^n = u_n\left(-\frac{1}{n} z^2\right)$$

(les polynômes en z commutent dans E), on en déduit par passage à la limite que $e^z e^{-z} = 1$, c'est-à-dire que pour tout $z \in E$, e^z est inversible d'inverse e^{-z}.

Pour z, z' dans E, qui commutent dans E, on a :

$$u_n(z)\, u_n(z') = \left(\left(1 + \frac{1}{n} z\right)\left(1 + \frac{1}{n} z'\right)\right)^n = \left(1 + \frac{1}{n} z_n\right)^n = u_n(z_n)$$

avec $z_n = z + z' + \dfrac{1}{n} z z' \underset{n\to+\infty}{\to} z + z'$. Il en résulte que :

$$e^z e^{z'} = \lim_{n\to+\infty} u_n(z)\, u_n(z') = \lim_{n\to+\infty} u_n(z_n) = e^{z+z'}$$

Ce résultat est faux si z et z' ne commutent pas comme le montre l'exemple des matrices $A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & -1 \\ 0 & 0 \end{pmatrix}$ dans $\mathcal{M}_2(\mathbb{R})$.

6.5 La suite de fonctions $\left(\sum_{k=0}^{n} \dfrac{x^k}{k!}\right)_{n\in\mathbb{N}}$

Si y est solution du problème (6.1), elle est alors de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ avec $y^{(n)}(0) = 1$ pour tout $n \in \mathbb{N}$. Pour tout réel non nul x, la formule de Taylor à l'ordre n nous donne alors :

$$y(x) = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{x^{n+1}}{(n+1)!} y^{(n+1)}(\theta x) = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{x^{n+1}}{(n+1)!} y(\theta x)$$

avec $\theta \in \left]0, 1\right[$. Avec $\lim\limits_{n\to+\infty} \dfrac{x^{n+1}}{(n+1)!} = 0$, on en déduit alors que $y(x) = \sum\limits_{n=0}^{+\infty} \dfrac{x^n}{n!}$, c'est-à-dire que y est limite de la suite de fonctions $(w_n)_{n\in\mathbb{N}}$ définie par :

$$\forall x \in \mathbb{R},\ w_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}$$

Le lien entre cette suite et la suite $\left(\left(1 + \dfrac{x}{n}\right)^n\right)_{n\geq 1}$ peut être précisé comme suit.

Lemme 6.7 *Pour tous $x > 0$ et n, p dans $\mathbb{N}^*$ on a :*

$$u_n(x) \leq w_n(x) \leq u_{n+p}(x)$$

Démonstration. Avec $\binom{n}{k} = 0$ pour $k > n$ et $\dfrac{\binom{n}{k}}{n^k} = \dfrac{1}{k!} \prod_{j=0}^{k-1} \left(1 - \dfrac{j}{n}\right) \leq \dfrac{1}{k!}$ pour $1 \leq k \leq n$, on déduit que $u_n(x) \leq w_n(x)$ et avec $\binom{n+p}{0} = 1$, $\binom{n+p}{k} = \dfrac{1}{k!} \prod_{j=0}^{k-1} (n+p-j) > \dfrac{1}{k!}$ pour $1 \leq k \leq n$, on déduit que $w_n(x) \leq u_{n+p}(x)$. ■

En particulier, pour $p = n^2$, on a :

$$u_{n+n^2}(x) < \left(1 + \frac{x}{n^2}\right)^{n^2} \left(1 + \frac{x}{n^2}\right)^n = u_{n^2}(x)\, u_n\left(\frac{x}{n}\right)$$

et de $u_n(x) \leq w_n(x) \leq u_{n^2}(x)\, u_n\left(\dfrac{x}{n}\right)$, on déduit que $\lim\limits_{n \to +\infty} w_n(x) = e^x$ (on a $\lim\limits_{n \to +\infty} u_n\left(\dfrac{x}{n}\right) = e^0 = 1$), c'est-à-dire que $e^x = \sum\limits_{n=0}^{+\infty} \dfrac{x^n}{n!}$ pour tout réel $x > 0$, ce résultat étant encore vrai pour $x = 0$.

Si on se place maintenant sur une algèbre de Banach E, on a pour tout $z \in E$ et $n \in \mathbb{N}$:

$$\begin{aligned} \|w_n(z) - u_n(z)\| &= \left\| \sum_{k=0}^{n} \left(\frac{1}{k!} - \frac{\binom{n}{k}}{n^k} \right) z^k \right\| \leq \sum_{k=0}^{n} \left(\frac{1}{k!} - \frac{\binom{n}{k}}{n^k} \right) \|z\|^k \\ &\leq w_n(\|z\|) - u_n(\|z\|) \underset{n \to +\infty}{\to} 0 \end{aligned}$$

et donc $e^z = \sum\limits_{n=0}^{+\infty} \dfrac{1}{n!} z^n$ pour tout $z \in E$ (ce qui montre au passage, en prenant $E = \mathbb{R}$, que ce résultat est vrai pour les réel négatifs).

6.6 Une définition du logarithme

Si on veut définir la fonction logarithme comme réciproque de l'exponentielle à partir d'une suite de fonctions, on part de l'équation $y = u_n(x) = \left(1 + \dfrac{x}{n}\right)^n$ où $y > 0$ et $n \geq 1$ sont donnés. Une solution de cette équation est $x = n\left(\sqrt[n]{y} - 1\right)$.

On est donc ainsi amené à considérer la suite de fonctions $(w_n)_{n \geq 1}$ définie par :

$$\forall n \geq 1,\ \forall x \in \mathbb{R}^{+,*},\ w_n(x) = n\left(\sqrt[n]{x} - 1\right)$$

la fonction $\sqrt[n]{\cdot}$ étant la fonction réciproque sur $\mathbb{R}^{+,*}$ de la fonction $x \mapsto x^n$. On rappelle que cette fonction est indéfiniment dérivable et strictement croissante de $\mathbb{R}^{+,*}$ sur $\mathbb{R}^{+,*}$.

Le résultat suivant nous sera utile.

Lemme 6.8 *Pour tout réel $x > 0$, on a $\lim\limits_{n \to +\infty} \sqrt[n]{x} = 1$.*

Démonstration. Avec $\sqrt[n]{1} = 1$ et $\sqrt[n]{\dfrac{1}{x}} = \dfrac{1}{\sqrt[n]{x}}$, il suffit de montrer le résultat pour $x > 1$. Dans ce cas la suite $(\sqrt[n]{x})_{n \in \mathbb{N}^*}$ est strictement décroissante ($\sqrt[n+1]{x} < \sqrt[n]{x}$ équivaut à $x < x^{\frac{n+1}{n}} =$

$x\sqrt[n]{x}$ encore équivalent à $\sqrt[n]{x} > 1$) et minorée par 1, elle converge donc vers un réel $\ell \geq 1$. Si $\ell > 1$, pour $\lambda \in]1, \ell[$ il existe un entier n_0 tel que $\sqrt[n]{x} > \lambda$ pour tout $n \geq n_0$, ce qui entraîne $x > \lambda^n$ pour tout $n \geq n_0$ qui est incompatible avec $\lim\limits_{n \to +\infty} \lambda^n = +\infty$. ■

Théorème 6.5 *Pour tout réel $x > 0$ la suite $(w_n(x))_{n \in \mathbb{N}^*}$ converge vers un réel $\ell(x)$ tel que $\ell(1) = 0$, $\ell(x) > 0$ pour $x > 1$ et $\ell(x) < 0$ pour $x \in]0, 1[$.*

Démonstration. Pour $x = 1$, la suite est stationnaire sur 0.

En notant $\alpha_n(x) = w_n(x) - w_{n+1}(x)$, on définit une fonction indéfiniment dérivable sur $\mathbb{R}^{+,*}$ avec :

$$\forall x > 0,\ \alpha_n'(x) = x^{\frac{1}{n}-1} - x^{\frac{1}{n+1}-1} = x^{-\frac{n}{n+1}} \left(x^{\frac{1}{n(n+1)}} - 1 \right)$$

La fonction α_n est donc strictement décroissante sur $]0, 1[$, strictement croissante sur $]1, \infty[$ avec $\alpha_n(1) = 0$, il en résulte que $\alpha_n(x) > 0$ pour tout $x \in \mathbb{R}^{+,*} \setminus \{1\}$.

La suite $(w_n(x))_{n \in \mathbb{N}^*}$ est donc strictement décroissante minorée par 0 pour tout $x > 1$, elle converge donc vers un réel $\ell(x) \geq 0$.

Si $x \in]0, 1[$, il s'écrit $x = \dfrac{1}{y}$ avec $y > 1$ et :

$$w_n(x) = n\left(\frac{1}{\sqrt[n]{y}} - 1\right) = -\frac{w_n(y)}{\sqrt[n]{y}} \underset{n \to +\infty}{\to} \ell(x) = -\ell(y) \leq 0$$

Pour $x \in]0, 1[$ la suite $(w_n(x))_{n \in \mathbb{N}^*}$ étant strictement décroissante à valeurs strictement négatives on a $\ell(x) < w_1(x) < 0$. Il en résulte que $\ell(x) = -\ell\left(\dfrac{1}{x}\right) > 0$ pour $x > 1$. ■

On définit donc une fonction ℓ sur $\mathbb{R}^{+,*}$ avec :

$$\forall x > 0,\ \ell(x) = \lim_{n \to +\infty} n\left(\sqrt[n]{x} - 1\right)$$

Théorème 6.6 *La fonction ℓ est solution sur $\mathbb{R}^{+,*}$ de l'équation fonctionnelle :*

$$\ell(xy) = \ell(x) + \ell(y)$$

Démonstration. Pour x, y dans $\mathbb{R}^{+,*}$ et $n \geq 1$, on a :

$$\begin{aligned} w_n(xy) &= n(\sqrt[n]{xy} - 1) = n\left(\sqrt[n]{x} - 1\right)\sqrt[n]{y} + n(\sqrt[n]{y} - 1) \\ &= w_n(x)\sqrt[n]{y} + w_n(y) \end{aligned}$$

et en passant à la limite on a le résultat. ■

On retrouve $\ell(1) = 0$ et $\ell\left(\dfrac{1}{x}\right) = -\ell(x)$.

Corollaire 6.2 *Pour tout réel $x > 0$ et tout nombre rationnel r, on a $\ell(x^r) = r\ell(x)$.*

Démonstration. Avec l'équation fonctionnelle, on montre facilement par récurrence que :

$$\forall p \in \mathbb{N},\ \ell(x^p) = p\ell(x)$$

En écrivant, pour tout $p \geq 1$, que $\ell(x) = \ell\left(\left(x^{\frac{1}{p}}\right)^p\right)$, on déduit que $\ell\left(x^{\frac{1}{p}}\right) = \dfrac{1}{p}\ell(x)$ et il en résulte que pour tout rationnel positif $r = \dfrac{p}{q}$, avec $p \in \mathbb{N}$ et $q \in \mathbb{N} \setminus \{0\}$, on a :

$$\ell(x^r) = \ell\left(\left(x^{\frac{1}{q}}\right)^p\right) = p\ell\left(x^{\frac{1}{q}}\right) = \frac{p}{q}\ell(x) = r\ell(x)$$

Enfin avec $\ell\left(\frac{1}{x}\right)=-\ell(x)$, on déduit que le résultat précédent est encore valable pour tout rationnel négatif. ■

Lemme 6.9 *La fonction ℓ est strictement croissante sur $\mathbb{R}^{+,*}$.*

Démonstration. La croissance de ℓ résulte immédiatement de celle des w_n.
On peut aussi écrire pour $0<x<y$ que :

$$0<\ell\left(y\frac{1}{x}\right)=\ell(y)-\ell(x)$$

ce qui donne la stricte croissance. ■

Lemme 6.10 *La fonction ℓ est continue en* 1.

Démonstration. Avec $0<\ell\left(2^{\frac{1}{n}}\right)=\frac{1}{n}\ell(2)$, on déduit que $\lim\limits_{n\to+\infty}\ell\left(2^{\frac{1}{n}}\right)=0$ et avec $\ell\left(2^{-\frac{1}{n}}\right)=-\frac{1}{n}\ell(2)<0$ que $\lim\limits_{n\to+\infty}\ell\left(2^{-\frac{1}{n}}\right)=0$. Pour tout réel $\varepsilon>0$ on peut donc trouver un entier n_0 tel que $-\varepsilon<\ell\left(2^{-\frac{1}{n}}\right)<\ell\left(2^{\frac{1}{n}}\right)<\varepsilon$ pour tout $n\geq n_0$. Pour tout x dans le voisinage ouvert de 1, $I_\varepsilon=\left]2^{-\frac{1}{n_0}},2^{\frac{1}{n_0}}\right[$, on a du fait de la croissance de ℓ :

$$-\varepsilon<\ell\left(2^{-\frac{1}{n_0}}\right)<\ell(x)<\ell\left(2^{\frac{1}{n_0}}\right)<\varepsilon$$

On a donc ainsi montré que ℓ est continue en 1 (on a $\ell(1)=0$). ■

Théorème 6.7 *La fonction ℓ est continue sur $\mathbb{R}^{+,*}$.*

Démonstration. Résulte de :

$$\ell(x)-\ell(x_0)=\ell\left(\frac{x}{x_0}\right)-\ell(1)$$

■

Théorème 6.8 *La fonction ℓ est concave sur $\mathbb{R}^{+,*}$.*

Démonstration. Résulte de la concavité des w_n. ■
On retrouve ainsi la continuité de ℓ.
On peut en fait montrer très simplement, sans les considérations précédentes (excepté le lemme 6.8), que la suite $(w_n(x))_{n\in\mathbb{N}^*}$ converge uniformément sur tout compact de $\mathbb{R}^{+,*}$ vers une fonction dérivable de dérivée égale à $\frac{1}{x}$.

Théorème 6.9 *La suite $(w_n(x))_{n\in\mathbb{N}^*}$ converge sur $\mathbb{R}^{+,*}$ vers la primitive nulle en 1 de la fonction $x\mapsto\frac{1}{x}$.*

Démonstration. Chaque fonction w_n est dérivable et pour tout $x > 0$, on a $\lim\limits_{n \to +\infty} w'_n(x) = \lim\limits_{n \to +\infty} \frac{\sqrt[n]{x}}{x} = \frac{1}{x}$. De plus pour tout $R > 1$ et tout $x \in \left[\frac{1}{R}, R\right]$, on a :

$$\left|w'_n(x) - \frac{1}{x}\right| = \frac{1}{x} \left|\sqrt[n]{x} - 1\right| \leq R\left(\sqrt[n]{R} - 1\right)$$

avec $\lim\limits_{n \to +\infty} \left(\sqrt[n]{R} - 1\right) = 0$, ce qui signifie que la suite $(w'_n(x))_{n \in \mathbb{N}^*}$ converge uniformément vers $\frac{1}{x}$ sur $\left[\frac{1}{R}, R\right]$. Comme de plus la suite $(w_n(1))_{n \in \mathbb{N}^*}$ est convergente, on en déduit que la suite $(w_n(x))_{n \in \mathbb{N}^*}$ converge uniformément sur $\left[\frac{1}{R}, R\right]$ vers une fonction dérivable ℓ de dérivée égale à $\frac{1}{x}$. On a donc $\ell(x) = \int_1^x \frac{dt}{t}$ pour tout $x \in \left[\frac{1}{R}, R\right]$. Ce résultat étant valable pour tout $R > 1$, il est vrai sur $\mathbb{R}^{+,*}$. ■

On retrouve donc ainsi la définition classique du logarithme et ses propriétés usuelles s'en déduisent.

6.7 Les fonctions ch, sh, cos et sin

On définit les fonctions cosinus et sinus réels par :

$$\forall t \in \mathbb{R},\ \cos(t) = \Re\left(e^{it}\right) = \frac{e^{it} + e^{-it}}{2} \text{ et } \sin(t) = \Im\left(e^{it}\right) = \frac{e^{it} - e^{-it}}{2i}$$

L'égalité $|e^{it}| = 1$ se traduit alors par :

$$\forall t \in \mathbb{R},\ \cos^2(t) + \sin^2(t) = 1$$

et ces fonctions sont donc à valeurs dans $[-1, 1]$.

De la définition de l'exponentielle complexe et de la continuité des fonction partie réelle et partie imaginaire, on déduit que ces fonctions sont développables en séries entières sur $\mathbb{R}$ avec :

$$\cos(t) = \Re\left(\sum_{n=0}^{+\infty} \frac{i^n t^n}{n!}\right) = \sum_{n=0}^{+\infty} \Re\left(\frac{i^n t^n}{n!}\right) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} t^{2n}$$

et :

$$\sin(t) = \Im\left(\sum_{n=0}^{+\infty} \frac{i^n t^n}{n!}\right) = \sum_{n=0}^{+\infty} \Im\left(\frac{i^n t^n}{n!}\right) = \sum_{n=0}^{+\infty} \frac{(-1)^n}{(2n+1)!} t^{2n+1}$$

On déduit également que ces fonctions sont indéfiniment dérivables sur $\mathbb{R}$ avec :

$$\forall t \in \mathbb{R},\ \cos'(t) = \Re\left(ie^{it}\right) = -\sin(t) \text{ et } \sin'(t) = \Im\left(ie^{it}\right) = \cos(t)$$

En fait, on définit plus généralement les fonctions cos, sin, ch et sh sur $\mathbb{C}$ par :

$$\forall z \in \mathbb{C},\ \begin{cases} \operatorname{ch}(z) = \dfrac{e^z + e^{-z}}{2} = \sum\limits_{n=0}^{+\infty} \dfrac{1}{(2n)!} z^{2n} \\ \operatorname{sh}(z) = \dfrac{e^z - e^{-z}}{2} = \sum\limits_{n=0}^{+\infty} \dfrac{1}{(2n+1)!} z^{2n+1} \\ \cos(z) = \operatorname{ch}(iz) = \dfrac{e^{iz} + e^{-iz}}{2} = \sum\limits_{n=0}^{+\infty} \dfrac{(-1)^n}{(2n)!} z^{2n} \\ \sin(z) = -i \operatorname{sh}(iz) = \dfrac{e^{iz} - e^{-iz}}{2i} = \sum\limits_{n=0}^{+\infty} \dfrac{(-1)^n}{(2n+1)!} z^{2n+1} \end{cases}$$

On a donc :

$$\forall z \in \mathbb{C}, \left\{ \begin{array}{l} \operatorname{ch}^2(z) - \operatorname{sh}^2(z) = 1 \\ \cos^2(z) + \sin^2(z) = 1 \end{array} \right.$$

Les fonctions $\cos, \operatorname{ch}$ sont paires et les fonctions $\sin, \operatorname{sh}$ sont impaires.

En se limitant à l'ensemble des réels, les fonctions ch et sh sont indéfiniment dérivables sur $\mathbb{R}$ avec :

$$\forall x \in \mathbb{R},\ \operatorname{ch}'(x) = \operatorname{sh}(x) \text{ et } \operatorname{sh}'(x) = \operatorname{ch}(x)$$

De la définition $\operatorname{ch}(x) = \dfrac{e^x + e^{-x}}{2}$, on déduit que $\operatorname{ch}(x) > 0$ et avec $\operatorname{ch}^2(x) = \operatorname{sh}^2(x) + 1$, que $\operatorname{ch}(x) \geq 1$ pour tout réel x, la valeur 1 étant atteinte pour $x = 0$. Il en résulte que sh est strictement croissante sur $\mathbb{R}$, donc $\operatorname{sh}(x) > \operatorname{sh}(0) = 0$ pour $x > 0$ et ch est strictement croissante sur $\mathbb{R}^+$. Enfin avec $\lim\limits_{x \to +\infty} \operatorname{ch}(x) = \lim\limits_{x \to +\infty} \operatorname{sh}(x) = +\infty$ et l'argument de parité, on peut tracer les graphes de ces fonctions (figure 6.1).

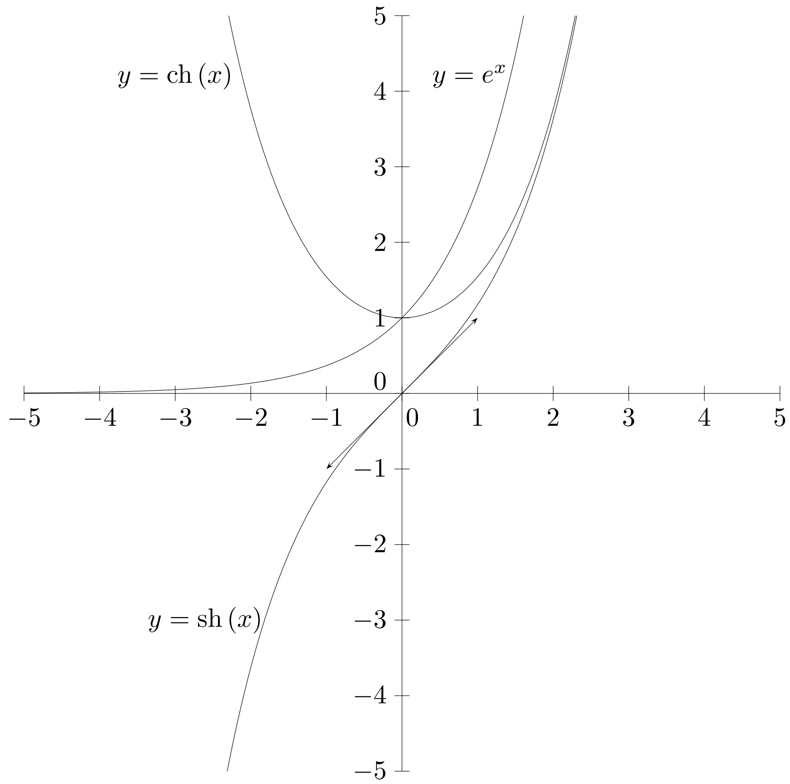

FIGURE 6.1 – fonctions e^x, $\operatorname{ch}(x)$ et $\operatorname{sh}(x)$

On vérifie facilement que :

$$\forall z \in \mathbb{C}, \left\{ \begin{array}{l} e^z = \operatorname{ch}(z) + \operatorname{sh}(z) \\ e^{-z} = \operatorname{ch}(z) - \operatorname{sh}(z) \\ e^{iz} = \cos(z) + i\sin(z) \\ e^{-iz} = \cos(z) - i\sin(z) \end{array} \right.$$

De l'équation fonctionnelle vérifiée par la fonction exponentielle, on déduit les relations suivantes valables pour tous nombres complexes a, b :

$$\left\{\begin{array}{l}\operatorname{ch}(a+b)=\operatorname{ch}(a)\operatorname{ch}(b)+\operatorname{sh}(a)\operatorname{sh}(b)\\ \cos(a+b)=\cos(a)\cos(b)-\sin(a)\sin(b)\\ \operatorname{sh}(a+b)=\operatorname{sh}(a)\operatorname{ch}(b)+\operatorname{ch}(a)\operatorname{sh}(b)\\ \sin(a+b)=\sin(a)\cos(b)+\cos(a)\sin(b)\end{array}\right.$$

et de ces formules on déduit les classiques formules de trigonométrie circulaire et hyperbolique.

La démonstration de la première formule peut se faire comme suit.

Pour a, b dans $\mathbb{C}$, on a :

$$\begin{aligned}2\operatorname{ch}(a+b)&=e^{a+b}+e^{-a-b}=e^{a}e^{b}+e^{-a}e^{-b}\\&=(\operatorname{ch}(a)+\operatorname{sh}(a))(\operatorname{ch}(b)+\operatorname{sh}(b))+(\operatorname{ch}(a)-\operatorname{sh}(a))(\operatorname{ch}(b)-\operatorname{sh}(b))\\&=2(\operatorname{ch}(a)\operatorname{ch}(b)+\operatorname{sh}(a)\operatorname{sh}(b))\end{aligned}$$

La deuxième s'en suit :

$$\begin{aligned}\cos(a+b)&=\operatorname{ch}(ia+ib)=\operatorname{ch}(ia)\operatorname{ch}(ib)+\operatorname{sh}(ia)\operatorname{sh}(ib)\\&=\cos(a)\cos(b)-\sin(a)\sin(b)\end{aligned}$$

Les deux dernières formules se montrent de manière analogue.

Avec les arguments de parité, on déduit alors les formules suivantes :

$$\left\{\begin{array}{l}\operatorname{ch}(a-b)=\operatorname{ch}(a)\operatorname{ch}(b)-\operatorname{sh}(a)\operatorname{sh}(b)\\ \cos(a-b)=\cos(a)\cos(b)+\sin(a)\sin(b)\\ \operatorname{sh}(a-b)=\operatorname{sh}(a)\operatorname{ch}(b)-\operatorname{ch}(a)\operatorname{sh}(b)\\ \sin(a-b)=\sin(a)\cos(b)-\cos(a)\sin(b)\end{array}\right.$$

Prenant $a=b$, on a :

$$\left\{\begin{array}{l}\operatorname{ch}(2a)=\operatorname{ch}^2(a)+\operatorname{sh}^2(a)=1+2\operatorname{sh}^2(a)=2\operatorname{ch}^2(a)-1\\ \cos(2a)=\cos^2(a)-\sin^2(a)=2\cos^2(a)-1=1-2\sin^2(a)\\ \operatorname{sh}(2a)=2\operatorname{sh}(a)\operatorname{ch}(b)\\ \sin(2a)=2\sin(a)\cos(b)\end{array}\right.$$

Si les fonctions cos et sin restreintes à $\mathbb{R}$ bornées, il n'en n'est pas de même pour ces fonctions définies sur $\mathbb{C}$. Précisément pour $z=x+iy\in\mathbb{C}$, on a :

$$\begin{aligned}\cos(z)&=\cos(x)\cos(iy)-\sin(x)\sin(iy)\\&=\cos(x)\operatorname{ch}(y)+i\sin(x)\operatorname{sh}(y)\end{aligned}$$

et :

$$\begin{aligned}|\cos(z)|^2&=\cos^2(x)\operatorname{ch}^2(y)+\sin^2(x)\operatorname{sh}^2(y)\\&=\cos^2(x)\operatorname{ch}^2(y)+\big(1-\cos^2(x)\big)\operatorname{sh}^2(y)\\&=\cos^2(x)\big(\operatorname{ch}^2(y)-\operatorname{sh}^2(y)\big)+\operatorname{sh}^2(y)\\&=\cos^2(x)+\operatorname{sh}^2(y)\geq\operatorname{sh}^2(y)\end{aligned}$$

la fonction sh étant non majorée sur $\mathbb{R}$.

6.8 Le nombre π

6.8.1 Une définition du nombre π

On désigne par Γ l'ensemble des nombres complexes de module égal à 1. C'est un sous-groupe du groupe multiplicatif $(\mathbb{C}^*, \cdot)$.

De l'équation fonctionnelle vérifiée par la fonction exponentielle complexe, on déduit le résultat suivant.

Théorème 6.10 *L'application $\varphi : t \mapsto e^{it}$ réalise un morphisme de groupes de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$.*

Nous allons voir que ce morphisme φ est surjectif et que son noyau n'est pas réduit à $\{0\}$, comme c'est un sous-groupe de $(\mathbb{R}, +)$, il est dense ou discret.

Nous allons voir qu'il est discret, c'est-à-dire de la forme $\mathbb{Z}\alpha$, où α est un réel strictement positif.

Lemme 6.11 *On a* $\cos(2) < 0$.

Démonstration. $\cos(2)$ est la somme de la série alternée $\sum\limits_{n=0}^{+\infty} \frac{(-1)^n}{(2n)!} 2^{2n}$, donc en notant S_n la somme partielle d'indice n de cette série, on a pour tout entier $n \geq 0$:

$$S_{2n+1} \leq \cos(2) \leq S_{2n}$$

et en particulier :

$$\cos(2) \leq S_4 = 1 - \frac{2^2}{2} + \frac{2^4}{24} = -\frac{1}{3} < 0$$

■

Lemme 6.12 *L'ensemble $E = \{t \in [0, 2] \mid \cos(t) = 0\}$ est non vide et admet une borne inférieure $\alpha \in]0, 2[\cap E$.*

Démonstration. Comme E est contenu dans $[0, 2]$, il est borné. Il reste à montrer qu'il est non vide.

Comme la fonction cos est continue sur $\mathbb{R}$ avec $\cos(0) = 1 > 0$ et $\cos(2) < 0$ le théorème des valeurs intermédiaires nous dit qu'il existe un réel $t \in]0, 2[$ tel que $\cos(t) = 0$.

L'ensemble E étant non vide et minoré admet une borne inférieure α et cette borne inférieure est dans E puisque cet ensemble est fermé ($E = [0, 2] \cap \cos^{-1}\{0\}$ avec cos continue). On a donc $\cos(\alpha) = 0$ et $\alpha \in]0, 2[$ puisque $\cos(0) \neq 0$ et $\cos(2) \neq 0$ (on peut aussi dire, par définition de la borne inférieure, qu'il existe une suite $(t_n)_{n\in\mathbb{N}}$ dans E qui converge vers α, donc $\cos(\alpha) = \lim\limits_{n\to+\infty} \cos(t_n) = 0$ et $\alpha \in E$). ■

On définit le nombre π par $\pi = 2\alpha$.

$\frac{\pi}{2}$ est donc le plus petit réel positif qui vérifie $\cos\left(\frac{\pi}{2}\right) = 0$.

Lemme 6.13 *On a $\cos(t) > 0$ pour tout $t \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, la fonction* sin *est strictement croissante de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ sur $[-1, 1]$, $\sin(t) > 0$ pour tout $t \in]0, \pi[$ et la fonction* cos *est strictement décroissante de $[0, \pi]$ sur $[-1, 1]$.*

Démonstration. Par définition de α, on a $\cos(t) > 0$ pour tout $t \in \left]0, \frac{\pi}{2}\right[$. En effet, on a $\cos(t) \neq 0$ pour tout $t \in \left]0, \frac{\pi}{2}\right[$ par définition de $\alpha = \frac{\pi}{2}$ comme borne inférieure de E. La fonction continue cos est donc de signe constant sur cet intervalle et avec $\cos(0) = 1 > 0$, on déduit que $\cos(t) > 0$ pour $t > 0$ voisin de 0 et $\cos(t) > 0$ pour tout $t \in \left]0, \frac{\pi}{2}\right[$. Comme la fonction cos est paire avec $\cos(0) = 1 > 0$, on déduit que $\cos(t) > 0$ pour tout $t \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$.

Comme $\sin'(t) = \cos(t) > 0$ pour tout $t \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, la fonction sin est strictement croissante sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$.

Avec $\cos^2\left(\frac{\pi}{2}\right) + \sin^2\left(\frac{\pi}{2}\right) = 1$ et $\cos\left(\frac{\pi}{2}\right) = 0$, on déduit que $\sin\left(\frac{\pi}{2}\right) = \pm 1$ et avec $\sin\left(\frac{\pi}{2}\right) > \sin(0) = 0$, que $\sin\left(\frac{\pi}{2}\right) = 1$. La fonction sin étant impaire, on a $\sin\left(-\frac{\pi}{2}\right) = -1$ et l'image de $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ par sin est bien $[-1, 1]$.

Pour $t \in \left]0, \frac{\pi}{2}\right[$, on a $\sin(t) > \sin(0) = 0$ et pour $t = \frac{\pi}{2} + t' \in \left]\frac{\pi}{2}, \pi\right[$, $\sin(t) = \sin\left(\frac{\pi}{2}\right)\cos(t') + \cos\left(\frac{\pi}{2}\right)\sin(t') = \cos(t') > 0$.

Comme $\cos'(t) = -\sin(t) < 0$ pour tout $t \in]0, \pi[$, la fonction cos est strictement décroissante sur $[0, \pi]$.

Avec $\cos(0) = 1$, $\cos(\pi) = \cos\left(2\frac{\pi}{2}\right) = \cos^2\left(\frac{\pi}{2}\right) - \sin^2\left(\frac{\pi}{2}\right) = -1$, on déduit que l'image de $[0, \pi]$ par cos est bien $[-1, 1]$. ∎

Remarque 6.6 *La fonction* cos *[resp.* sin*] est continue strictement décroissante [resp. croissante] de* $[0, \pi]$ *[resp.* $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$*] sur* $[-1, 1]$, *elle réalise donc un homéomorphisme de* $[0, \pi]$ *[resp.* $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$*] sur* $[-1, 1]$. *Sa fonction réciproque est notée* arccos *[resp.* arcsin*]. Comme* cos *[resp.* sin*] est dérivable de dérivée non nulle sur* $]0, \pi[$ *[resp.* $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$*] la fonction* arccos *[resp.* arcsin*] est dérivable sur* $]-1, 1[$ *de dérivée* $\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}$ *[resp.* $\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}$*].*

Lemme 6.14 *Pour tout* $t \in]0, 2\pi[$, *on a* $e^{it} \neq 1$.

Démonstration. Soient $t \in]0, 2\pi[$ et $z = e^{i\frac{t}{4}} = \cos\left(\frac{t}{4}\right) + i\sin\left(\frac{t}{4}\right)$. Comme $\frac{t}{4} \in \left]0, \frac{\pi}{2}\right[$, on a $\cos\left(\frac{t}{4}\right) > 0$ et $\sin\left(\frac{t}{4}\right) > 0$.

En écrivant que :

$$\begin{aligned} e^{it} = \left(e^{i\frac{t}{4}}\right)^4 &= \cos^4\left(\frac{t}{4}\right) - 6\cos^2\left(\frac{t}{4}\right)\sin^2\left(\frac{t}{4}\right) + \sin^4\left(\frac{t}{4}\right) \\ &+ 4i\cos\left(\frac{t}{4}\right)\sin\left(\frac{t}{4}\right)\left(\cos^2\left(\frac{t}{4}\right) - \sin^2\left(\frac{t}{4}\right)\right) \end{aligned}$$

on déduit que l'égalité $e^{it} = 1$ entraîne $\cos^2\left(\frac{t}{4}\right) = \sin^2\left(\frac{t}{4}\right)$ et avec $\cos^2\left(\frac{t}{4}\right) + \sin^2\left(\frac{t}{4}\right) = 1$, cela impose $\cos^2\left(\frac{t}{4}\right) = \sin^2\left(\frac{t}{4}\right) = \frac{1}{2}$ et :

$$\begin{aligned} e^{it} &= \cos^4\left(\frac{t}{4}\right) - 6\cos^2\left(\frac{t}{4}\right)\sin^2\left(\frac{t}{4}\right) + \sin^4\left(\frac{t}{4}\right) \\ &= \frac{1}{4} + \frac{1}{4} - \frac{6}{4} = -1 \end{aligned}$$

ce qui est une contradiction. ■

On en déduit immédiatement le résultat suivant.

Lemme 6.15 *On a :*

$$e^{i\frac{\pi}{2}} = i,\ e^{i\pi} = -1 \text{ et } e^{2i\pi} = 1$$

et pour tout réel t :

$$\begin{cases} \cos\left(t + \frac{\pi}{2}\right) = -\sin(t) \\ \sin\left(t + \frac{\pi}{2}\right) = \cos(t) \\ \cos(t+\pi) = -\cos(t) \\ \sin(t+\pi) = -\sin(t) \end{cases}$$

Les fonctions cos *et* sin *sont périodiques de plus petite période 2π.*

Démonstration. Les deux premières égalités se déduisent de (et sont même équivalentes à) :

$$\cos\left(\frac{\pi}{2}\right) = 0,\ \sin\left(\frac{\pi}{2}\right) = 1,\ \cos(\pi) = -1,\ \sin(\pi) = 0$$

($\cos(\pi) = -1$ a été montré avec le lemme précédent et $\sin^2(\pi) = 1 - \cos^2(\pi) = 0$).

Il en résulte que $e^{2i\pi} = (e^{i\pi})^2 = 1$, ce qui équivaut à $\cos(2\pi) = 1$ et $\sin(2\pi) = 0$.

Les formules de trigonométries nous donnent les dernières égalité et la 2π-périodicité de cos et sin.

Si $T \in]0, 2\pi[$ est une période plus petite, on a alors $\cos(T) = 1, \sin(T) = 0$, soit $e^{iT} = 1$ avec $T \in]0, 2\pi[$, ce qui est impossible. ■

Théorème 6.11 *Le noyau du morphisme de groupes* $\exp : z \mapsto e^z$, *de* $(\mathbb{C}, +)$ *dans* $(\mathbb{C}^*, \cdot)$, *est* $2i\pi\mathbb{Z}$.

Démonstration. Pour tout entier naturel k, on a $e^{2ik\pi} = (e^{2i\pi})^k = 1$ et avec $e^{-2ik\pi} = \frac{1}{e^{2ik\pi}}$, on déduit que le résultat est valable pour tout entier relatif k. La fonction $\psi : z \mapsto e^z - 1$ s'annule donc sur $2i\pi\mathbb{Z}$.

Si $e^z = 1$ avec $z = x + iy$, on a $|e^z| = e^x |e^{iy}| = e^x = 1$ et $x = 0$, donc $z = iy$ et $e^{iy} = 1$. Si $y \notin 2\pi\mathbb{Z}$, il existe un entier relatif k tel que $2k\pi < y < 2(k+1)\pi$ ($k = E\left(\frac{y}{2\pi}\right)$) donc $y - 2k\pi \in]0, 2\pi[$ et $e^{iy} = e^{i(y-2k\pi)} \neq 1$ d'après le lemme précédent. On a donc $y \in 2\pi\mathbb{Z}$ et $z = iy \in 2i\pi\mathbb{Z}$. ■

Le théorème précédent se traduit en disant que la fonction $z \mapsto e^z$ est périodique de période $2i\pi$. Il se traduit aussi en disant que l'égalité $e^z = 1$ est réalisée si, et seulement si, il existe un entier relatif k tel que $z = 2ik\pi$.

Remarque 6.7 *L'égalité* $\left(e^{2i\pi}\right)^k = 1$ *pour* k *non entier n'est pas vraiment valable, sans quoi on montrerait que* $-1 = 1$ *comme suit :*

$$\left(1 = e^{2i\pi}\right) \Rightarrow \left(1 = 1^{\frac{1}{2}} = \left(e^{2i\pi}\right)^{\frac{1}{2}} = e^{\frac{1}{2}2i\pi} = e^{i\pi} = -1\right)$$

Corollaire 6.3 *L'application* $\varphi : t \mapsto e^{it}$ *réalise un morphisme continu de groupes de* $(\mathbb{R}, +)$ *dans* $(\Gamma, \cdot)$ *de noyau* $\ker(\varphi) = 2\pi\mathbb{Z}$.

On peut aussi montrer directement qu'il existe un réel $\alpha > 0$ tel que $\ker(\varphi) = \alpha\mathbb{Z}$ et définir π par $\alpha = 2\pi$ (voir [8] ou [17]).

Montrons enfin que φ est surjectif.

Théorème 6.12 *L'application* $\varphi : t \mapsto e^{it}$ *réalise un morphisme continu de groupes surjectif de* $(\mathbb{R}, +)$ *sur* $(\Gamma, \cdot)$ *de noyau* $\ker(\varphi) = 2\pi\mathbb{Z}$.

Démonstration. Soit $z = x + iy$ dans Γ. On distingue les cas de figure suivants.

1. Si $x \geq 0$ et $y \geq 0$, alors avec $|z|^2 = x^2 + y^2 = 1$, on déduit que x et y sont dans $[0, 1]$. Comme la fonction cos est strictement décroissante sur $\left[0, \frac{\pi}{2}\right]$ avec $\cos(0) = 1$ et $\cos\left(\frac{\pi}{2}\right) = 0$, elle réalise une bijection de $\left[0, \frac{\pi}{2}\right]$ sur $[0, 1]$ et il existe un unique réel $t \in \left[0, \frac{\pi}{2}\right]$ tel que $x = \cos(t)$. On a alors $y^2 = 1 - x^2 = 1 - \cos^2(t) = \sin^2(t)$ et $y = \sin(t)$ puisque ces deux quantités sont positives. Il en résulte que $z = e^{it}$.
2. Si $x < 0$ et $y \geq 0$, alors $-iz = y - ix$ se trouve dans le premier cas de figure, il s'écrit donc $-iz = e^{it}$ et $z = ie^{it} = e^{i\left(t + \frac{\pi}{2}\right)}$.
3. Si $y < 0$ et x est réel quelconque, alors $-z$ se trouve dans le premier ou deuxième cas de figure, il s'écrit donc $-z = e^{it}$ et $z = -e^{it} = e^{i(t+\pi)}$.

■

Le résultat précédent nous dit en fait que pour tout nombre complexe $z \in \Gamma$ il existe un unique réel $t \in [0, 2\pi[$ tel que $z = e^{it}$. En effet, il existe un réel y tel que $z = e^{iy}$ et désignant par k l'entier relatif tel que $2k\pi \leq y < 2(k+1)\pi$ ($k = E\left(\frac{y}{2\pi}\right)$), on a $t = y - 2k\pi \in [0, 2\pi[$ et $e^{it} = e^{i(y-2k\pi)} = e^{iy} = z$. Si $t_1 \leq t_2$ sont deux tels réels, alors $t_2 - t_1 \in [0, 2\pi[$ est dans $\ker(\varphi) = 2\pi\mathbb{Z}$, donc nécessairement nul.

En notant, pour $z \in \Gamma$, t_0 le réel dans $[0, 2\pi[$ tel que $z = e^{it_0}$, on a $e^{it} = z$ avec t réel si, et seulement si, $t = t_0 + 2k\pi$ avec k entier relatif.

On peut aussi résumer cela en disant que le groupe multiplicatif Γ est isomorphe au groupe additif $\frac{\mathbb{R}}{\ker(\varphi)} = \frac{\mathbb{R}}{2\pi\mathbb{Z}}$.

Corollaire 6.4 *L'application* $\exp : z \mapsto e^z$ *réalise un morphisme de groupes surjectif de* $(\mathbb{C}, +)$ *sur* $(\mathbb{C}^*, \cdot)$ *de noyau* $\ker(\exp) = 2i\pi\mathbb{Z}$.

Démonstration. On sait déjà que exp un morphisme de groupes de $(\mathbb{C}, +)$ dans $(\mathbb{C}^*, \cdot)$ de noyau $2i\pi\mathbb{Z}$.

Pour tout nombre complexe non nul, on a $\frac{z}{|z|} \in \Gamma$ et il existe un réel y tel que $\frac{z}{|z|} = e^{iy}$, soit $z = |z|\, e^{iy}$ où $\rho = |z|$ est un réel strictement positif. Mais on a vu que l'exponentielle réelle réalise une bijection de $\mathbb{R}$ sur $\mathbb{R}^{+,*}$, il existe donc un unique réel x tel que $|z| = e^x$ et $z = e^{x+iy}$.
■

6.8.2 Les fonction $\tan$ et $\arctan$

Pour $x \in \mathbb{R}$, l'égalité $\cos(x) = 0$ équivaut à $x = \dfrac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$. En effet, pour $k \in \mathbb{Z}$ on a $e^{i\left(\frac{\pi}{2}+k\pi\right)} = e^{i\frac{\pi}{2}} \left(e^{i\pi}\right)^k = (-1)^k i$ et $\cos\left(\dfrac{\pi}{2} + k\pi\right) = 0$.

D'autre part pour $x \in \mathbb{R}$ ne s'écrivant pas $\dfrac{\pi}{2} + k\pi$ avec $k \in \mathbb{Z}$, il existe un entier p tel que $x + p\pi \in \left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$ et $\cos(x + p\pi) = (-1)^p \cos(x) \neq 0$ (de $e^{ip\pi} = \left(e^{i\pi}\right)^p = (-1)^p$, on déduit que $\cos(p\pi) = (-1)^p$ et $\sin(p\pi) = 0$). On a donc ainsi toutes les racines réelles de $\cos$.

On définit alors la fonction tangente sur $\mathbb{R} \setminus \left\{\dfrac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\right\}$ par $\tan(x) = \dfrac{\sin(x)}{\cos(x)}$.

Cette fonction est indéfiniment dérivable sur son domaine de définition de dérivée $\tan'(x) = 1 + \tan^2(x) = \dfrac{1}{\cos^2(x)}$.

Cette fonction est impaire, strictement croissante sur $\left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$ (dérivée strictement positive) avec $\lim\limits_{x \to \frac{\pi}{2}^-} \tan(x) = +\infty$. Elle définit donc un homéomorphisme de $\left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$ sur $\mathbb{R}$. Sa fonction réciproque est notée $\arctan$, c'est la fonction arc-tangente. Elle est dérivable de dérivé $\dfrac{1}{1+x^2}$.

6.8.3 Le lien avec le nombre π des géomètres

Le cercle unité du plan affine euclidien, identifié à Γ, peut être paramétré par :

$$\gamma : t \in [0, 2\pi] \mapsto (\cos(t), \sin(t))$$

Théorème 6.13 *Le périmètre du cercle unité du plan euclidien vaut 2π.*

Démonstration. On rappelle que la longueur d'un arc géométrique paramétré par une application γ de classe $\mathcal{C}^1$ de $[a, b]$ dans $\mathbb{R}^2$ est $\ell(\gamma) = \displaystyle\int_a^b \|\gamma'(t)\| \, dt$, où $\|\cdot\|$ désigne la norme euclidienne usuelle, ce qui donne pour le cercle :

$$\ell(\gamma) = \int_0^{2\pi} \sqrt{\sin^2(t) + \cos^2(t)} dt = 2\pi.$$

■

6.8.4 Les fonctions argument principal et logarithme

On a montré que tout nombre complexe non nul z s'écrit de manière unique $z = \rho e^{i\theta}$ avec $\rho > 0$ ($\rho = |z|$) et $\theta \in [0, 2\pi[$. Le réel ρ est le module de z.

Avec ces notations, on aura $z = \rho e^{it}$ si, et seulement si, $\rho e^{it} = \rho e^{i\theta}$, ce équivaut à $e^{i(t-\theta)} = 1$ ou encore à $t = \theta + 2k\pi$ avec k entier relatif. On dit alors que t est un argument de z. En se fixant k un tel argument est unique dans $[2k\pi, 2(k+1)\pi[$.

En utilisant les arguments, on peut montrer que les applications $t \mapsto e^{i\alpha t}$ sont les seuls morphismes de groupes continus de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$.

Lemme 6.16 *Les seuls morphismes de groupes continus de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$ sont les applications $x \mapsto e^{i\alpha x}$ avec $\alpha \in \mathbb{R}$.*

Démonstration. Avec ce qui précède, on voit que pour tout $\alpha \in \mathbb{R}$ l'application $x \mapsto e^{i\alpha x}$ est un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$.

Réciproquement si $f : \mathbb{R} \to \Gamma$ est un morphisme de groupes, il existe alors un unique réel $\alpha \in [0, 2\pi[$ tel que $f(1) = e^{i\alpha}$. Par récurrence on vérifie facilement que $f(n) = e^{in\alpha}$ pour tout $n \in \mathbb{N}$, puis avec $e^{i\alpha} = f\left(n\frac{1}{n}\right) = f\left(\frac{1}{n}\right)^n$ on déduit que $f\left(\frac{1}{n}\right) = e^{i\frac{\alpha}{n}}$ pour tout $n \geq 1$. Il en résulte que $f(r) = e^{ir\alpha}$ pour tout $r \in \mathbb{Q}^+$. Si de plus f est continue, avec la densité de $\mathbb{Q}^+$ dans $\mathbb{R}^+$, on déduit que $f(x) = e^{i\alpha x}$ pour tout $x \in \mathbb{R}^+$.

Enfin avec $1 = f(x - x) = f(x) f(-x)$ (le neutre est transformé en neutre), on déduit que $f(-x) = \dfrac{1}{f(x)} = \overline{f(x)}$ pour tout $x \in \mathbb{R}$. Il en résulte que $f(x) = e^{i\alpha x}$ pour tout $x \in \mathbb{R}$. ∎

Connaissant tous les morphismes de groupes continus de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, \cdot)$ (ce sont les $x \mapsto e^{ax}$ avec a réel), on déduit le résultat suivant.

Théorème 6.14 *Les seuls morphismes de groupes continus de* $(\mathbb{R}, +)$ *dans* $(\mathbb{C}^*, \cdot)$ *sont les applications* $x \mapsto e^{\alpha x}$ *avec* $\alpha \in \mathbb{C}$.

Démonstration. Si f est un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\mathbb{C}^*, \cdot)$, alors $|f|$ est un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, \cdot)$, il existe donc un réel a tel que $|f(x)| = e^{ax}$ pour tout $x \in \mathbb{R}$. La fonction $g : x \mapsto f(x) e^{-ax}$ est alors un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$, il existe donc un réel b tel que $f(x) e^{-ax} = e^{ibx}$ pour tout $x \in \mathbb{R}$. On a donc $f(x) = e^{\alpha x}$ pour tout $x \in \mathbb{R}$ avec $\alpha = a + ib \in \mathbb{C}$.

La réciproque est évidente. ∎

En fait un tel argument peut être uniquement déterminé dans tout intervalle de longueur 2π, $[\theta_0, \theta_0 + 2\pi[$ où θ_0 est un réel fixé. En effet en désignant par θ l'argument de $z \in \mathbb{C}^*$ dans $[0, 2\pi[$, il existe un entier k tel que $t = \theta - 2k\pi \in [\theta_0, \theta_0 + 2\pi[$ (il s'agit de réaliser $\theta_0 \leq \theta - 2k\pi < \theta_0 + 2\pi$, soit $2k\pi \leq \theta - \theta_0 < 2(k+1)\pi$ ou encore $k \leq \dfrac{\theta - \theta_0}{2\pi} < k + 1$, ce qui définit $k = E\left(\dfrac{\theta - \theta_0}{2\pi}\right)$) et $z = \rho e^{i\theta} = \rho e^{it}$. Si $t \leq t'$ sont deux tels arguments, on a $0 \leq t' - t < 2\pi$ et $t' - t = 2k\pi$ avec k entier, ce qui impose $k = 0$.

Le choix de $\theta_0 = -\pi$ définit l'argument principal dans $[-\pi, \pi[$ d'un nombre complexe non nul z.

En résumé tout nombre complexe non nul s'écrit de manière unique $z = \rho e^{i\theta}$ où ρ est un réel strictement positif et θ un réel dans $[-\pi, \pi[$. On note $\theta = \arg(z)$ cet argument principal.

On aura $\arg(z) = -\pi$ si, et seulement si, z est un réel strictement négatif.

Tout nombre complexe $z \in \mathbb{C} \setminus \mathbb{R}^-$ s'écrit donc $z = x + iy = \rho e^{i\theta}$ avec $\rho = \sqrt{x^2 + y^2}$, $\theta \in]-\pi, \pi[$ et $x = \rho \cos(\theta)$, $y = \rho \sin(\theta)$, ce qui donne :

$$\begin{cases} 2\cos^2\left(\dfrac{\theta}{2}\right) = \cos(\theta) + 1 = \dfrac{x}{\rho} + 1 \\ 2\sin\left(\dfrac{\theta}{2}\right)\cos\left(\dfrac{\theta}{2}\right) = \sin(\theta) = \dfrac{y}{\rho} \end{cases}$$

Comme $\dfrac{\theta}{2} \in \left]-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right[$, on a $\cos\left(\dfrac{\theta}{2}\right) > 0$ et divisant la première égalité par la seconde, on obtient :

$$\tan\left(\frac{\theta}{2}\right) = \frac{\frac{y}{\rho}}{\frac{x}{\rho} + 1} = \frac{y}{x + \sqrt{x^2 + y^2}}$$

soit :

$$\theta = \arg(z) = 2\arctan\left(\frac{y}{x + \sqrt{x^2 + y^2}}\right) = 2\arctan\left(\frac{\Im(z)}{\Re(z) + |z|}\right)$$

On peut donc définir la fonction argument principal par :

$$\begin{array}{rccc} \arg : & \mathbb{C} \setminus \mathbb{R}^- & \to &]-\pi, \pi[\\ & z & \mapsto & 2 \arctan \left(\frac{\Im(z)}{\Re(z)+|z|} \right) \end{array}$$

et cette fonction est continue sur $\mathbb{C} \setminus \mathbb{R}^-$.

Cette fonction est donc définie par $\arg(z) \in]-\pi, \pi[$ et $z = |z| e^{i \arg(z)}$ pour tout $z \in \mathbb{C} \setminus \mathbb{R}^-$.

Remarque 6.8 *Une telle fonction argument principal ne peut pas être prolongée en une fonction continue sur un ouvert contenant strictement $\mathbb{C} \setminus \mathbb{R}^-$.*
En effet, supposons que cette fonction arg *se prolonge en une fonction continue φ sur une partie Ω de $\mathbb{C}$ contenant strictement $\mathbb{C} \setminus \mathbb{R}\_$. L'ensemble Ω contient alors un réel $x < 0$ et le cercle $\mathcal{C}(0, |x|)$ est tout entier contenu dans Ω (les points autres que x sont dans $\mathbb{C} \setminus \mathbb{R}\_ \subset \Omega$). Considérant les suites $(u_n)_{n \geq 1}$ et $(v_n)_{n \geq 1}$ de $\mathcal{C}(0, |x|) \setminus \{x\}$ définies par :*

$$u_n = |x| e^{i\pi\left(1 - \frac{1}{n}\right)}, \; v_n = |x| e^{-i\pi\left(1 + \frac{1}{n}\right)}$$

on a $\lim\limits_{n \to +\infty} u_n = \lim\limits_{n \to +\infty} v_n = -|x| = x$ (la fonction exp *est continue sur $\mathbb{C}$) avec :*

$$\varphi(u_n) = \arg(u_n) = \pi \left(1 - \frac{1}{n}\right) \underset{n \to +\infty}{\to} \pi$$

et :

$$\varphi(v_n) = \arg(v_n) = -\pi \left(1 + \frac{1}{n}\right) \underset{n \to +\infty}{\to} -\pi$$

(on a $u_n = |u_n| e^{i \arg(u_n)} = |x| e^{i\pi\left(1 - \frac{1}{n}\right)}$ avec $\arg(u_n)$ et $\pi \left(1 - \dfrac{1}{n}\right)$ dans $]-\pi, \pi[$, donc $\arg(u_n) = \pi \left(1 - \dfrac{1}{n}\right)$, de même pour v_n), ce qui est incompatible avec la continuité de φ.

On est maintenant en mesure de définir une fonction logarithme complexe sur $\mathbb{C} \setminus \mathbb{R}\_$.

Précisément, on va définir une fonction notée ln sur $\mathbb{C} \setminus \mathbb{R}\_$ telle que $e^{\ln(z)} = z$ pour tout $z \in \mathbb{C} \setminus \mathbb{R}\_$.

Supposons le résultat acquis. En notant respectivement P et Q les parties réelle et imaginaire de ln, on doit avoir $z = e^{\ln(z)} = e^{P(z)+iQ(z)}$, donc $|z| = \left|e^{P(z)+iQ(z)}\right| = e^{P(z)}$ et $P(z) = \ln(|z|)$ où ln est la fonction logarithme réel réciproque de l'exponentielle réelle. En écrivant que $z = |z| e^{i \arg(z)}$, on a $e^{P(z)} e^{iQ(z)} = |z| e^{i \arg(z)}$ avec $|z| = e^{P(z)}$, donc $e^{iQ(z)} = e^{i \arg(z)}$ et $Q(z) = \arg(z) + 2k\pi$ avec $k = k(z)$ entier. Si on souhaite la fonction ln continue sur $\mathbb{C} \setminus \mathbb{R}\_$, il doit en être de même des fonctions $Q = \Im(\ln)$ et $k = \dfrac{Q - \arg}{2\pi}$. En définitive, k est une fonction continue à valeurs entières sur le connexe $\mathbb{C} \setminus \mathbb{R}\_$, elle est nécessairement constante.

On définit donc, au vu de cette analyse la détermination principale du logarithme par :

$$\begin{array}{rccc} \ln : & \mathbb{C} \setminus \mathbb{R}^- & \to & \mathbb{C} \\ & z & \mapsto & \ln(|z|) + i \arg(z) = \ln(|z|) + 2i \arctan \left(\frac{\Im(z)}{\Re(z)+|z|} \right) \end{array}$$

et cette fonction est continue sur $\mathbb{C} \setminus \mathbb{R}^-$ avec :

$$\forall z \in \mathbb{C} \setminus \mathbb{R}^-, \; e^{\ln(z)} = z.$$

Par exemple, on a $\ln(i) = i\dfrac{\pi}{2}$.

6.8.5 Mesure des angles

On note $\mathcal{P}$ le plan affine euclidien muni d'un repère orthonormé $\mathcal{R} = (O, \overrightarrow{e_1}, \overrightarrow{e_2})$ et $\overrightarrow{\mathcal{P}}$ est le plan vectoriel associé à $\mathcal{P}$.

Théorème 6.15 *Si θ est un argument de $z \in \mathbb{C}^*$ affixe d'un vecteur non nul $\overrightarrow{v}$, c'est alors une mesure de l'angle orienté $(\widehat{\overrightarrow{e_1}, \overrightarrow{v}})$.*

Démonstration. Par définition d'une mesure θ' de l'angle orienté $(\widehat{\overrightarrow{e_1}, \overrightarrow{v}})$, il existe un unique automorphisme orthogonal direct u tel que $u(\overrightarrow{e_1}) = \frac{1}{\|\overrightarrow{v}\|}\overrightarrow{v}$. Dans la base $(\overrightarrow{e_1}, \overrightarrow{e_2})$ la matrice de u est $\begin{pmatrix} \cos(\theta') & -\sin(\theta') \\ \sin(\theta') & \cos(\theta') \end{pmatrix}$ et si $z = x + iy$ est l'affixe de $\overrightarrow{v}$, on a alors :

$$\begin{aligned} \overrightarrow{v} &= x\overrightarrow{e_1} + y\overrightarrow{e_2} = \|\overrightarrow{v}\| \, u(\overrightarrow{e_1}) \\ &= |z| \left(a\overrightarrow{e_1} + b\overrightarrow{e_2} \right) = |z| \left(\cos(\theta') \, \overrightarrow{e_1} + \sin(\theta') \, \overrightarrow{e_2} \right) \end{aligned}$$

ce qui entraîne $x = |z| \cos(\theta')$, $y = |z| \sin(\theta')$ et $\theta' \equiv \theta \ (2\pi)$. ∎

Remarque 6.9 *Le choix d'une orientation de $\overrightarrow{\mathcal{P}}$ nous permet de définir sans ambiguïté la mesure principale dans $[-\pi, \pi[$ d'un angle de vecteurs. Ce choix d'une orientation correspond au choix d'une racine carrée i de -1 dans $\mathbb{C}$.*

Plus généralement on a le résultat suivant.

Théorème 6.16 *Si $\overrightarrow{v_1}$ et $\overrightarrow{v_2}$ sont deux vecteurs non nuls d'affixes respectives z_1 et z_2 alors un argument de $\frac{z_2}{z_1}$ est une mesure de l'angle orienté $\theta = (\widehat{\overrightarrow{v_1}, \overrightarrow{v_2}})$ et on a :*

$$\begin{cases} \cos(\theta) = \dfrac{\overrightarrow{v_1} \cdot \overrightarrow{v_2}}{\|\overrightarrow{v_1}\| \, \|\overrightarrow{v_2}\|} \\ \sin(\theta) = \dfrac{\det(\overrightarrow{v_1}, \overrightarrow{v_2})}{\|\overrightarrow{v_1}\| \, \|\overrightarrow{v_2}\|} \end{cases}$$

Démonstration. On a $\frac{1}{\|\overrightarrow{v_2}\|}\overrightarrow{v_2} = u\left(\frac{1}{\|\overrightarrow{v_1}\|}\overrightarrow{v_1}\right)$ où l'automorphisme orthogonal direct u a pour matrice $\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ dans la base $(\overrightarrow{e_1}, \overrightarrow{e_2})$, ce qui donne :

$$\begin{cases} x_2 = \dfrac{\|\overrightarrow{v_2}\|}{\|\overrightarrow{v_2}\|} \left(\cos(\theta) \, x_1 - \sin(\theta) \, y_1 \right) \\ \quad = \dfrac{|z_2|}{|z_1|} \left(\cos(\theta) \, x_1 - \sin(\theta) \, y_1 \right) \\ y_2 = \dfrac{\|\overrightarrow{v_2}\|}{\|\overrightarrow{v_2}\|} \left(\sin(\theta) \, x_1 + \cos(\theta) \, y_1 \right) \\ \quad = \dfrac{|z_2|}{|z_1|} \left(\sin(\theta) \, x_1 + \cos(\theta) \, y_1 \right) \end{cases}$$

et :

$$\begin{aligned}
z_2 &= x_2 + iy_2 \\
&= \frac{|z_2|}{|z_1|}\left(\left(\cos(\theta)\,x_1 - \sin(\theta)\,y_1\right) + i\left(\sin(\theta)\,x_1 + \cos(\theta)\,y_1\right)\right) \\
&= \frac{|z_2|}{|z_1|}\left(x_1 + iy_1\right)\left(\cos(\theta) + i\sin(\theta)\right) \\
&= \frac{|z_2|}{|z_1|} z_1\left(\cos(\theta) + i\sin(\theta)\right)
\end{aligned}$$

soit :

$$\frac{z_2}{z_1} = \frac{|z_2|}{|z_1|}\left(\cos(\theta) + i\sin(\theta)\right),$$

ce qui signifie que θ est un argument de $\dfrac{z_2}{z_1}$.

On rappelle que :

$$\overrightarrow{v_1} \cdot \overrightarrow{v_2} = \Re\left(\overline{z_1} z_2\right) = |z_1|^2 \, \Re\left(\frac{z_2}{z_1}\right)$$

et :

$$\det\left(\overrightarrow{v_1}, \overrightarrow{v_2}\right) = \Im\left(\overline{z_1} z_2\right) = |z_1|^2 \, \Im\left(\frac{z_2}{z_1}\right)$$

avec $\Re\left(\dfrac{z_2}{z_1}\right) = \left|\dfrac{z_2}{z_1}\right| \cos(\theta)$ et $\Im\left(\dfrac{z_2}{z_1}\right) = \left|\dfrac{z_2}{z_1}\right| \sin(\theta)$, ce qui donne compte tenu de $|z_1| = \|\overrightarrow{v_1}\|$ et $|z_2| = \|\overrightarrow{v_2}\|$:

$$\overrightarrow{v_1} \cdot \overrightarrow{v_2} = \|\overrightarrow{v_1}\| \, \|\overrightarrow{v_2}\| \cos(\theta) \text{ et } \det\left(\overrightarrow{v_1}, \overrightarrow{v_2}\right) = \|\overrightarrow{v_1}\| \, \|\overrightarrow{v_2}\| \sin(\theta)$$

■

On déduit de ce théorème que $\left(\lambda\overrightarrow{v_2}, \lambda\overrightarrow{v_1}\right) \equiv \left(\overrightarrow{v_2}, \overrightarrow{v_1}\right)$ modulo 2π pour tout réel non nul et en particulier $\left(-\overrightarrow{v_2}, -\overrightarrow{v_1}\right) \equiv \left(\overrightarrow{v_2}, \overrightarrow{v_1}\right)$ modulo 2π.

Corollaire 6.5 *Si les points A, B, C sont deux à deux distincts alors un argument de $\dfrac{c-a}{b-a}$ est une mesure de l'angle orienté $\theta_A = \left(\widehat{\overrightarrow{AB}, \overrightarrow{AC}}\right)$ et on a :*

$$\begin{cases}
\cos(\theta_A) = \dfrac{\overrightarrow{AB} \cdot \overrightarrow{AC}}{AB \cdot AC} \\
\sin(\theta_A) = \dfrac{\det\left(\overrightarrow{AB}, \overrightarrow{AC}\right)}{AB \cdot AC}
\end{cases}$$

En particulier, on a $\arg(b-a) = \left(\widehat{\overrightarrow{e_1}, \overrightarrow{AB}}\right)$ (modulo 2π).

7

Équations fonctionnelles

7.1 L'équation fonctionnelle de Cauchy : $f(x+y) = f(x) + f(y)$

On s'intéresse ici l'équation fonctionnelle de Cauchy qui consiste à déterminer toutes les fonctions $f : \mathbb{R} \to \mathbb{R}$ telles que :

$$\forall (x, y) \in \mathbb{R}^2,\ f(x+y) = f(x) + f(y) \tag{7.1}$$

Cela revient à étudier les morphismes du groupe additif $(\mathbb{R}, +)$ dans lui même.

Exemple 7.1 *Une fonction linéaire $f : x \mapsto ax$, où a est un réel donné, vérifie cette équation.*

Lemme 7.1 *Les morphismes de groupes de $(\mathbb{Q}, +)$ dans $(\mathbb{R}, +)$ sont les applications définies par :*

$$\forall r \in \mathbb{Q},\ f(r) = \lambda r$$

où λ est un réel.

Démonstration. Un morphisme de groupes $f : (\mathbb{Q}, +) \to (\mathbb{R}, +)$ est une application telle que :

$$\forall (r, s) \in \mathbb{Q}^2,\ f(r+s) = f(x) + f(y) \tag{7.2}$$

En prenant $(r, s) = (0, 0)$, on obtient $f(0) = 2f(0)$, ce qui équivaut à $f(0) = 0$ (un morphisme de groupes transforme le neutre en neutre).

En prenant $(r, s) = (r, -r)$, on obtient $f(r) + f(-r) = 0$. On a donc $f(-r) = -f(r)$ pour tout $r \in \mathbb{Q}$ (un morphisme de groupes transforme l'opposé en opposé).

De (7.2) on déduit que pour tout $s \in \mathbb{Q}$, on a :

$$\forall n \in \mathbb{N},\ f(ns) = nf(s)$$

En effet, le résultat est vrai pour $n = 0$ et le supposant vrai pour $n \geq 0$, on a :

$$f((n+1)s) = f(ns) + f(s) = nf(s) + f(s) = (n+1)f(s)$$

il est donc vrai pour tout $n \in \mathbb{N}$.

En écrivant, pour tout entier $n \in \mathbb{N} \setminus \{0\}$, que $f(s) = f\left(n\frac{s}{n}\right) = nf\left(\frac{s}{n}\right)$, on déduit que $f\left(\frac{s}{n}\right) = \frac{1}{n}f(s)$ pour tout $s \in \mathbb{Q}$ et tout $n \in \mathbb{N} \setminus \{0\}$.

Il en résulte que pour tout rationnel s et tout rationnel positif $r = \frac{p}{q}$, avec $p \in \mathbb{N}$ et $q \in \mathbb{N} \setminus \{0\}$, on a :

$$f(rs) = f\left(p\frac{s}{q}\right) = pf\left(\frac{s}{q}\right) = \frac{p}{q}f(s) = rf(s)$$

Enfin avec l'imparité de f, on déduit que ce dernier résultat est encore vrai pour les rationnels négatifs.

On a donc $f(rs) = rf(s)$ pour tout $(r, s) \in \mathbb{Q}^2$.

En prenant $s = 1$ et en notant $\lambda = f(1)$, on a $f(r) = \lambda r$ pour tout $r \in \mathbb{Q}$.

Réciproquement, il est clair que de telles applications sont morphismes de groupes de $(\mathbb{Q}, +)$ dans $(\mathbb{R}, +)$. ■

En considérant $\mathbb{R}$ comme un espace vectoriel sur le corps $\mathbb{Q}$ des rationnels, le lemme précédent s'exprime en disant que tout morphisme du groupe additif $(\mathbb{Q}, +)$ dans $(\mathbb{R}, +)$ est une application $\mathbb{Q}$-linéaire.

Dans ce qui suit nous allons montrer que si f est un morphisme du groupe additif $(\mathbb{R}, +)$ dans lui même, une hypothèse supplémentaire (continuité en un point, monotonie, f bornée sur un intervalle) va entraîner que f est $\mathbb{R}$-linéaire, ce qui revient à dire que c'est une homothétie.

On peut montrer, en utilisant le lemme de Zorn, qu'il existe une fonction vérifiant l'équation fonctionnelle de Cauchy qui n'est pas une homothétie. Pour ce faire, on se donne une $\mathbb{Q}$-base $(e_i)_{i\in I}$ de $\mathbb{R}$ (utilisation du lemme de Zorn) et l'application $\mathbb{Q}$-linéaire f définie sur $\mathbb{R}$ par $f(e_j) = 1$ et $f(e_i) = 0$ pour tout $i \in I \setminus \{j\}$ où j est fixé dans I, vérifie bien l'équation fonctionnelle (7.1) sans être une homothétie.

En utilisant les approximations décimales par défaut et par excès d'un réel, on obtient le résultat suivant.

Théorème 7.1 *Les morphismes du groupe additif* $(\mathbb{R}, +)$ *dans lui même qui sont monotones sont les homothéties.*

Démonstration. Si f est un morphisme de groupes de $(\mathbb{R}, +)$ dans $(\mathbb{R}, +)$, c'est aussi un morphisme de groupes de $(\mathbb{Q}, +)$ dans $(\mathbb{R}, +)$ et on a $f(r) = \lambda r$ pour tout $r \in \mathbb{Q}$, où $\lambda = f(1)$.

Si de plus f est croissante, on a alors $\lambda = f(1) \geq f(0) = 0$.

Pour $x \in \mathbb{R}$, on désigne par $(r_n)_{n\in\mathbb{N}}$ et $(s_n)_{n\in\mathbb{N}}$ des suites d'approximations décimales de x par défaut et par excès.

De telles suites peuvent être définies par :

$$\forall n \in \mathbb{N},\ r_n = \frac{[10^n x]}{10^n},\ s_n = r_n + \frac{1}{10^n}$$

Pour tout $n \in \mathbb{N}$, on a :

$$\lambda r_n = f(r_n) \leq f(x) \leq f(s_n) = \lambda s_n$$

et faisant tendre n vers l'infini, on en déduit que $f(x) = \lambda x$.

On procède de manière analogue pour f décroissante. ■

Corollaire 7.1 *L'identité est le seul morphisme de corps non identiquement nul de* $\mathbb{R}$ *dans lui même.*

Démonstration. Si f est morphisme du corps $\mathbb{R}$ dans lui même, on a alors $f(1) = 1$ et $f(x + y) = f(x) + f(y)$ et $f(xy) = f(x)f(y)$ pour tous x, y dans $\mathbb{R}$.

Avec $f(x^2) = (f(x))^2 \geq 0$, on déduit que $f(x) \geq 0$ pour tout $x \geq 0$ et pour $x \geq y$ dans $\mathbb{R}$, on a $f(x) - f(y) = f(x-y) \geq 0$, ce qui signifie que f est croissante.

On déduit alors du théorème précédent que $f(x) = x$ pour tout $x \in \mathbb{R}$ ($\lambda = f(1) = 1$). ■

Ce résultat sera utilisé dans l'étude de l'équation fonctionnelle $f(x \wedge y) = f(x) \wedge f(y)$ sur $\mathbb{R}^3$.

Exercice 7.1 *Déterminer toutes les fonctions $f : \mathbb{R} \to \mathbb{R}$ continues en 0 telles que :*

$$\forall (x,y) \in \mathbb{R}^2,\ f(x+y) - f(x-y) = f(x) f(y) \tag{7.3}$$

Solution 7.1 *Prenant $(x,y) = (0,0)$ dans (7.3), on obtient $f(0) = 0$.*
En prenant $(x,y) = (x,x)$, où x est un réel quelconque, on obtient $f(2x) = f(x)^2$ et en conséquence f est à valeurs positives ou nulles.
Par récurrence, on obtient $f(2^n x) = (f(x))^{2^n}$ pour tout entier naturel n et tout réel x, ce qui peut aussi s'écrire $f(x) = \left(f\left(\dfrac{x}{2^n}\right)\right)^{2^n}$, ou encore $f\left(\dfrac{x}{2^n}\right) = (f(x))^{\frac{1}{2^n}}$.
S'il existe un réel x tel que $f(x) > 0$, on a alors avec la continuité de f en 0 :

$$0 = f(0) = \lim_{n\to+\infty} f\left(\frac{x}{2^n}\right) = \lim_{n\to+\infty} (f(x))^{\frac{1}{2^n}} = 1$$

ce qui est impossible.
En conclusion f est la fonction nulle.

Exercice 7.2 *Déterminer les fonctions $f : \mathbb{R} \to \mathbb{R}$ telle que $f(1) = 1$, $f(x+y) = f(x) + f(y)$ pour tout $(x,y) \in \mathbb{R}^2$ et $f(x) f\left(\dfrac{1}{x}\right) = 1$ pour tout $x \in \mathbb{R}^*$.*

Solution 7.2 *Avec $f(x+y) = f(x) + f(y)$ pour tout $(x,y) \in \mathbb{R}^2$ on déduit que f est impaire (donc $f(0) = 0$) et avec $f(x) f\left(\dfrac{1}{x}\right) = 1$ pour tout $x \in \mathbb{R}^*$ que $f(x) \neq 0$ pour tout $x \in \mathbb{R}^*$.*
Pour tout $x \in \mathbb{R} \setminus \{0,1\}$, on a :

$$f\left(\frac{1}{x(1-x)}\right) = f\left(\frac{1}{x} + \frac{1}{1-x}\right) = f\left(\frac{1}{x}\right) + f\left(\frac{1}{1-x}\right)$$

ce qui s'écrit aussi :

$$\frac{1}{f(x(1-x))} = \frac{1}{f(x)} + \frac{1}{f(1-x)} = \frac{f(x) + f(1-x)}{f(x) f(1-x)} = \frac{1}{f(x) f(1-x)}$$

ou encore :

$$f(x) - f\left(x^2\right) = f\left(x - x^2\right) = f(x) f(1-x) = f(x)(1 - f(x))$$

soit $f(x^2) = (f(x))^2$ pour tout $x \in \mathbb{R} \setminus \{0,1\}$.
Ce résultat étant encore vrai pour $x = 0$ et $x = 1$, on en déduit que $f(x) > 0$ pour tout $x \in \mathbb{R}^{+,*}$ et pour $x \geq y$ on a alors $f(x) - f(y) = f(x-y) \geq 0$, ce qui signifie que la fonction f est croissante sur $\mathbb{R}$.
Il en résulte que $f(x) = f(1) x = x$ pour tout $x \in \mathbb{R}$.

Exercice 7.3 *Montrer que si $f : \mathbb{R} \to \mathbb{R}$ est une fonction monotone telle que :*

$$\forall (x,y) \in \mathbb{R}^2,\ f\left(\frac{x+y}{2}\right) = \frac{f(x) + f(y)}{2}$$

(équation fonctionnelle de Jensen), elle est alors affine.

Solution 7.3 *La relation* $f\left(\frac{x+y}{2}\right) = \frac{f(x)+f(y)}{2}$ *appliquée au couple* $(2x,0)$ *pour* $x \in \mathbb{R}$ *donne :*

$$f(x) = \frac{f(2x)+f(0)}{2}$$

ou encore $f(2x) = 2f(x) - f(0)$.
On note g *la fonction définie sur* $\mathbb{R}$ *par* $g(x) = f(x) - f(0)$.
Pour $(x,y) \in \mathbb{R}^2$ *on a :*

$$\begin{aligned} g(x+y) &= f\left(\frac{2x+2y}{2}\right) - f(0) = \frac{f(2x)+f(2y)}{2} - f(0) \\ &= \frac{g(2x)+g(2y)}{2} \end{aligned}$$

et pour $x \in \mathbb{R}$ *on a :*

$$g(2x) = f(2x) - f(0) = 2(f(x) - f(0)) = 2g(x)$$

On a donc $g(x+y) = g(x) + g(y)$.
Si f *est monotone, il en est alors de même de* g*, donc* $g(x) = ax$ *et* $f(x) = ax+b$ *avec* $b = f(0)$.

Exercice 7.4 *On dit qu'une suite de réels* $(x_n)_{n\in\mathbb{N}}$ *converge au sens de Cesàro vers* ℓ *si la suite* $\left(\frac{1}{n}\sum_{k=1}^{n} x_k\right)_{n\in\mathbb{N}^*}$ *est convergente vers* ℓ.
On dit qu'une fonction $f : \mathbb{R} \to \mathbb{R}$ *conserve la convergence au sens de Cesàro si pour toute suite* $(x_n)_{n\in\mathbb{N}}$ *qui converge au sens de Cesàro vers* ℓ*, la suite* $(f(x_n))_{n\in\mathbb{N}}$ *converge au sens de Cesàro vers* $f(\ell)$.
On se donne une fonction $f : \mathbb{R} \to \mathbb{R}$ *qui conserve la convergence au sens de Cesàro.*

1. *Montrer que* f *vérifie l'équation fonctionnelle de Jensen.*
2. *Montrer que pour tout couple de réels* (x,y) *et tout couple d'entiers* (n,k) *tels* $n \in \mathbb{N}$ *et* $0 \leq k \leq 2^n$*, on a :*

$$f\left(x + \frac{k}{2^n}(y-x)\right) = f(x) + \frac{k}{2^n}(f(y) - f(x))$$

3. *Montrer que* f *est une fonction affine.*

Solution 7.4

1. *Pour* $(x,y) \in \mathbb{R}^2$ *on définit la suite* $(x_n)_{n\in\mathbb{N}}$ *par :*

$$\forall k \in \mathbb{N}, \left\{ \begin{array}{l} x_{2k} = x \\ x_{2k+1} = y \end{array} \right.$$

   Avec :

$$\left\{ \begin{array}{l} \frac{1}{2n}\sum_{k=1}^{2n} x_k = \frac{1}{2n}\left(\sum_{k=1}^{n} x_{2k} + \sum_{k=0}^{n-1} x_{2k+1}\right) = \frac{x+y}{2} \\ \frac{1}{2n+1}\sum_{k=1}^{2n+1} x_k = \frac{1}{2n+1}\left(\sum_{k=1}^{n} x_{2k} + \sum_{k=0}^{n} x_{2k+1}\right) = \frac{n}{2n+1}(x+y) + \frac{y}{2n+1} \end{array} \right.$$

on déduit que cette suite converge au sens de Cesàro vers $\dfrac{x+y}{2}$ *et donc la suite* $(f(x_n))_{n\in\mathbb{N}}$ *converge au sens de Cesàro vers* $f\left(\dfrac{x+y}{2}\right)$.

On a donc :

$$f\left(\frac{x+y}{2}\right) = \lim_{n\to+\infty} \frac{1}{2n}\sum_{k=1}^{2n} f(x_k) = \frac{f(x)+f(y)}{2}$$

c'est-à-dire que f *vérifie l'équation fonctionnelle de Jensen.*

2. *Pour tout* $n \in \mathbb{N}$ *et* $k \in \{0, 2^n\}$, *le résultat est trivial.*
Pour $n = 1$ *et* $k = 1$, *c'est l'équation fonctionnelle de Jensen.*
La propriété est donc vérifiée pour $n = 0$ *et* $n = 1$.
Supposons la vraie pour $n \geq 1$ *et tout couple de réels* (x, y).
Pour tout entier k *compris entre* 1 *et* $2^{n+1} - 1$ *on a :*

$$\begin{aligned} f\left(x + \frac{k}{2^{n+1}}(y-x)\right) &= f\left(\frac{1}{2}\left(x + \frac{k}{2^n}(y-x)\right) + \frac{x}{2}\right) \\ &= \frac{1}{2} f\left(x + \frac{k}{2^n}(y-x)\right) + \frac{f(x)}{2} \end{aligned}$$

Si k *est inférieur à* 2^n, *on peut alors utiliser l'hypothèse de récurrence et on obtient :*

$$\begin{aligned} f\left(x + \frac{k}{2^{n+1}}(y-x)\right) &= \frac{1}{2}\left(f(x) + \frac{k}{2^n}(f(y) - f(x))\right) + \frac{f(x)}{2} \\ &= f(x) + \frac{k}{2^{n+1}}(f(y) - f(x)) \end{aligned}$$

Si k *est compris entre* $2^n + 1$ *et* $2^{n+1} - 1$, *en effectuant la division euclidienne de* k *par* 2^n, *on a* $k = 2^n + r$ *avec* r *compris entre* 1 *et* $2^n - 1$ *et :*

$$\begin{aligned} f\left(x + \frac{k}{2^{n+1}}(y-x)\right) &= f\left(\frac{y}{2} + \frac{1}{2}\left(\frac{r}{2^n}(y-x) + x\right)\right) \\ &= \frac{f(y)}{2} + \frac{1}{2} f\left(x + \frac{r}{2^n}(y-x)\right) \\ &= \frac{f(y)}{2} + \frac{1}{2}\left(f(x) + \frac{r}{2^n}(f(y) - f(x))\right) \\ &= f(x) + \frac{k}{2^{n+1}}(f(y) - f(x)) \end{aligned}$$

3. *Si* λ *est un réel compris entre* 0 *et* 1, *en notant* $(k_n)_{n\in\mathbb{N}}$ *la suite d'entiers définie par* $k_n = [2^n\lambda]$, *on a* $0 \leq k_n \leq 2^n$ *et* $\lim\limits_{n\to+} \dfrac{k_n}{2^n} = \lambda$.
Pour tout couple de réels (x, y) *la suite* $(z_n)_{n\in\mathbb{N}}$ *définie par* $z_n = x + \dfrac{k_n}{2^n}(y - x)$ *converge vers* $z = x + \lambda(y - x)$, *elle converge donc vers* z *au sens de Cesàro et la suite* $(f(z_n))_{n\in\mathbb{N}}$ *va converger au sens de Cesàro vers* $f(z)$.
En écrivant que :

$$f(z_n) = f(x) + \frac{k_n}{2^n}(f(y) - f(x))$$

on constate que cette suite converge au sens usuel et donc au sens de Cesàro vers $f(x) + \lambda(f(y) - f(x))$.
Avec l'unicité de la limite, on peut donc conclure que :

$$f(x + \lambda(y-x)) = f(x) + \lambda(f(y) - f(x))$$

On a donc ainsi montré que la fonction f est convexe sur $\mathbb{R}$ et en conséquence elle est continue.
En définitive f est continue et vérifie l'équation de Jensen, c'est donc une fonction affine.

Lemme 7.2 *Si f vérifiant* (7.1) *est continue à droite [resp. à gauche] en un point x_0, elle est alors continue à droite [resp. à gauche] en tout point de $\mathbb{R}$.*

Démonstration. Supposant que f soit continue à droite en x_0.

Pour $x \in \mathbb{R}$ et $h > 0$, on a $f(x+h) = f(x_0+h) - f(x_0) + f(x)$ et avec la continuité à droite en x_0 de f, on déduit que $\lim\limits_{h \to 0^+} f(x+h) = f(x)$, ce qui signifie que f est continue à droite en x. ■

Théorème 7.2 *Si f vérifiant* (7.1) *est continue à droite [resp. à gauche] en un point x_0, c'est alors une homothétie.*

Démonstration. Supposant que f soit continue à droite en x_0. Dans ce cas, elle l'est sur tout $\mathbb{R}$.

En notant $\lambda = f(1)$ et en désignant, pour $x \in \mathbb{R}$, par $(s_n)_{n\in\mathbb{N}}$ une suite d'approximations décimales par excès de x, on a :

$$f(x) = \lim_{n\to+\infty} f(s_n) = \lim_{n\to+\infty} \lambda s_n = \lambda x$$

■

Théorème 7.3 *Si f vérifiant* (7.1) *est bornée sur un intervalle $[a,b]$ avec $a < b$, c'est alors une homothétie.*

Démonstration. On suppose qu'il existe deux réels $m \le M$ tels que $m \le f(x) \le M$ pour tout $x \in [a,b]$.

1. On vérifie que f est bornée sur $[0, b-a]$.
   Pour tout $x \in [0, b-a]$, on a $x + a \in [a,b]$, donc :
   $$m \le f(x+a) = f(x) + f(a) \le M$$
   soit :
   $$m - f(a) \le f(x) \le M - f(a)$$
   pour tout $x \in [0, b-a]$.
2. On en déduit f est continue à droite en 0, donc sur tout $\mathbb{R}$.
   Pour $n \in \mathbb{N}^*$ et $x \in \left[0, \dfrac{b-a}{n}\right]$, on a $nx \in [0, b-a]$, de sorte que :
   $$m' = m - f(a) \le f(nx) = nf(x) \le M' = M - f(a)$$
   et $\dfrac{m'}{n} \le f(x) \le \dfrac{M'}{n}$.
   Pour tout réel $\varepsilon > 0$, il existe un entier $n \ge 1$ tel que $\left[\dfrac{m'}{n}, \dfrac{M'}{n}\right]$ soit contenu dans $]-\varepsilon, \varepsilon[$, donc pour tout $x \in \left[0, \dfrac{b-a}{n}\right]$, on a $-\varepsilon < f(x) < \varepsilon$. La fonction f est donc continue à droite en 0.

3. Il en résulte que f est linéaire.

■

On peut aussi s'intéresser à l'équation fonction (7.1) pour les fonctions à valeurs complexes, ce qui revient à étudier les morphismes de groupes de $(\mathbb{R}, +)$ dans $(\mathbb{C}, +)$. Pour un tel morphisme f, $g = \Re(f)$ (partie réelle de f) et $h = \Im(f)$ (partie imaginaire de f) sont des morphismes du groupes $(\mathbb{R}, +)$ dans lui même. De l'étude précédente, on déduit alors le résultat suivant.

Théorème 7.4 *Soit f un morphisme de groupes de $(\mathbb{R}, +)$ dans $(\mathbb{C}, +)$. Si l'une des deux conditions suivantes est réalisée :*

- *f est continue en un point x_0 ;*
- *f est bornée sur un intervalle non réduit à un point;*

il existe alors un nombre complexe α tel que $f(x) = \alpha x$ pour tout réel x.

Démonstration. Si f est continue en un point x_0 [resp. bornée sur un intervalle non réduit à un point], il en est alors de même des fonctions $g = \Re(f)$ et $h = \Im(f)$.

Ces deux fonctions vérifiant l'équation fonctionnelle de Cauchy, ce sont des homothéties, ce qui entraîne $f(x) = ax + ibx = \alpha x$. ■

On peut aussi s'intéresser à l'équation fonction (7.1) pour les fonctions de $\mathbb{C}$ dans $\mathbb{C}$, ce qui revient à étudier les morphismes de groupes de $(\mathbb{C}, +)$ dans lui même. Pour une telle application, on a pour tout $z = x + iy \in \mathbb{C}$, $f(z) = f(x) + f(iy)$.

Si, par exemple, f est continue en un point x_0, il en est de même des fonctions d'une variable réelle $x \mapsto f(x)$ et $y \mapsto f(iy)$ et il existe deux constantes complexes α, β telle que $f(x) = \alpha x$ et $f(iy) = \beta y$ pour tous réels x, y. On a donc :

$$f(z) = \alpha \frac{z + \overline{z}}{2} + \beta \frac{z - \overline{z}}{2i} = \gamma z + \delta \overline{z}$$

où γ, δ sont deux constantes complexes.

On a un résultat analogue pour f bornée sur un disque centré en 0 et non réduit à un point.

7.2 L'équation fonctionnelle $f(xy) = f(x) + f(y)$ sur $\mathbb{R}^*$

Si on s'intéresse à l'équation fonctionnelle qui consiste à déterminer toutes les fonctions $f : \mathbb{R} \to \mathbb{R}$ telles que :

$$\forall (x, y) \in \mathbb{R}^2,\ f(xy) = f(x) + f(y)$$

On voit facilement que la fonction identiquement nulle est la seule solution de cette équation sur $\mathbb{R}$.

En effet la fonction identiquement nulle est solution et réciproquement pour une telle fonction on a $f(0) = f(x \cdot 0) = f(x) + f(0)$ et $f(x) = 0$ pour tout réel x.

On s'intéresse donc à cette équation fonctionnelle pour les fonctions définies sur $\mathbb{R}^*$.

Si f est une solution, avec $f(1) = f(1^2) = 2f(1)$, on déduit alors que $f(1) = 0$, puis avec $0 = f(1) = f\left((-1)^2\right) = 2f(-1)$ que $f(-1) = 0$ et enfin avec $f(-x) = f(-1) + f(x)$ que $f(-x) = f(x)$ pour tout $x \in \mathbb{R}$, c'est-à-dire que f est paire.

Il nous suffit donc d'étudier l'équation fonctionnelle :

$$\forall (x, y) \in \left(\mathbb{R}^{+,*}\right)^2,\ f(xy) = f(x) + f(y) \tag{7.4}$$

On s'intéresse tout d'abord aux solutions dérivables sur $\mathbb{R}^{+,*}$ de cette équation fonctionnelle.

Lemme 7.3 *Les solutions dérivables sur* $\mathbb{R}^{+,*}$ *de* (7.4) *sont les fonctions définies par :*

$$\forall x \in \mathbb{R}^{+,*},\ f(x) = \lambda \int_1^x \frac{dt}{t}$$

où λ *est une constante réelle.*

Démonstration. En se fixant $x > 0$ et en dérivant par rapport à la variable y, on a $xf'(xy) = f'(y)$, ce qui donne en prenant $y = 1$ et en notant $\lambda = f'(1)$, $f'(x) = \frac{\lambda}{x}$ pour tout $x > 0$. Considérant que $f(1) = 0$, on en déduit que $f(x) = \lambda \int_1^x \frac{dt}{t}$. Réciproquement si f est ainsi définie sur $\mathbb{R}^{+,*}$, alors $f(1) = 0$ et, pour tout $y > 0$ fixé, la dérivée de la fonction $x \mapsto f(xy)$ est égale à $yf'(xy) = \lambda y \frac{1}{xy} = \lambda \frac{1}{x} = f'(x)$, ce qui donne $f(xy) = f(x) + C$, la constante C étant égale à $f(y) - f(1) = f(y)$. La fonction f vérifie bien l'équation fonctionnelle (7.4). ■

Ce résultat nous conduit à définir la fonction logarithme comme la primitive sur $\mathbb{R}^{+,*}$ nulle en 1 de la fonction $x \mapsto \frac{1}{x}$, soit $\ln(x) = \int_1^x \frac{dt}{t}$ pour tout $x > 0$. Les résultats précédents nous disent alors que les solutions dérivables sur $\mathbb{R}^*$ de l'équation fonctionnelle (7.4) sont les fonctions définies par $f(x) = \lambda \ln(|x|)$ pour tout $x \in \mathbb{R}^*$.

Avec $\ln'(x) > 0$ sur $\mathbb{R}^{+,*}$ on déduit que ln est strictement croissante sur $\mathbb{R}^{+,*}$ et avec $\ln(2) > \ln(1) = 0$, $\ln(2^n) = n\ln(2)$ que cette fonction n'est pas bornée. On a donc $\lim\limits_{x \to +\infty} \ln(x) = +\infty$.

Avec $\ln\left(\frac{1}{x}\right) = -\ln(x)$ (conséquence de $\ln\left(x\frac{1}{x}\right) = 0$), on déduit que $\lim\limits_{x \to 0} \ln(x) = -\infty$.

La fonction ln est donc continue strictement croissante de $\mathbb{R}^{+,*}$ sur $\mathbb{R}$, c'est donc un homéomorphisme et sa fonction réciproque est la fonction exponentielle $x \mapsto e^x$. Cette fonction est donc définie par :

$$(x \in \mathbb{R} \text{ et } y = e^x) \Leftrightarrow (y > 0 \text{ et } x = \ln(y))$$

Avec $\ln'(x) > 0$, on déduit du théorème de dérivation des fonctions réciproques que la fonction exponentielle est dérivable avec $(e^x)' = e^x$ pour tout $x \in \mathbb{R}$.

De $\ln(xy) = \ln(x) + \ln(y)$ pour tous x, y dans $\mathbb{R}^{+,*}$, on déduit que la fonction exponentielle est une solution de l'équation fonctionnelle :

$$\forall (x, y) \in \mathbb{R}^2,\ f(x + y) = f(x) f(y) \tag{7.5}$$

Si f est solution de (7.4) sur $\mathbb{R}^{+,*}$, alors la fonction g définie sur $\mathbb{R}$ par $g(t) = f(e^t)$ vérifie l'équation fonctionnelle de Cauchy et on déduit le résultat suivant.

Théorème 7.5 *Soit* f *vérifiant l'équation fonctionnelle* (7.4) *sur* $\mathbb{R}^{+,*}$. *Si l'une des conditions suivantes est vérifiée :*

- f *est monotone ;*
- f *est continue à droite en un point* $x_0 \in \mathbb{R}^{+,*}$ *;*
- f *est bornée sur un intervalle non réduit à un point dans* $\mathbb{R}^{+,*}$ *;*

il existe alors une constante réelle λ *telle que* $f(x) = \lambda \ln(x)$ *pour tout réel* $x \in \mathbb{R}^{+,*}$.

Démonstration. Si f est monotone [resp. continue à droite en $x_0 \in \mathbb{R}^{+,*}$, resp. bornée sur $[a, b]$ avec $0 < a < b$], la fonction g est alors monotone (la fonction exponentielle est croissante), [resp. continue à droite en $\ln(x_0)$, resp. bornée sur $[\ln(a), \ln(b)]$] donc $g(t) = \alpha t$ et pour tout $x > 0$, on a :

$$f(x) = f\left(e^t\right) = g(t) = \alpha t = \alpha \ln(x)$$

■

7.3 L'équation fonctionnelle $f(x+y) = f(x) f(y)$ sur $\mathbb{R}$

En utilisant la fonction ln et les résultats sur l'équation fonctionnelle de Cauchy, on peut étudier l'équation fonctionnelle (7.5).

Lemme 7.4 *Si f non identiquement nulle vérifie l'équation fonctionnelle* (7.5), *elle est alors à valeurs strictement positives et la fonction $g = \ln \circ f$ vérifie l'équation fonctionnelle de Cauchy.*

Démonstration. Pour $x \in \mathbb{R}$ on a $f(x) = f\left(\frac{x}{2} + \frac{x}{2}\right) = \left(f\left(\frac{x}{2}\right)\right)^2 \geq 0.$

Si il existe $x_0 \in \mathbb{R}$ tel que $f(x_0) = 0$, on a alors pour tout tout réel x :

$$f(x) = f(x - x_0 + x_0) = f(x - x_0) f(x_0) = 0$$

et f est identiquement nulle.

On a donc $f(x) > 0$ pour tout $x \in \mathbb{R}$ si f non identiquement nulle vérifie (7.5).

La fonction $g = \ln \circ f$ est alors bien définie et vérifie (7.1). ∎

Théorème 7.6 *Soit f non identiquement nulle vérifiant l'équation fonctionnelle* (7.5). *Si l'une des conditions suivante est réalisée :*

- *f est monotone ;*
- *f est continue à droite en un point x_0 ;*
- *f est bornée sur un intervalle réel non réduit à un point ;*

il existe alors un réel α tel que $f(x) = e^{\alpha x}$ pour tout $x \in \mathbb{R}$.

Démonstration.

1. Si f est monotone, il en est de même de $g = \ln \circ f$ (la fonction ln est croissante). On a donc $g(x) = \alpha x$ pour tout $x \in \mathbb{R}$ avec $\alpha = g(1) = \ln(f(1))$ et $f(x) = e^{g(x)} = e^{\alpha x}$.
2. Avec la continuité de la fonction ln on déduit que si f est continue à droite en x_0, il en est alors de même de $g = \ln \circ f$ ($f(x_0) > 0$ si f n'est pas identiquement nulle). D'où le résultat.
3. Si f est bornée sur un intervalle $[a, b]$, on désigne par M un majorant strictement positif de f sur cet intervalle. Pour tout réel $x \in [0, b-a]$ on a, $a + x \in [a, b]$, $b - x \in [a, b]$ et :

$$f(x) = \frac{f(b)}{f(b-x)} \geqslant \frac{f(b)}{M} > 0, \ f(x) = \frac{f(x+a)}{f(a)} \leq \frac{M}{f(a)}$$

($f(t) > 0$ pour tout $t \in \mathbb{R}$). La fonction f est donc bornée sur $[0, b-a]$ avec une borne inférieure strictement positive. Il en résulte que la fonction $g = \ln \circ f$ est bornée sur $[0, b-a]$ (l'image d'un compact par l'application continue ln est un compact) et on peut conclure.

∎

7.4 L'équation fonctionnelle $f(xy) = f(x) f(y)$ sur $\mathbb{R}^{+,*}$

Si f est solution de l'équation fonctionnelle :

$$\forall (x, y) \in \left(\mathbb{R}^{+,*}\right)^2, \ f(xy) = f(x) f(y)$$

la fonction g définie sur $\mathbb{R}$ par $g(x) = f(e^x)$ vérifie alors l'équation fonctionnelle (7.5).

Si f est monotone, il en est de même de g (la fonction exponentielle est croissante), donc $g(x) = e^{\alpha x}$ et pour tout $t > 0$, on a :

$$f(t) = f(e^x) = g(x) = e^{\alpha x} = e^{\alpha \ln(t)} = t^\alpha$$

Si f est continue à droite en un point $x_0 \in \mathbb{R}^{+,*}$, alors g est continue à droite en $\ln(x_0)$ et on a $f(t) = t^\alpha$ pour tout $t > 0$.

Si f est bornée sur un intervalle $[a, b]$ avec $0 < a < b$, la fonction g est alors bornée sur $[\ln(a), \ln(b)]$ et on a $f(t) = t^\alpha$ pour tout $t > 0$.

Exercice 7.5 *Soit $f : \mathbb{C} \to \mathbb{R}$ telle que $f(1) = 1$ et $f(z_1 z_2) = f(z_1) f(z_2)$ pour tous z_1, z_2 dans $\mathbb{C}$.*

1. *Montrer que $f(z) > 0$ pour tout $z \in \mathbb{C}^*$.*
   On suppose que f est continue en 0 et en 1.
2. *Montrer qu'il existe un réel α positif ou nul tel que $f(x) = x^\alpha$ pour tout $x \in \mathbb{R}^{+,*}$.*
3. *Montrer que $f(e^{it}) = 1$ pour tout réel t.*
4. *Conclure.*

Solution 7.5

1. *Tout $z \in \mathbb{C}^*$ peut s'écrire $z = u^2$ avec $u \in \mathbb{C}^*$ et $f(z) = (f(u))^2 \geq 0$.*
   De plus avec $1 = f\left(z\frac{1}{z}\right) = f(z) f\left(\frac{1}{z}\right)$, on déduit que $f(z) > 0$ pour tout $z \in \mathbb{C}^*$.
2. *La restriction g de la fonction f à $\mathbb{R}^{+,*}$ est continue en 1 et vérifie l'équation fonctionnelle $g(xy) = g(x) g(y)$, il en résulte que g est une fonction exponentielle, c'est-à-dire qu'il existe un réel α tel que $f(x) = x^\alpha$ pour tout $x \in \mathbb{R}^{+,*}$.*
   Si de plus f est continue en 0, on a alors $\lim\limits_{x \to 0^+} g(x) = f(0) \in \mathbb{R}$ et α est nécessairement positif ou nul.
3. *La fonction $h : t \mapsto f(e^{it})$ est définie sur $\mathbb{R}$, continue en 0 et vérifie l'équation fonctionnelle $h(x+y) = h(x) h(y)$, il en résulte l'existence d'un réel β tel que $h(t) = e^{\beta t}$ pour tout $t \in \mathbb{R}$.*
   Avec $h(2n\pi) = f(1) = 1$, on déduit que $e^{\beta 2n\pi} = 1$ pour tout entier naturel n et nécessairement $\beta = 0$.
   On a donc $f(e^{it}) = 1$ pour tout réel t.
4. *Pour $z = \rho e^{it}$ dans $\mathbb{C}^*$, on a $f(z) = f(\rho) f(e^{it}) = \rho^\alpha = |z|^\alpha$ avec $\alpha \geq 0$.*
   Si $\alpha = 0$ alors $f(z) = 1$ pour tout $z \in \mathbb{C}^*$ et le résultat est encore vrai en 0 par continuité.
   Si $\alpha > 0$ alors $f(0) = \lim\limits_{z \to 0} f(z) = 0$.
   On a donc, pour tout $z \in \mathbb{C}$, $f(z) = |z|^\alpha$ avec $\alpha \geq 0$.

7.5 L'équation fonctionnelle $f(x+y) + f(x-y) = 2f(x) f(y)$ sur $\mathbb{R}$

On suppose connues les fonctions trigonométriques et hyperboliques.

Les fonctions cos et ch vérifient l'équation fonctionnelle :

$$\forall (x, y) \in \mathbb{R}^2, \ f(x+y) + f(x-y) = 2f(x) f(y) \tag{7.6}$$

Lemme 7.5 *Si f est une fonction non identiquement nulle vérifiant l'équation fonctionnelle* (7.6), *on a alors $f(0)=1$ et la fonction f est paire.*

Démonstration. L'équation (7.6) appliquée au couple $(x,0)$ donne $f(x)=f(x)f(0)$ pour tout $x\in\mathbb{R}$, ce qui équivaut à $f(0)=1$ puisque f n'est pas identiquement nulle.

Cette équation appliquée au couple $(0,y)$ pour $y\in\mathbb{R}$ donne $f(y)+f(-y)=2f(y)$, soit $f(-y)=f(y)$, la fonction f est donc paire. ∎

Lemme 7.6 *Si f est une fonction continue sur $\mathbb{R}$ qui vérifie l'équation fonctionnelle* (7.6), *elle est de alors de classe $\mathcal{C}^2$ sur $\mathbb{R}$ et solution de l'équation différentielle $f''=f''(0)f$ avec les conditions initiales $f(0)=1$ et $f'(0)=0$.*

Démonstration. On note F la primitive de f nulle en 0.

Pour $y\in\mathbb{R}$ fixé, on a par intégration, pour tout réel x :

$$\int_0^x f(t+y)\,dt+\int_0^x f(t-y)\,dt=2f(y)\int_0^x f(t)\,dt$$

ou encore :

$$\int_y^{x+y} f(z)\,dz+\int_{-y}^{x-y} f(z)\,dz=2f(y)\int_0^x f(t)\,dt$$

soit :

$$F(x+y)-F(y)+F(x-y)-F(-y)=2f(y)F(x)$$

Si $F=0$, on a alors $f=F'=0$.

Si $F\neq 0$, il existe alors x_0 tel que $F(x_0)\neq 0$ et on en déduit que f est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Avec :

$$\begin{aligned}2f'(y)F(x_0)&=F'(x+y)-F'(y)-F'(x-y)+F'(-y)\\&=f(x+y)-f(y)-f(x-y)+f(-y)\\&=f(x+y)-f(x-y)\end{aligned}$$

on en déduit que f est de classe $\mathcal{C}^2$ sur $\mathbb{R}$.

En dérivant deux fois l'équation (7.6) par rapport à x, pour y réel fixé, on obtient :

$$f''(x+y)+f''(x-y)=2f''(x)f(y)$$

et pour $x=0$, en tenant compte de la parité de f, on a $f''(y)=f''(0)f(y)$ avec $f(0)=1$.

En dérivant par rapport à y et en faisant $y=0$, on obtient $2f(x)f'(0)=0$ pour tout réel x, ce qui entraîne $f'(0)=0$ puisque f n'est pas la fonction nulle.

En définitive, f est solution de $f''=f''(0)f$ avec $f(0)=1$ et $f'(0)=0$. ∎

En résolvant l'équation différentielle $f''=f''(0)f$, on aboutit au résultat suivant.

Théorème 7.7 *Si f est une fonction continue sur $\mathbb{R}$ qui vérifie l'équation fonctionnelle* (7.6), *on a alors pour tout réel x, $f(x)=\cos(\lambda x)$ si $f''(0)=-\lambda^2\leq 0$ ou $f(x)=\operatorname{ch}(\lambda x)$ si $f''(0)=\lambda^2\geq 0$.*

L'étude de l'équation fonctionnelle (7.6) peut aussi se faire en utilisant les approximations dyadiques des réels.

Exercice 7.6 *On se donne une fonction continue non identiquement nulle $f:\mathbb{R}\to\mathbb{R}$ qui vérifie l'équation fonctionnelle* (7.6).
On sait déjà que $f(0)=1$ et que f est paire.

1. *Justifier l'existence d'un réel* $\alpha > 0$ *tel que* $f(x) > 0$ *pour tout* $x \in [-\alpha, \alpha]$.
2. *On suppose que* $0 < f(\alpha) \leq 1$ *et* $\theta \in \left[0, \frac{\pi}{2}\right]$ *est défini par* $f(\alpha) = \cos(\theta)$.
   (a) *Montrer que :*
   $$\forall n \in \mathbb{N},\ f\left(\frac{\alpha}{2^n}\right) = \cos\left(\frac{\theta}{2^n}\right)$$
   (b) *Montrer que, pour tout* n *fixé dans* $\mathbb{N}$, *on a :*
   $$\forall p \in \mathbb{N},\ f\left(p\frac{\alpha}{2^n}\right) = \cos\left(p\frac{\theta}{2^n}\right)$$
   (c) *En déduire qu'il existe un réel* $\lambda > 0$ *tel que* $f(x) = \cos(\lambda x)$ *pour tout* $x \in \mathbb{R}$.
3. *Dans le cas où* $|f(\alpha)| > 1$, *montrer qu'il existe un réel* $\lambda > 0$ *tel que* $f(x) = \mathrm{ch}(\lambda x)$.

Solution 7.6

1. *Avec la continuité en* 0 *et* $f(0) = 1$ *on déduit qu'il existe un réel* $\alpha > 0$ *tel que* $f(x) > 0$ *pour tout* $x \in [-\alpha, \alpha]$.
2.
   (a) *Le résultat est vrai pour* $n = 0$ *par définition de* θ *et le supposant vrai pour* $n \geq 0$, *en appliquant* (7.6) *au couple* $\left(\frac{\alpha}{2^{n+1}}, \frac{\alpha}{2^{n+1}}\right)$, *on obtient :*
   $$f^2\left(\frac{\alpha}{2^{n+1}}\right) = \frac{f\left(\frac{\alpha}{2^n}\right) + 1}{2} = \frac{\cos\left(\frac{\theta}{2^n}\right) + 1}{2} = \cos^2\left(\frac{\theta}{2^{n+1}}\right)$$
   ce qui entraîne $f\left(\frac{\alpha}{2^{n+1}}\right) = \cos\left(\frac{\theta}{2^{n+1}}\right)$ *du fait de la positivité de* f *sur* $[0, \alpha]$ *et de* $\cos$ *sur* $\left[0, \frac{\pi}{2}\right]$.

   (b) *Le résultat est vrai pour* $p = 0$, $p = 1$ *et le supposant vrai pour* $p \geq 1$, *on a :*
   $$\begin{aligned} f\left((p+1)\frac{\alpha}{2^n}\right) &= 2f\left(p\frac{\alpha}{2^n}\right) f\left(\frac{\alpha}{2^n}\right) - f\left((p-1)\frac{\alpha}{2^n}\right) \\ &= 2\cos\left(p\frac{\theta}{2^n}\right)\cos\left(\frac{\theta}{2^n}\right) - \cos\left((p-1)\frac{\theta}{2^n}\right) \\ &= \cos\left((p+1)\frac{\theta}{2^n}\right) \end{aligned}$$

   (c) *Avec* $f(0) = \cos(0) = 1$ *et la parité des fonctions* f *et* $\cos$ *il nous suffit de montrer le résultat pour* $x > 0$.
   Pour $x > 0$, *on définit les suites* $(r_n)_{n\in\mathbb{N}}$ *et* $(s_n)_{n\in\mathbb{N}}$ *par :*
   $$\forall n \in \mathbb{N},\ r_n = \frac{[2^n x]}{2^n} = \frac{p_n}{2^n},\ s_n = r_n + \frac{1}{2^n}$$
   On a $0 \leq x - r_n < \frac{1}{2^n}$ *pour tout* $n \in \mathbb{N}$, *ce qui entraîne la convergence de ces suites vers* x *(développement dyadique par défaut et par excès). Avec la continuité des fonctions* f *et* $\cos$, *on en déduit que :*
   $$\begin{aligned} f(\alpha x) &= \lim_{n\to+\infty} f(\alpha r_n) = \lim_{n\to+\infty} f\left(p_n\frac{\alpha}{2^n}\right) = \lim_{n\to+\infty} \cos\left(p_n\frac{\theta}{2^n}\right) \\ &= \lim_{n\to+\infty} \cos(r_n\theta) = \cos(\theta x) \end{aligned}$$

Il en résulte que $f(x) = \cos\left(\dfrac{\theta}{\alpha}x\right) = \cos(\lambda x)$.

(d) Si $|f(\alpha)| > 1$ *en écrivant que* $f(\alpha) = \mathrm{ch}(\theta)$ *avec* $\theta > 0$, *on aboutit à* $f(x) = \mathrm{ch}(\lambda x)$.

7.6 L'exponentielle complexe

On note $\Gamma = \{z \in \mathbb{C} \mid |z| = 1\}$ le groupe unitaire dans $\mathbb{C}^*$.

Lemme 7.7 *Les seuls morphismes de groupes continus de* $(\mathbb{R}, +)$ *dans* $(\Gamma, \cdot)$ *sont les applications* $x \mapsto e^{i\alpha x}$ *avec* $\alpha \in \mathbb{R}$.

Démonstration. Il est clair que pour tout $\alpha \in \mathbb{R}$ l'application $x \mapsto e^{i\alpha x}$ est un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$.

Réciproquement si $f : \mathbb{R} \to \Gamma$ est un morphisme de groupes, il existe alors un unique réel $\alpha \in [0, 2\pi[$ tel que $f(1) = e^{i\alpha}$.

Par récurrence on vérifie facilement que $f(n) = e^{in\alpha}$ pour tout $n \in \mathbb{N}$, puis avec $e^{i\alpha} = f\left(n\dfrac{1}{n}\right) = f\left(\dfrac{1}{n}\right)^n$ on déduit que $f\left(\dfrac{1}{n}\right) = e^{i\frac{\alpha}{n}}$ pour tout $n \geq 1$.

Il en résulte que $f(r) = e^{ir\alpha}$ pour tout $r \in \mathbb{Q}$.

Si de plus f est continue, avec la densité de $\mathbb{Q}$ dans $\mathbb{R}$, on déduit que $f(x) = e^{i\alpha x}$ pour tout $x \in \mathbb{R}^+$.

Enfin avec $1 = f(x - x) = f(x) f(-x)$ (le neutre est transformé en neutre), on déduit que $f(-x) = \dfrac{1}{f(x)} = \overline{f(x)}$ pour tout $x \in \mathbb{R}$.

Il en résulte que $f(x) = e^{i\alpha x}$ pour tout $x \in \mathbb{R}$. ■

Théorème 7.8 *Les seuls morphismes de groupes continus de* $(\mathbb{R}, +)$ *dans* $(\mathbb{C}^*, \cdot)$ *sont les applications* $x \mapsto e^{\alpha x}$ *avec* $\alpha \in \mathbb{C}$.

Démonstration. Si f est un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\mathbb{C}^*, \cdot)$, la fonction $|f|$ est alors un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\mathbb{R}^*, \cdot)$, donc il existe un réel a tel que $|f(x)| = e^{ax}$ pour tout $x \in \mathbb{R}$.

La fonction $g : x \mapsto f(x) e^{-ax}$ est alors un morphisme de groupes continu de $(\mathbb{R}, +)$ dans $(\Gamma, \cdot)$ et il existe un réel b tel que $f(x) e^{-ax} = e^{ibx}$ pour tout $x \in \mathbb{R}$.

On a donc $f(x) = e^{\alpha x}$ pour tout $x \in \mathbb{R}$ avec $\alpha = a + ib \in \mathbb{C}$.

La réciproque est évidente. ■

7.7 L'équation fonctionnelle $f(x+1) = xf(x)$. Fonction Γ

On rappelle d'abord quelques résultats sur la fonction Γ d'Euler que l'on se contente de définir sur $\mathbb{R}^{+,*}$.

On se reportera au volume sur l'intégration et les équations différentielles pour les outils d'intégration utilisés.

Définition 7.1 *La fonction gamma d'Euler est la fonction définie sur* $\mathbb{R}^{+,*}$ *par :*

$$\forall z \in \mathbb{R}^{+,*},\ \Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t} dt$$

Pour tout $x > 0$, la fonction $t \mapsto t^{x-1}e^{-t}$ est continue sur $\mathbb{R}^{+,*}$ et avec $0 < t^{x-1}e^{-t} = \underset{t\to+\infty}{o}\left(e^{-\frac{t}{2}}\right)$, on déduit que l'intégrale $\displaystyle\int_1^{+\infty} t^{x-1}e^{-t}dt$ converge.

Avec $0 < t^{x-1}e^{-t} \underset{t\to 0^+}{\sim} t^{x-1}$, on déduit que l'intégrale $\displaystyle\int_0^1 t^{x-1}e^{-t}dt$ converge.

Théorème 7.9 *La fonction gamma vérifie l'équation fonctionnelle :*

$$\forall z \in \mathbb{R}^{+,*},\ \Gamma(x+1) = x\Gamma(x) \tag{7.7}$$

Démonstration. Une intégration par parties donne pour $z \in \mathbb{R}^{+,*}$ et $0 < \varepsilon < R$:

$$\int_\varepsilon^R t^x e^{-t}dt = \left[-t^x e^{-t}\right]_\varepsilon^R + x\int_\varepsilon^R t^{x-1}e^{-t}dt$$

et le passage à la limite quand (ε, R) tend vers $(0, +\infty)$ donne le résultat. ■

Une définition équivalente de la fonction gamma est donnée par le théorème d'Euler qui suit.

Pour tout entier $n \geq 1$ et tout réel $x > 0$, on note :

$$I_n(x) = \frac{n!n^x}{x(x+1)\cdots(x+n)}$$

Lemme 7.8 *Pour tout entier $n \geq 1$ et tout réel $x > 0$, on a :*

$$\int_0^n \left(1-\frac{t}{n}\right)^n t^{x-1}dt = I_n(x)$$

Démonstration. Pour $n \geq 1$ et $x > 0$, le changement de variable $t = nu$ nous donne :

$$\int_0^n \left(1-\frac{t}{n}\right)^n t^{x-1}dt = \int_0^1 (1-u)^n u^{x-1}n^x du = n^x J_n(x)$$

Une intégration par parties nous donne :

$$J_{n+1}(x) = \int_0^1 (1-u)^{n+1} u^{x-1}du = \frac{n+1}{x}\int_0^1 (1-u)^n u^x du = \frac{n+1}{x}J_n(x+1)$$

et par récurrence, on déduit que :

$$\begin{aligned} J_n(x) &= \frac{n!}{x(x+1)\cdots(x+n-1)}J_0(x+n) \\ &= \frac{n!}{x(x+1)\cdots(x+n-1)}\int_0^1 t^{x+n-1}dt \\ &= \frac{n!}{x(x+1)\cdots(x+n-1)(x+n)} \end{aligned}$$

On peut aussi écrire que :

$$\int_0^n \left(1-\frac{t}{n}\right)^n t^{x-1}dt = \int_0^n \sum_{k=0}^n \binom{n}{k}(-1)^k \frac{t^{k+x-1}}{n^k}dt = n^x\sum_{k=0}^n \frac{\binom{n}{k}(-1)^k}{k+x}$$

et utiliser la décomposition en éléments simples :

$$P_n(x) = \frac{1}{x(x+1)\cdots(x+n)} = \sum_{k=0}^n \frac{\alpha_k}{x+z}$$

les coefficients α_k étant donnés par :

$$\alpha_k = \left((x+k) P_n(x)\right)_{|x=-k} = \frac{(-1)^k}{k!\,(n-k)!} = \frac{(-1)^k \binom{n}{k}}{n!}$$

■

Théorème 7.10 (Euler) *On a :*

$$\forall x \in \mathbb{R}^{+,*},\ \Gamma(x) = \lim_{n\to+\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}$$

Démonstration. On désigne par $(f_n)_{n\in\mathbb{N}^*}$ la suite de fonctions définies sur $\mathbb{R}^{+,*}$ par :

$$f_n(t) = \begin{cases} \left(1 - \dfrac{t}{n}\right)^n t^{x-1} \text{ si } t \in \left]0, n\right[\\ 0 \text{ si } t \geq n \end{cases}$$

Chaque fonction f_n est continue et intégrable sur $\left]0, +\infty\right[$ avec :

$$\forall t \in \left]0, +\infty\right[,\ \begin{cases} \lim\limits_{n\to+\infty} f_n(t) = e^{-t}t^{x-1} \\ \forall n \geq 1,\ 0 \leq f_n(t) \leq e^{-t}t^{x-1} = f(t) \end{cases}$$

(pour tout réel $t \in [0, n]$, on a $\left(1 - \dfrac{t}{n}\right)^n \leq e^{-t}$), la fonction f étant continue et intégrable sur $\left]0, +\infty\right[$.

On déduit alors du théorème de convergence dominée que :

$$\lim_{n\to+\infty} \int_0^n \left(1 - \frac{t}{n}\right)^n t^{x-1}dt = \int_0^{+\infty} e^{-t}t^{x-1}dt = \Gamma(x)$$

soit :

$$\Gamma(x) = \lim_{n\to+\infty} I_n(x) = \lim_{n\to+\infty} \frac{n!n^x}{x(x+1)\cdots(x+n)}$$

■

En utilisant des résultats sur la convexité et quelques propriétés de la fonction Γ d'Euler, on donne une caractérisation de cette fonction par l'équation fonctionnelle $f(x+1) = xf(x)$ sur $\mathbb{R}^{+,*}$.

On se donne une fonction $f : \mathbb{R}^{+,*} \to \mathbb{R}^{+,*}$ telle que :

(i) $\forall x \in \mathbb{R}^{+,*},\ f(x+1) = xf(x)$;

(ii) $f(1) = 1$;

(iii) f est logarithmiquement convexe.

On note $g = \ln(f)$.

La fonction g étant convexe sur $\mathbb{R}^{+,*}$, elle est continue et il en est de même de la fonction $f = e^g$.

Lemme 7.9 *Pour tout réel $x \in \mathbb{R}^{+,*}$ et tout entier $n \in \mathbb{N}$, on a :*

$$f(n+1) = n! \text{ et } g(n+1+x) - g(n+1) = \ln\left(\frac{f(x)}{n!}\prod_{k=0}^{n}(x+k)\right)$$

Démonstration. De (i), on déduit par récurrence sur $n \in \mathbb{N}$ que :

$$f(n+1+x) = f(x)\prod_{k=0}^{n}(x+k)$$

($x \in \mathbb{R}^{+,*}$ étant fixé).

En effet, pour $n=0$, c'est la condition (i) et supposant le résultat acquis pour $n-1 \geq 0$, on a :

$$f(n+1+x) = (n+x)f(n+x) = (n+x)f(x)\prod_{k=0}^{n-1}(x+k) = f(x)\prod_{k=0}^{n}(x+k)$$

Prenant $x=1$, on obtient $f(n+1) = f(1)\,n! = n!$ (condition (ii)).

Il en résulte que :

$$g(n+1+x) - g(n+1) = \ln\left(\frac{f(n+1+x)}{f(n+1)}\right) = \ln\left(\frac{f(x)}{n!}\prod_{k=0}^{n}(x+k)\right)$$

∎

Lemme 7.10 *Pour tout réel $x \in \left]0,1\right]$ et tout entier $n \in \mathbb{N}^*$, on a :*

$$\ln(n) \leq \frac{g(n+1+x) - g(n+1)}{x} \leq \ln(n+1)$$

Démonstration. On note, pour tous réels $x \neq y$ dans $\mathbb{R}^{+,*}$, $p(x,y) = \dfrac{g(x)-g(y)}{x-y}$.

On remarque que $p(x,y) = p(y,x)$ pour tous réels $x \neq y$ dans $\mathbb{R}^{+,*}$.

Dire que la fonction g est convexe sur $\mathbb{R}^{+,*}$ équivaut à dire que, pour tout $y \in \mathbb{R}^{+,*}$, la fonction $x \mapsto p(x,y) = p(y,x)$ est croissante sur $\mathbb{R}^{+,*} - \{y\}$.

Il en résulte que, pour tout réel $x \in \left]0,1\right]$, on a :

$$p(n,n+1) \leq p(n+1,n+1+x) \leq p(n+1,n+2)$$

soit :

$$\ln\left(\frac{f(n+1)}{f(n)}\right) = \ln(n) \leq \frac{g(n+1+x)-g(n+1)}{x} \leq \ln\left(\frac{f(n+2)}{f(n+1)}\right) = \ln(n+1)$$

∎

Lemme 7.11 *Pour tout réel $x \in \left]0,1\right]$ et tout entier $n \in \mathbb{N}^*$, on a :*

$$n^x \leq \frac{f(x)}{n!}\prod_{k=0}^{n}(x+k) \leq (n+1)^x$$

Démonstration. Des résultats précédents, on déduit que :

$$\ln(n) \leq \frac{1}{x}\ln\left(\frac{f(x)}{n!}\prod_{k=0}^{n}(x+k)\right) \leq \ln(n+1)$$

soit :

$$\ln(n^x) \leq \ln\left(\frac{f(x)}{n!}\prod_{k=0}^{n}(x+k)\right) \leq \ln(n+1)^x$$

ou encore :

$$n^x \leq \frac{f(x)}{n!}\prod_{k=0}^{n}(x+k) \leq (n+1)^x$$

■

Théorème 7.11 (Bohr-Mollerup) *La fonction Γ d'Euler est l'unique fonction qui vérifie les conditions (i), (ii) et (iii).*

Démonstration. En utilisant les lemmes précédents, on déduit que si $f : \mathbb{R}^{+,*} \to \mathbb{R}^{+,*}$ vérifie les conditions (i), (ii) et (iii), on a alors pour tout $x \in]0,1]$ et tout $n \in \mathbb{N}^*$:

$$1 \leq f(x)\frac{x(x+1)\cdots(x+n)}{n^x n!} \leq \left(1+\frac{1}{n}\right)^x$$

et faisant tendre n vers l'infini on déduit que :

$$f(x) = \lim_{n\to+\infty}\frac{n!n^x}{x(x+1)\cdots(x+n)} = \Gamma(x)$$

(théorème d'Euler).

En utilisant l'équation fonctionnelle (i) vérifiée par les deux fonctions f et Γ, on déduit que $f(x) = \Gamma(x)$ pour tout $x > 0$. ■

7.8 L'équation fonctionnelle $f(x \wedge y) = f(x) \wedge f(y)$ sur $\mathbb{R}^3$

On montre ici que les rotations de $\mathbb{R}^3$ sont les seules solutions non identiquement nulles de l'équation fonctionnelle :

$$\forall (x,y) \in \left(\mathbb{R}^3\right)^2,\ f(x \wedge y) = f(x) \wedge f(y) \tag{7.8}$$

L'espace vectoriel $\mathbb{R}^3$ est muni de sa base canonique $\mathcal{B} = (e_1, e_2, e_3)$ et de sa structure euclidienne canonique.

On note $\det(x,y,z)$ le déterminant du système de vecteurs (x,y,z) dans la base $\mathcal{B}$.

Pour x, y dans E, $x \wedge y$ désigne le produit vectoriel de x et y.

Les propriétés suivantes du produit vectoriel nous seront utiles.

Lemme 7.12 *Pour tous vecteurs x, y, z, t dans $\mathbb{R}^3$, on a :*

$$\langle x \wedge y \mid z \wedge t\rangle = \begin{vmatrix} \langle x \mid z\rangle & \langle y \mid z\rangle \\ \langle x \mid t\rangle & \langle y \mid t\rangle \end{vmatrix}$$

$$\|x \wedge y\|^2 = \|x\|^2\|y\|^2 - \langle x \mid y\rangle^2$$

$$x \wedge (y \wedge z) = \langle x \mid z\rangle y - \langle x \mid y\rangle z \text{ (formule de Grassmann)}$$

Démonstration. Par 4-linéarité il suffit de vérifier la première formule sur les vecteurs de base canonique, ce qui ne pose pas de problème.

De cette formule, on déduit que :

$$\|x \wedge y\|^2 = \begin{vmatrix} \|x\|^2 & \langle y \mid x\rangle \\ \langle x \mid y\rangle & \|y\|^2 \end{vmatrix} = \|x\|^2 \|y\|^2 - \langle x \mid y\rangle^2$$

Par 3-linéarité il suffit de vérifier la formule de Grassmann sur les vecteurs de base canonique, ce qui ne pose pas de problème. ■

Lemme 7.13 *Soit x un vecteur non nul dans $\mathbb{R}^3$. Pour tout vecteur $y \in \mathbb{R}^3$ orthogonal à x il existe un vecteur $z \in \mathbb{R}^3$ tel $x \wedge z = y$.*

Démonstration. Si $\langle x \mid y\rangle = 0$, on a alors $x \wedge (y \wedge x) = \|x\|^2 y$ et en posant $z = \dfrac{1}{\|x\|^2} y \wedge x$, on a $x \wedge z = y$. ■

Théorème 7.12 *Si $f : \mathbb{R}^3 \to \mathbb{R}^3$ est une rotation, on a alors $f(x \wedge y) = f(x) \wedge f(y)$ pour tous vecteurs x, y dans $\mathbb{R}^3$.*

Démonstration. Soit f une rotation de $\mathbb{R}^3$.

Pour tous vecteurs x, y, z dans $\mathbb{R}^3$ on a :

$$\begin{aligned} \det(f(x), f(y), f(z)) &= \det(f)\det(x, y, z) = \det(x, y, z) \\ &= \langle x \wedge y \mid z\rangle = \langle f(x \wedge y) \mid f(z)\rangle \end{aligned}$$

et du fait que f est un isomorphisme, cela peut s'écrire $\det(f(x), f(y), w) = \langle f(x \wedge y) \mid w\rangle$ pour tout $w \in \mathbb{R}^3$, ce qui revient à dire que $f(x \wedge y) = f(x) \wedge f(y)$ par définition du produit vectoriel. ■

Pour ce qui suit, on se donne une application $f : \mathbb{R}^3 \to \mathbb{R}^3$ solution de (7.8).

Lemme 7.14 *On a $f(0) = 0$ et si f s'annule en un vecteur non nul x_0, elle est alors identiquement nulle.*

Démonstration. Avec $f(0) = f(0) \wedge f(0)$, on déduit que $f(0)$ est orthogonal à lui même, donc que $\|f(0)\|^2 = \langle f(0) \mid f(0)\rangle = 0$ et $f(0) = 0$.

En notant $D = \mathbb{R}x_0$ la droite vectorielle engendrée par x_0 et $H = D^\perp$ le plan orthogonal à D, on a $\mathbb{R}^3 = D \oplus H$ et tout vecteur $x \in \mathbb{R}^3$ s'écrit de manière unique $x = \lambda x_0 + h$ avec $\lambda \in \mathbb{R}$ et $h \in H$.

Si $\lambda = 0$, on a alors $x = h$ et x est orthogonal à x_0, il existe donc un vecteur $z \in E$ tel $x_0 \wedge z = x$ et $f(x) = f(x_0) \wedge f(z) = 0$.

On a donc $f(h) = 0$ pour tout $h \in H$.

En désignant par h_0 un vecteur non nul dans H, pour tout $\lambda \in \mathbb{R}$ on peut trouver $z \in \mathbb{R}^3$ tel $h_0 \wedge z = \lambda x_0$ et $f(\lambda x_0) = f(h_0) \wedge f(z) = 0$.

On a donc $f(x) = 0$ pour tout $x \in D$.

Enfin si $x = \lambda x_0 + h$ avec $\lambda \in \mathbb{R}^*$ et $h \in H - \{0\}$, le vecteur $z_0 = x_0 \wedge h$ est non nul (x_0 et h sont linéairement indépendants), $f(z_0) = 0$ et en écrivant $x = z_0 \wedge t$, on déduit que $f(x) = 0$.

En définitive on a $f(x) = 0$ pour tout $x \in \mathbb{R}^3$, c'est-à-dire que f est identiquement nulle. ■

Dans ce qui suit on suppose que f est une solution non identiquement nulle de l'équation fonctionnelle (7.8).

Lemme 7.15 *Les vecteurs $f(x)$ et $f(y)$ sont liés si, et seulement si, les vecteurs x et y le sont.*

Démonstration. On a :

$$\begin{aligned}(f(x), f(y) \text{ liés}) &\Leftrightarrow (f(x) \wedge f(y) = 0) \Leftrightarrow (f(x \wedge y) = 0) \\ &\Leftrightarrow (x \wedge y = 0) \Leftrightarrow (x, y \text{ liés})\end{aligned}$$

■

Lemme 7.16 *L'application f est linéaire.*

Démonstration. Soient x, y dans $\mathbb{R}^3$ et λ dans $\mathbb{R}$.

1. Pour $x = 0$ ou $y = 0$, il est clair que $f(x+y) = f(x) + f(y)$.
   On suppose donc que $x \neq 0$ et $y \neq 0$. Avec :

   $$\begin{aligned}f(x+y) \wedge f(x) &= f((x+y) \wedge x) = f(y \wedge x) = f(y) \wedge f(x) \\ &= (f(x) + f(y)) \wedge f(x)\end{aligned}$$

   on déduit que $(f(x+y) - f(x) - f(y)) \wedge f(x) = 0$, ce qui équivaut à dire que les vecteurs $f(x+y) - f(x) - f(y)$ et $f(x)$ sont liés.
   Pour $x \neq 0$ on a $f(x) \neq 0$ et il existe alors un réel α tel que $f(x+y) - f(x) - f(y) = \alpha f(x)$.
   Les vecteurs x et y jouant des rôles symétriques on obtient de même l'existence d'un réel β tel que $f(x+y) - f(x) - f(y) = \beta f(y)$.
   Si x et y sont linéairement indépendants il en est de même de $f(x)$ et $f(y)$ et l'égalité $\alpha f(x) - \beta f(y) = 0$ se traduit par $\alpha = \beta = 0$ et $f(x+y) = f(x) + f(y)$.
   Si les vecteurs x, y sont liés on a $y = \lambda x$ avec $\lambda \in \mathbb{R}^*$ et pour $z \in \mathbb{R}^3$ linéairement indépendant de x (et de y), le système $(x+y, z)$ est soit libre soit égal à $(0, z)$, ce qui donne dans tous les cas :

   $$f(x+y+z) = f(x+y) + f(z)$$

   D'autre part avec $(x+z, y)$ et (x, z) libres on a aussi :

   $$f(x+y+z) = f(x+z) + f(y) = f(x) + f(z) + f(y)$$

   ce qui permet de conclure à $f(x+y) = f(x) + f(y)$.
2. Pour $x = 0$ ou $\lambda = 0$, il est clair que $f(\lambda x) = \lambda f(x)$.
   Pour $x \neq 0$ fixé les vecteurs x et λx étant liés il en est de même des vecteurs $f(x)$ et $f(\lambda x)$ avec $f(x) \neq 0$, il existe donc un réel $\mu_x(\lambda)$ tel que $f(\lambda x) = \mu_x(\lambda) f(x)$.
   Il est clair que $\mu_x(0) = 0$ et $\mu_x(1) = 1$.
   Pour λ_1, λ_2 dans $\mathbb{R}$ on a :

   $$\begin{aligned}f((\lambda_1 + \lambda_2) x) &= \mu_x(\lambda_1 + \lambda_2) f(x) \\ &= f(\lambda_1 x + \lambda_2 x) = f(\lambda_1 x) + f(\lambda_2 x) \\ &= (\mu_x(\lambda_1) + \mu_x(\lambda_2)) f(x)\end{aligned}$$

   ce qui entraîne $\mu_x(\lambda_1 + \lambda_2) = \mu_x(\lambda_1) + \mu_x(\lambda_2)$.
   Si y est un vecteur linéairement indépendant de x, avec :

   $$\begin{aligned}f(\lambda x \wedge y) &= f(\lambda x) \wedge f(y) = \mu_x(\lambda) f(x) \wedge f(y) \\ &= f(x \wedge \lambda y) = f(x) \wedge f(\lambda y) = \mu_y(\lambda) f(x) \wedge f(y)\end{aligned}$$

et avec $f(x) \wedge f(y) \neq 0$ (puisque $f(x)$ et $f(y)$ sont libres comme x et y), on déduit que $\mu_x(\lambda) = \mu_y(\lambda)$ pour tout réel λ. En posant $z = x \wedge y$, le système (x, y, z) est libre, donc $\mu_x = \mu_y = \mu_z$ et pour λ_1, λ_2 dans $\mathbb{R}$ on a :

$$\begin{aligned} f(\lambda_1\lambda_2 z) &= \mu_z(\lambda_1\lambda_2) f(z) = \mu_x(\lambda_1\lambda_2) f(z) \\ &= f(\lambda_1 x \wedge \lambda_2 y) = f(\lambda_1 x) \wedge f(\lambda_2 y) \\ &= \mu_x(\lambda_1)\mu_x(\lambda_2) f(x) \wedge f(y) = \mu_x(\lambda_1)\mu_x(\lambda_2) f(z) \end{aligned}$$

ce qui entraîne $\mu_x(\lambda_1\lambda_2) = \mu_x(\lambda_1)\mu_x(\lambda_2)$.
En définitive μ_x est un morphisme de corps de $\mathbb{R}$ sur lui même, c'est donc l'identité.
On a donc $f(\lambda x) = \lambda x$ pour tout $x \in E$.
L'application f est donc linéaire.

■

Lemme 7.17 *L'application f conserve l'orthogonalité et pour tout $x \in \mathbb{R}^3$ tel que $\|x\| = 1$, on a $\|f(x)\| = 1$.*

Démonstration. Soient x, y orthogonaux dans $\mathbb{R}^3$.
Si $x = 0$, on a alors $f(x) = 0$ qui est orthogonal à $f(y)$.
Si $x \neq 0$, comme y est orthogonal à x, on peut trouver $z \in \mathbb{R}^3$ tel que $y = x \wedge z$ et $f(y) = f(x \wedge z) = f(x) \wedge f(z)$ est orthogonal à $f(x)$.
Soit $x \in \mathbb{R}^3$ tel que $\|x\| = 1$.
On complète x en une base orthonormée directe (x, y, z) et avec $x = y \wedge z$, on a :

$$\begin{aligned} \|f(x)\|^2 = \|f(y) \wedge f(z)\|^2 &= \|f(y)\|^2 \|f(z)\|^2 - \langle f(y) \mid f(z)\rangle^2 \\ &= \|f(y)\|^2 \|f(z)\|^2 \end{aligned}$$

soit $\|f(x)\| = \|f(y)\| \|f(z)\|$.
De manière analogue, on montre que $\|f(y)\| = \|f(x)\| \|f(z)\|$ et $\|f(z)\| = \|f(x)\| \|f(y)\|$, ce qui donne :

$$\|f(y)\| = \|f(x)\|^2 \|f(y)\|$$

avec $f(y) \neq 0$, ce qui entraîne $\|f(x)\| = 1$. ■

Théorème 7.13 *Les seules solutions non identiquement nulles de l'équation fonctionnelle* (7.8) *sont les rotations.*

Démonstration. Soit f une solution non identiquement nulle de l'équation fonctionnelle (7.8).
Cette application est linéaire et transforme la base canonique en une base orthonormée (les lemmes qui précèdent), donc c'est un endomorphisme orthogonal.
Avec :

$$\begin{aligned} \det(f) &= \det(f(e_1), f(e_2), f(e_3)) = \det(f(e_1), f(e_2), f(e_1 \wedge e_2)) \\ &= \det(f(e_1), f(e_2), f(e_1) \wedge f(e_2)) = \|f(e_1) \wedge f(e_2)\|^2 > 0 \end{aligned}$$

on déduit que c'est une rotation. ■

8

Algorithmes de résolution approchée d'une équation numérique

8.1 Position des problèmes

I est un intervalle réel d'extrémités $a < b$ dans $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ et $\overset{\circ}{I} =]a, b[$ est l'intérieur de I.

On désigne par $\mathcal{C}^0(I)$ l'espace vectoriel réel des fonctions définies, continues sur I et à valeurs réelles et pour tout entier p strictement positif, on note $\mathcal{C}^p(I)$ le sous-espace vectoriel de $\mathcal{C}^0(I)$ constitué des fonctions p fois continûment dérivables sur I.

Enfin, on note $\mathcal{C}^\infty(I)$ le sous-espace vectoriel de $\mathcal{C}^0(I)$ constitué des fonctions indéfiniment dérivables sur I.

On s'intéresse ici à l'approximation des solutions, quand elles existent, d'une équation numérique $f(x) = 0$, où f est une fonction à valeurs réelles définie sur un intervalle d'intérieur non vide.

Dans le cas général on s'intéresse aux solutions réelles localisées dans un intervalle I. Dans le cas particulier des fonctions polynomiales on peut chercher toutes les solutions réelles ou complexes.

Trois problèmes essentiels vont se poser.

- Tout d'abord le problème qualitatif de l'existence et de l'unicité d'une solution de l'équation $f(x) = 0$ dans un intervalle donné contenu dans I.
- Ensuite vient le problème de l'approximation numérique d'une telle solution. On cherchera à construire une suite $(x_n)_{n\in\mathbb{N}}$ de points de I qui converge vers la solution en question.
- Pour juger de la performance d'une méthode il est important d'avoir une majoration aussi fine que possible de l'erreur d'approximation.

Les outils d'analyse réelle utilisés pour cette étude sont les suivants :

- propriétés des suites réelles monotones ;
- suites adjacentes ;
- suites bornées ;
- critère de Cauchy sur les suites réelles ou complexes ;
- le théorème des valeurs intermédiaires pour les fonctions continues d'une variable réelle ;
- le théorème et l'inégalité des accroissements finis ;
- la formule de Taylor-Lagrange.

8.2 La méthode de dichotomie

Cette méthode est basée sur le théorème des valeurs intermédiaires qui suit.

Théorème 8.1 *Si $f \in \mathcal{C}^0([a,b])$ est telle que $f(a)f(b) < 0$ alors l'équation $f(x) = 0$ admet au moins une solution $\alpha \in \,]a,b[$. Si de plus f est strictement monotone alors cette solution est unique dans l'intervalle $]a,b[$.*

Démonstration. On peut le déduire du théorème des segments emboîtés (voir le volume sur les suites et les séries numériques) ou du théorème de la borne supérieure (théorème 2.1). ■

Remarque 8.1 *On rappelle qu'une fonction $f \in \mathcal{C}^0([a,b])$ est injective si et seulement si elle est strictement monotone (théorème 2.11).*

Remarque 8.2 *Si $I = \mathbb{R}$ et f a une limite ℓ_1 en $-\infty$ et une limite ℓ_2 en $+\infty$ telles que $\ell_1\ell_2 < 0$, on déduit alors du théorème des valeurs intermédiaires que l'équation $f(x) = 0$ a au moins une solution sur $\mathbb{R}$.*

Exercice 8.1 *Montrer qu'un polynôme non nul $P \in \mathbb{R}[X]$ de degré impair a au moins une racine réelle.*

Solution 8.1 *En supposant que le terme de plus degré de P soit positif, dans le cas où* $\deg(P)$ *est impair, on a $\lim\limits_{x\to-\infty} P(x) = -\infty$ et $\lim\limits_{x\to+\infty} P(x) = +\infty$. Le théorème des valeurs intermédiaires que l'équation $P(x) = 0$ a au moins une solution sur $\mathbb{R}$.*

On suppose pour ce paragraphe que $f \in \mathcal{C}^0([a,b])$ est strictement monotone avec $f(a)f(b) < 0$. Le théorème des valeurs intermédiaires nous permet alors de construire deux suites adjacentes qui convergent vers la solution $\alpha \in \,]a,b[$ de l'équation $f(x) = 0$.

L'idée de la méthode de dichotomie est de découper l'intervalle en deux parties égales et de conserver celle qui contient la solution de $f(x) = 0$. On détermine cette partie en étudiant le signe de $f(u)f(v)$ où $[u,v] \subset [a,b]$. On répète ensuite le processus.

On construit donc une suite $([a_n,b_n])_{n\in\mathbb{N}}$ d'intervalles emboîtés dans $[a,b]$ de la manière suivante.

$[a_0,b_0] = [a,b]$.

En supposant construit un intervalle $[a_n,b_n] \subset [a,b]$ de longueur $\dfrac{b-a}{2^n}$ qui contient la racine α de $f(x) = 0$, on note $c_n = \dfrac{a_n+b_n}{2}$ (milieu de $[a_n,b_n]$) et on fait le test suivant :

si $f(a_n)f(c_n) \leq 0$
alors $[a_{n+1},b_{n+1}] = [a_n,c_n]$;
sinon $[a_{n+1},b_{n+1}] = [c_n,b_n]$.

Dans tous les cas on a $\alpha \in [a_{n+1},b_{n+1}]$, $b_{n+1} - a_{n+1} = \dfrac{b_n-a_n}{2} = \dfrac{b-a}{2^{n+1}}$.

On donc pour tout $n \in \mathbb{N}$:

$$\begin{aligned} &a_n \leq a_{n+1} \leq \alpha \leq b_{n+1} \leq b_n \\ &b_n - a_n = \frac{b-a}{2^n} \underset{k\to+\infty}{\longrightarrow} 0 \end{aligned}$$

C'est-à-dire que les suites $(a_n)_{k\in\mathbb{N}}$ et $(b_n)_{k\in\mathbb{N}}$ sont adjacentes et convergent vers la solution α de $f(x) = 0$.

Remarque 8.3 *La suite $(c_n)_{k\in\mathbb{N}}$ converge également vers α.*

Pour chacune des trois suites, on la majoration de l'erreur d'approximation :

$$\forall n \in \mathbb{N},\ |\alpha - x_n| \leq \frac{b-a}{2^n}$$

ce qui permet de déterminer un nombre suffisant d'itérations pour atteindre une précision $\varepsilon > 0$ donnée. On peut prendre $n_0 = \log_2 \left(\frac{b-a}{\varepsilon}\right)$.

En pratique on préfère utiliser le test d'arrêt $|b_n - a_n| < \varepsilon$.

Avec les hypothèses ci-dessus la méthode converge toujours, mais cette convergence peut être lente. De plus cette méthode ne se généralise pas au cas des systèmes non linéaires d'équations.

On peut écrire la programmation structurée qui suit.

Entrée f : Fonction ; a, b, ε : Réel ; Sortie α : Réel ;
Début $y = f(a)$;
  Répéter
    $c = \dfrac{a+b}{2}$;
    $z = f(c)$;
    Si $y \cdot z \leq 0$ *Alors*
      $b = c$;
    Sinon $a = c$;
      $y = z$;
    Fin Si ;
  Jusqu'à $(|b-a| < \epsilon)$ *ou* $(|z| < \epsilon)$;
  $\alpha = c$;
Fin ;

Remarque 8.4 *Pour calculer toutes les racines réelles de l'équation $f(x) = 0$, sur un intervalle $[a, b]$, on peut le découper en un nombre suffisamment grand de sous intervalles $[u_i, v_i]$ et appliquer la procédure précédente sur chacun de ces intervalles.*

Exercice 8.2 *Montrer que l'équation $e^x + x = 2$ a une unique solution sur $\mathbb{R}$. Localiser cette solution α avec un outil graphique, puis donner une approximation de α à 10^{-1} près.*

Solution 8.2 *La fonction $f : x \mapsto e^x + x - 2$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ avec $f'(x) = e^x + 1 > 0$ pour tout réel x, elle est donc strictement croissante. De plus avec $\lim\limits_{x\to-\infty} f((x) = -\infty$ et $\lim\limits_{x\to+\infty} f((x) = +\infty$, on déduit qu'il existe un unique réel α tel que $f(\alpha) = 0$.*
Graphiquement, on voit que $\alpha \in \left]0, 1\right[$ (figure 8.1).
Avec $f(0) = -1$, $f(1) = 0.718\,28$, $f\left(\frac{1}{2}\right) = 0.148\,72$, on a $\alpha \in \left]0, \frac{0.5}{1}\right[$;
avec $f\left(\frac{1}{4}\right) = -0.465\,97$, on a $\alpha \in \left]\frac{1}{4}, \frac{1}{2}\right[$;
avec $f\left(\frac{3}{8}\right) = -0.170\,01$, on a $\alpha \in \left]\frac{3}{8}, \frac{1}{2}\right[$;
avec $f\left(\frac{7}{16}\right) = -1.367\,0 \times 10^{-2}$, on a $\alpha \in \left]\frac{7}{16}, \frac{1}{2}\right[$.
En définitive, $c = \frac{15}{32} \approx 0.468\,75$ est une approximation de α avec $|c - \alpha| \leq \frac{1}{16} < 10^{-1}$.
Le logiciel Maple nous donne $\alpha \approx 0.44285440100238858314$.

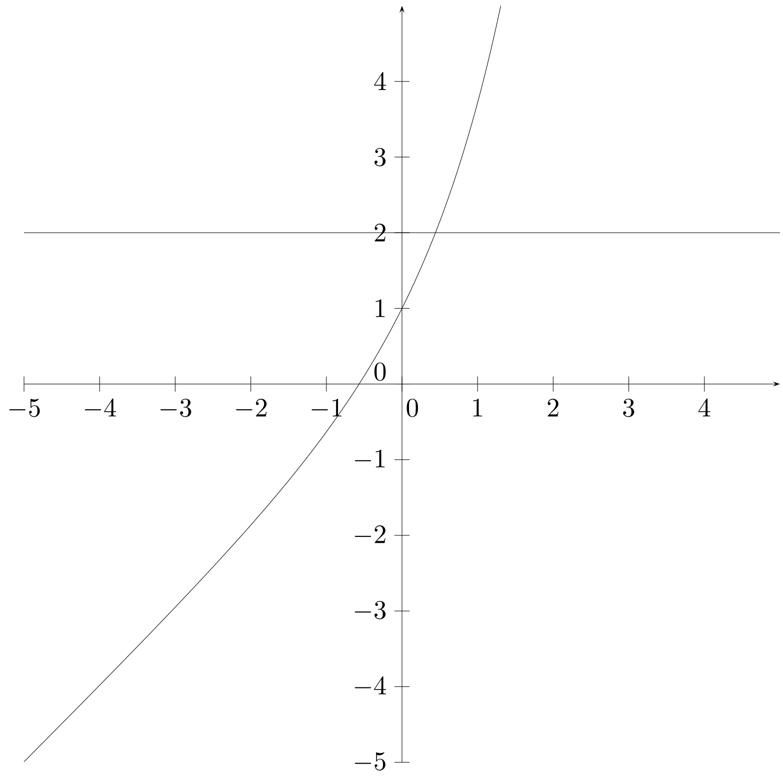

FIGURE 8.1 –

Exercice 8.3 *Soit* $f : x \mapsto 2\cos(1 + x^2) - x$.

1. *Montrer que l'équation* $f(x) = 1$ *a une unique solution* α *dans* $I = [0, 1]$.
2. *Donner une approximation de* α *à* 10^{-1} *près.*

Solution 8.3

1. *Graphiquement, on voit que l'équation* $2\cos(1 + x^2) = x + 1$ *a une solution* $\alpha \in]0, 1[$ *(figure 8.2).La fonction* $g : x \mapsto 2\cos(1 + x^2) - x - 1$ *est de classe* $\mathcal{C}^1$ *sur* I *avec :*

$$\forall x \in I,\ g'(x) = -4x \sin\left(1 + x^2\right) - 1 < 0$$

(pour $0 \leq x \leq 1$, *on a* $0 < 1 + x^2 \leq 2 < \pi$ *et* $\sin(1 + x^2) > 0$*), elle est donc strictement croissante.*
Comme $g(0) = 2\cos(1) - 1 \approx 8.060\,5 \times 10^{-2} < 0$ *(on peut aussi dire que* $1 > \dfrac{\pi}{3}$ *et* $\cos(1) < \cos\left(\dfrac{\pi}{3}\right) = \dfrac{1}{2}$ *puisque* $\cos$ *est strictement décroissante sur* $[0, \pi]$*) et* $g(1) = 2(\cos(2) - 1) < 0$, *le théorème des valeurs intermédiaires nous dit qu'il existe un unique réel* $\alpha \in]0, 1[$ *tel que* $g(\alpha) = 0$.
2. *On a* $\alpha \approx 0.07189199978$.

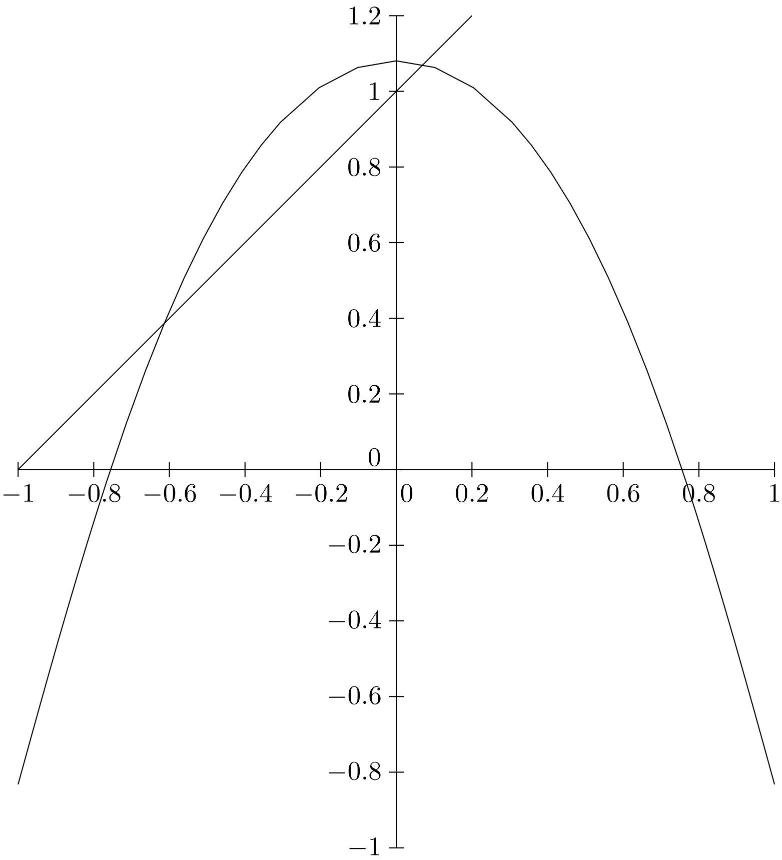

FIGURE 8.2 – $2\cos(1+x^2) = x+1$

8.3 La méthode des approximations successives

La résolution d'une équation $g(x) = 0$ sur I peut se ramener à un problème de recherche de point fixe en considérant la fonction f définie sur I par :

$$\forall x \in I,\ f(x) = x - \lambda(x)\, g(x)$$

où la fonction λ définie sur I est telle que $f(I) \subset I$ et $\lambda(x) \neq 0$ pour tout $x \in I$ (l'intervalle I devra éventuellement être remplacé par un intervalle plus petit).

On peut prendre pour λ une fonction constante non nulle (exercice 8.8), $\lambda(x) = \dfrac{1}{g'(x)}$ quand cela est possible (méthode de Newton), ou $\lambda(x) = \dfrac{b-x}{g(b)-g(x)}$ avec $b \in I$ quand cela est possible (méthode de la sécante).

Pour ce paragraphe I désigne un intervalle fermé de $\mathbb{R}$ d'intérieur non vide (non nécessairement borné) et f une fonction définie sur I à valeurs réelles telle que $f(I) \subset I$. On dit que l'intervalle I est stable par f.

Définition 8.1 *Si x est un réel dans I, on appelle orbite de x suivant f la suite $(x_n)_{n\in\mathbb{N}}$ de points de I définie par $x_0 = x$ et $x_{n+1} = f(x_n)$ pour tout $n \in \mathbb{N}$.*

Remarque 8.5 *C'est la stabilité de l'intervalle I qui permet de définir les orbites suivant f. Par exemple pour $f : x \mapsto \ln(x)$ et $I = \left[\frac{1}{2}, 1\right]$, on a $\ln(x) \leq 0$ pour tout $x \in I$ et il n'est pas possible de définir des orbites suivant f.*

Lemme 8.1 *Si, avec les notations précédentes, la suite $(x_n)_{n\in\mathbb{N}}$ converge alors sa limite α est dans I et si de plus la fonction f est continue sur I, alors $f(\alpha) = \alpha$.*

Démonstration. Si $\lim\limits_{n\to+\infty} x_n = \alpha$ alors $\alpha \in I$ puisque cet ensemble est fermé. Si de plus f est continue sur I, alors :

$$\alpha = \lim_{n\to+\infty} x_{n+1} = \lim_{n\to+\infty} f(x_n) = f(\alpha)$$

■

Définition 8.2 *On dit que $\alpha \in I$ est un point fixe de f si $f(\alpha) = \alpha$.*

Remarque 8.6 *La seule hypothèse de continuité de f n'assure pas la convergence de la suite $(x_n)_{n\in\mathbb{N}}$. Par exemple pour $f(x) = -x$ sur $[-1, 1]$ on a $x_n = (-1)^n x_0$ pour tout n et cette suite est divergente pour $x_0 \neq 0$.*

Remarque 8.7 *La fonction caractéristique de $\mathbb{Q}$ fournit un exemple de fonction discontinue en tout point de $\mathbb{R}$ pour laquelle la suite $(x_n)_{n\in\mathbb{N}}$ converge vers l'unique point fixe de f. En effet pour $x_0 \in \mathbb{Q}$, on a $x_n = 1$ pour tout $n \geq 1$ et pour $x_0 \notin \mathbb{Q}$ on a $x_1 = 0$ et $x_n = 1$ pour tout $n \geq 2$, donc dans tous les cas $\lim\limits_{n\to+\infty} x_n = 1$ est point fixe de f.*

Si l'orbite suivant f d'un point $x \in I$ est convergente vers un point fixe α de f, on dit alors que cette suite est une suite d'approximations successives de α de premier terme x.

Exercice 8.4 *Soit f une fonction continue de I dans I admettant un unique point fixe α dans I telle que :*

$$\forall x \in I \setminus \{\alpha\}, \ |f(x) - \alpha| < |x - \alpha|$$

Montrer que toute orbite suivant f converge vers α.

Solution 8.4 *Une orbite $(x_n)_{n\in\mathbb{N}}$ est bien définie du fait que I est stable par f et pour tout $n \in \mathbb{N}$, en notant $r_n = |x_n - \alpha|$, on a :*

$$0 \leq r_{n+1} = |f(x_n) - \alpha| \leq |x_n - \alpha| = r_n$$

(l'égalité est réalisée si $x_n = \alpha$), c'est-à-dire que la suite $(r_n)_{n\in\mathbb{N}}$ est décroissante minorée, elle est donc convergente.

Avec :

$$\forall n \in \mathbb{N}, \ |x_n| \leq r_n + |\alpha| \leq r_0 + |\alpha|$$

on déduit que la suite $(x_n)_{n\in\mathbb{N}}$ est bornée, on peut donc en extraire une sous-suite $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ convergente vers $\beta \in I$ (l'intervalle I est fermé).

Avec la continuité de f et $\lim\limits_{n\to+\infty} r_{\varphi(n)+1} = \lim\limits_{n\to+\infty} r_{\varphi(n)}$ on déduit que :

$$\begin{aligned} |f(\beta) - \alpha| &= \lim_{n\to+\infty} \left|f\left(x_{\varphi(n)}\right) - \alpha\right| = \lim_{n\to+\infty} \left|x_{\varphi(n)+1} - \alpha\right| \\ &= \lim_{n\to+\infty} r_{\varphi(n)+1} = \lim_{n\to+\infty} r_{\varphi(n)} = \lim_{n\to+\infty} \left|x_{\varphi(n)} - \alpha\right| = |\beta - \alpha| \end{aligned}$$

ce qui implique $\beta = \alpha$ puisque $\beta \neq \alpha$ donnerait $|f(\beta) - \alpha| < |\beta - \alpha|$.

En définitive, la suite $(x_n)_{n\in\mathbb{N}}$ est bornée avec α pour unique valeur d'adhérence, elle est donc convergente vers α (voir le volume sur les suite et les séries numériques).

8.3.1 Cas des fonctions monotones sur un segment

Dans le cas où la fonction f est monotone sur un intervalle compact, les problèmes de l'existence de points fixes et de la convergence des orbites peut s'étudier de manière relativement simple.

On a tout d'abord le résultat suivant, conséquence du théorème de la borne supérieure, qui nous assure l'existence d'un point fixe pour f croissante.

Théorème 8.2 *Si f est une fonction croissante de $I = [a, b]$ dans I, elle admet alors au moins un point fixe $\alpha \in I$.*

Démonstration. L'ensemble :

$$E = \{x \in [a, b] \mid f(x) \geq x\}$$

est non vide ($f(a) \in [a, b]$ entraîne $f(a) \geq a$) majoré par b, il admet donc une borne supérieure $\alpha \in [a, b]$.

Si $\alpha = a$ alors $f(x) < x$ pour tout $x \in I \setminus \{a\}$ et pour $n_0 \in \mathbb{N}^*$ tel que $a + \dfrac{1}{n_0} \leq b$ on a, du fait de la croissance de f sur I :

$$\forall n \geq n_0,\ a \leq f(a) \leq f\left(a + \frac{1}{n}\right) < a + \frac{1}{n}$$

qui par passage à la limite quand n tend vers l'infini donne $a = f(a)$.

Si $\alpha > a$, par définition de la borne supérieure, on peut trouver pour tout entier naturel non nul n un réel $x_n \in \left]\alpha - \dfrac{1}{n}, \alpha\right] \cap E$ et :

$$\forall n \geq 1,\ f(\alpha) \geq f(x_n) \geq x_n > \alpha - \frac{1}{n}$$

qui par passage à la limite quand n tend vers l'infini donne $f(\alpha) \geq \alpha$, c'est-à-dire que α est dans E.

Si $\alpha = b$ alors $\alpha \leq f(\alpha) \leq b = \alpha$ et $\alpha = f(\alpha)$.

Si $\alpha < b$, pour n assez grand on a $\alpha \leq f(\alpha) \leq f\left(\alpha + \dfrac{1}{n}\right) < \alpha + \dfrac{1}{n}$ qui par passage à la limite donne $\alpha = f(\alpha)$.

En définitive α est un point fixe de f dans I. ■

Remarque 8.8 *Le résultat précédent n'est pas valable pour f décroissante comme le montre l'exemple de la fonction f définie par $f(x) = 1$ si $x \in \left[0, \dfrac{1}{2}\right]$ et $f(x) = 0$ si $x \in \left]\dfrac{1}{2}, 1\right]$.*
Il n'est pas valable non plus sur un intervalle fermé non borné comme le montre l'exemple de $f(x) = x + 1$ sur $[0, +\infty[$.

En ce qui concerne la convergence des orbites on a le résultat suivant où $(x_n)_{n \in \mathbb{N}}$ est une orbite suivant f.

Théorème 8.3 *Si la fonction f est croissante alors la suite $(x_n)_{n \in \mathbb{N}}$ est croissante si $f(x_0) \geq x_0$ et décroissante si $f(x_0) \leq x_0$. Si l'intervalle I est borné alors cette suite est convergente et si de plus f est continue alors sa limite est un point fixe de f.*

Démonstration. Si $x_1 = f(x_0) \geq x_0$, une récurrence élémentaire permet de déduire de la croissance de f que la suite $(x_n)_{n\in\mathbb{N}}$ est croissante.

Si $I = [a,b]$ alors cette suite est croissante et majorée par b, elle converge donc vers un réel $\alpha \in [a,b]$.

Pour f continue sur I, de $x_{n+1} = f(x_n)$ pour tout $n \in \mathbb{N}$, on déduit que α est un point fixe de f dans I. ■

Pour f non monotone, la monotonie de la suite $(x_n)_{n\in\mathbb{N}}$, peut se déduire l'étude du signe de la fonction $x \mapsto f(x) - x$. Dans le cas où tous les x_n sont strictement positifs, on peut aussi s'intéresser à la position de $\dfrac{x_{n+1}}{x_n}$ par rapport à 1, en étudiant le signe de la fonction $x \mapsto \dfrac{f(x)}{x} - 1$.

Remarque 8.9 *Dans le théorème précédent sans l'hypothèse de continuité sur le compact $[a,b]$ la suite $(x_n)_{n\in\mathbb{N}}$ ne converge pas nécessairement vers un point fixe de f.*
Si par exemple f est définie sur $[0,1]$ par $f(x) = \dfrac{x}{2} + \dfrac{1}{4}$ sur $\left[0, \dfrac{1}{2}\right[$, $f(x) = 1$ sur $\left[\dfrac{1}{2}, 1\right]$, l'orbite de $x_0 = 0$ est définie par $x_n = \dfrac{1}{2}\left(1 - \dfrac{1}{2^n}\right)$ et cette suite converge vers $\dfrac{1}{2}$ qui n'est pas point fixe de f.

Dans le cas où f est décroissante on a, toujours avec les mêmes notations, le résultat suivant.

Théorème 8.4 *Si la fonction f est décroissante elle admet alors au plus un point fixe dans I. Si de plus l'intervalle I est borné et la fonction f continue, alors les suites extraites $(x_{2n})_{n\in\mathbb{N}}$ et $(x_{2n+1})_{n\in\mathbb{N}}$ convergent vers des points fixes de $f \circ f$.*
Dans le cas où $f \circ f$ admet un unique point fixe, la suite $(x_n)_{n\in\mathbb{N}}$ converge et sa limite est ce point fixe qui est également l'unique point fixe de f.

Démonstration. Si $\alpha \leq \beta$ sont deux points fixes de f dans I, on a alors pour f décroissante :

$$\alpha = f(\alpha) \geq f(\beta) = \beta$$

et $\alpha = \beta$. La fonction f admet donc au plus un point fixe dans I.

La fonction f étant décroissante il en résulte que $f \circ f$ est croissante et les orbites de g, $(x_{2n})_{n\in\mathbb{N}}$, $(x_{2n+1})_{n\in\mathbb{N}}$ sont monotones. Si I est borné, ces suites convergent alors vers des réels α et β dans I qui sont des points fixes de $f \circ f$ si f est continue.

Dans l'hypothèse où $f \circ f$ a un seul point fixe on a $\alpha = \beta$ et la suite $(x_n)_{n\in\mathbb{N}}$ converge vers ce point fixe qui est également l'unique point fixe de f. ■

Remarque 8.10 *Dans le cas où, avec les hypothèses précédentes, la fonction $f \circ f$ admet plusieurs points fixes la suite $(x_n)_{n\in\mathbb{N}}$ peut être divergente.*

Exercice 8.5 *Soit f la fonction définie sur $I = [0,1]$ par $f(x) = 1 - x^2$.*

1. *Montrer que l'intervalle I est stable par f et que f a un unique point fixe $\alpha \in I$.*
2. *Vérifier que $f \circ f$ a trois points fixes $\gamma < \alpha < \delta$ dans I.*
3. *Soit $(x_n)_{n\in\mathbb{N}}$ une orbite suivant f avec $x_0 \neq \alpha$.*
   - (a) *Montrer que $x_n \neq \alpha$ pour tout entier $n \geq 0$ et que la suite $(x_n)_{n\in\mathbb{N}}$ diverge.*
   - (b) *On suppose que $x_0 < \alpha$. Montrer que $\lim\limits_{n\to+\infty} x_{2n} = \gamma$ et $\lim\limits_{n\to+\infty} x_{2n+1} = \delta$.*

Solution 8.5

1. *L'équation* $f(x) = x$ *équivaut à* $x^2 + x - 1 = 0$. *Cette dernière équation a deux racines réelles* $\alpha = \dfrac{\sqrt{5}-1}{2}$ *et* $\beta = -\dfrac{\sqrt{5}+1}{2}$, *seule* α *est dans* I.
La fonction f *est strictement décroissante sur* I *(*$f'(x) = -2x < 0$ *pour* $0 < x < 1$*) avec* $f(0) = 1 \in I$ *et* $f(1) = 0 \in I$, *donc* $f(I) = [0,1] = I$.
2. *On a :*
$$g(x) - x = f \circ f(x) - x = 1 - \left(1 - x^2\right)^2 = 2x^2 - x^4 - x$$
Comme f *et* g *sont polynomiales et les points fixes de* f *sont aussi points fixes de* g, *il existe des réels* a, b, c *tels que :*
$$g(x) - x = (f(x) - x)\left(ax^2 + bx + c\right)$$
soit :
$$-x^4 + 2x^2 - x = \left(-x^2 - x + 1\right)\left(ax^2 + bx + c\right)$$
On trouve :
$$g(x) - x = (f(x) - x)\left(x^2 - x\right)$$
et trois points fixes pour g *dans* I *:*
$$\gamma = 0 < \alpha = \frac{\sqrt{5}-1}{2} < \delta = 1$$
3. *Si* $x_0 = \alpha$, *la suite* $(x_n)_{n\in\mathbb{N}}$ *est alors stationnaire sur* α.
   (a) *On a, pour tout entier* $n \geq 0$ *:*
   $$x_{n+1} - \alpha = f(x_n) - f(\alpha) = \left(\alpha^2 - x_n^2\right) = (\alpha - x_n)(\alpha + x_n) \tag{8.1}$$
   Pour $n = 0$, *on a* $x_0 \neq \alpha$ *et supposant que* $x_n \neq \alpha$ *pour* $n \geq 0$, *on a :*
   $$|x_{n+1} - \alpha| = |\alpha - x_n|(\alpha + x_n) \geq \alpha|\alpha - x_n| > 0$$
   On a donc $x_n \neq \alpha$ *pour tout entier* $n \geq 0$.
   Si la suite $(x_n)_{n\in\mathbb{N}}$ *converge, c'est nécessairement vers* α *(l'unique point fixe de* f *dans* I*), mais de* (8.1), *on déduit que :*
   $$\lim_{n\to+\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|} = 2\alpha = \sqrt{5} - 1 > 1$$
   ce qui implique que $\lim\limits_{n\to+\infty} |x_n - \alpha| = +\infty$ *(critère de d'Alembert) et est incompatible* $\lim\limits_{n\to+\infty} x_n = \alpha$.
   Cette suite est donc divergente.
   (b) *Supposons que* $0 \leq x_0 < \alpha$.
   Comme $f \circ f$ *est strictement croissante sur* I *et* α *est point fixe de* $f \circ f$, *on a* $0 \leq x_{2n} < \alpha$ *pour tout entier* n. *La limite de la suite* $(x_{2n})_{n\in\mathbb{N}}$ *est donc* γ *ou* α *(on sait que cette suite converge vers un point fixe de* $f \circ f$ *et elle est à valeurs dans le fermé* $[0, \alpha]$*). Si* $\lim\limits_{n\to+\infty} x_{2n} = \alpha$, *on a alors* $\lim\limits_{n\to+\infty} x_{2n+1} = \lim\limits_{n\to+\infty} f(x_{2n}) = f(\alpha) = \alpha$ *et* $\lim\limits_{n\to+\infty} x_n = \alpha$, *ce qui contredit la divergence de* $(x_n)_{n\in\mathbb{N}}$. *On a donc* $\lim\limits_{n\to+\infty} x_{2n} = \gamma = 0$ *et* $\lim\limits_{n\to+\infty} x_{2n+1} = f(\gamma) = \delta = 1$.
   Pour $\alpha < x_0 \leq 2$, *on vérifie que* $\lim\limits_{n\to+\infty} x_{2n} = \delta$ *et* $\lim\limits_{n\to+\infty} x_{2n+1} = f(\delta) = \gamma$.

8.3.2 Le théorème du point fixe

I est toujours un intervalle réel fermé.

Pour f continue sur un segment, l'existence d'un point fixe est assurée par le résultat suivant, conséquence immédiate du théorème des valeurs intermédiaires.

Lemme 8.2 *Une fonction continue de $[a,b]$ dans $[a,b]$ admet au moins un point fixe.*

Démonstration. La fonction g définie sur $[a,b]$ par $g(x) = f(x) - x$ est continue avec $g(a) = f(a) - a \geq 0$ et $g(b) = f(b) - b \leq 0$ dans le cas où $f([a,b]) \subset [a,b]$.

Le théorème des valeurs intermédiaires nous dit alors qu'il existe un point α dans $[a,b]$ tel que $g(\alpha) = 0$. ∎

Remarque 8.11 *L'exemple de $f(x) = x+1$ sur $[0,+\infty[$ nous montre que ce résultat n'est pas vrai pour I fermé non compact.*

Remarque 8.12 *L'exemple de $f(x) = 1 - x^2$ sur $[0,1]$, nous montre qu'une suite d'approximations successives ne converge pas nécessairement vers un point fixe de f, même si ce dernier est unique sur I.*

Une condition nécessaire et suffisante de convergence d'une orbite suivant une fonction continue sur un segment est donnée par le théorème 8.5. La démonstration de ce théorème utilise le résultat suivant.

Lemme 8.3 *Si $(x_n)_{n\in\mathbb{N}}$ est une suite de réels telle que $\lim\limits_{n\to+\infty} (x_{n+1} - x_n) = 0$ alors l'ensemble de ses valeurs d'adhérence est soit vide soit un intervalle.*

Démonstration. Notons $\mathcal{A}$ l'ensemble des valeurs d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$, c'est-à-dire l'ensemble des limites de suites extraites de $(x_n)_{n\in\mathbb{N}}$. On suppose que cet ensemble est non vide, non réduit à un point et on se donne $a < b$ dans $\mathcal{A}$. Il s'agit alors de montrer que $]a,b[$ est inclus dans $\mathcal{A}$.

Si $c \in\,]a,b[$ n'est pas dans $\mathcal{A}$, il existe alors un réel $\varepsilon > 0$ tel que $]c-\varepsilon, c+\varepsilon[$ est contenu dans $]a,b[$ et un entier n_0 tel que pour tout $n \geq n_0$, $x_n \notin\,]c-\varepsilon, c+\varepsilon[$. Avec $\lim\limits_{n\to+\infty} (x_{n+1} - x_n) = 0$ on déduit qu'il existe un entier $n_1 \geq n_0$ tel que $|x_{n+1} - x_n| < 2\varepsilon$ pour tout $n \geq n_1$. En particulier on a $x_{n_1} \notin\,]c-\varepsilon, c+\varepsilon[$. Si $x_{n_1} \leq c-\varepsilon$ alors $x_{n_1+1} < 2\varepsilon + x_{n_1} \leq c+\varepsilon$ et comme $x_{n_1+1} \notin\,]c-\varepsilon, c+\varepsilon[$ cela signifie que $x_{n_1+1} \leq c-\varepsilon$. Par récurrence on déduit alors que $x_n \leq c-\varepsilon$ pour tout $n \geq n_1$, ce qui contredit le fait que b est valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$. De manière analogue $x_{n_1} \geq c+\varepsilon$ va contredire le fait que a est valeur d'adhérence de la suite $(x_n)_{n\in\mathbb{N}}$. On aboutit donc ainsi à une contradiction. ∎

De manière générale il est facile de vérifier que l'ensemble des valeurs d'adhérence d'une suite réelle est un fermé. On peut donc préciser dans la démonstration du lemme précédent que $\mathcal{A}$ est un intervalle fermé.

Théorème 8.5 *Soient $f : [a,b] \to [a,b]$ continue et $(x_n)_{n\in\mathbb{N}}$ une orbite suivant f. Cette suite est convergente si et seulement si :*

$$\lim_{n\to+\infty} (x_{n+1} - x_n) = 0$$

Démonstration. Si $(x_n)_{n\in\mathbb{N}}$ converge il est facile de vérifier que $\lim\limits_{n\to+\infty}(x_{n+1}-x_n)=0.$

Réciproquement si $\lim\limits_{n\to+\infty}(x_{n+1}-x_n)=0$ alors toute valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$ est un point fixe de f. En effet si ℓ est une valeur d'adhérence de cette suite et $\left(x_{\varphi(n)}\right)_{n\in\mathbb{N}}$ une suite extraite convergente vers ℓ, alors $\lim\limits_{n\to+\infty}x_{\varphi(n)+1}=\lim\limits_{n\to+\infty}f\left(x_{\varphi(n)}\right)=f(\ell)$ (f est continue) et avec $\lim\limits_{n\to+\infty}\left(x_{\varphi(n)+1}-x_{\varphi(n)}\right)=0$ on déduit qu'on a aussi $\lim\limits_{n\to+\infty}x_{\varphi(n)+1}=\ell$ et donc $\ell=f(\ell)$.

D'autre part, la suite $(x_n)_{n\in\mathbb{N}}$ étant à valeurs dans un compact admet des valeurs d'adhérence et le lemme précédent nous dit que l'ensemble de ces valeurs d'adhérence est un intervalle. Si cet intervalle n'est pas réduit à un point on dispose de deux valeurs d'adhérence $\alpha<\beta$ et tout point $\gamma\in\,]\alpha,\beta[$ étant valeur d'adhérence on dispose d'une infinité d'indices n tels que $x_n\in\,]\alpha,\beta[$ et considérant qu'un tel x_n est alors valeur d'adhérence de $(x_n)_{n\in\mathbb{N}}$, donc point fixe de f, on a $x_{n+1}=f(x_n)=x_n$, c'est-à-dire que la suite $(x_n)_{n\in\mathbb{N}}$ est stationnaire à partir d'un certain rang, elle est donc convergente et n'a qu'une valeur d'adhérence contrairement à l'hypothèse de départ.

En définitive $(x_n)_{n\in\mathbb{N}}$ n'a qu'une valeur d'adhérence, elle est donc convergente et c'est nécessairement vers un point fixe de f. ■

Le théorème du point fixe sur un segment nous donne une condition suffisante d'existence et d'unicité du point fixe et nous assure de la convergence de toute orbite vers ce point fixe.

Définition 8.3 *On dit qu'une fonction $f:I\longrightarrow\mathbb{R}$ est strictement contractante s'il existe une constante $\lambda\in[0,1[$ telle que :*

$$\forall(x,y)\in I^2,\ |f(x)-f(y)|\leq\lambda|x-y|$$

On dit aussi que f est λ-contractante.

Il est facile de vérifier qu'une application contractante est uniformément continue.

Théorème 8.6 (du point fixe) *Si $f:I\to I$ est une application λ-contractante avec $\lambda\in[0,1[$, alors f admet un unique point fixe α dans I et ce point fixe est la limite de toute orbite suivant f.*
Une majoration de l'erreur d'approximation est donnée par :

$$\forall n\in\mathbb{N},\ |\alpha-x_n|\leq\frac{\lambda^n}{1-\lambda}|x_1-x_0|$$

Démonstration. Pour l'existence du point fixe, on vérifie que la suite $(x_n)_{n\in\mathbb{N}}$ est de Cauchy.

Pour $q>p\geq0$, on a :

$$|x_q-x_p|\leq|x_q-x_{q-1}|+\cdots+|x_{p+1}-x_p|$$

avec pour tout $k\geq1$:

$$\begin{aligned}|x_{k+1}-x_k|&=|f(x_k)-f(x_{k-1})|\\&\leq\lambda|x_k-x_{k-1}|\leq\cdots\leq\lambda^k|x_1-x_0|\end{aligned}$$

Ce qui nous donne :

$$\begin{aligned}|x_q-x_p|&\leq\left(\sum_{k=p}^{q-1}\lambda^k\right)|x_1-x_0|=\frac{\lambda^p\left(1-\lambda^{q-p}\right)}{1-\lambda}|x_1-x_0|\\&\leq\frac{\lambda^p}{1-\lambda}|x_1-x_0|\underset{p\to+\infty}{\to}0\end{aligned}$$

si $\lambda \in [0,1[$.

La suite $(x_n)_{n\in\mathbb{N}}$ étant de Cauchy, elle converge et sa limite α est dans I puisque cet intervalle est fermé. Comme f est continue, α est un point fixe de f dans I.

On peut aussi utiliser le théorème 8.5 pour conclure plus rapidement : de $|x_{k+1} - x_k| \leq \lambda^k |x_1 - x_0|$, on déduit que $\lim\limits_{k\to+\infty} (x_{k+1} - x_k) = 0$ et la suite $(x_n)_{n\in\mathbb{N}}$ converge vers un point fixe de f dans I.

En faisant tendre q vers l'infini dans la majoration précédente (à p fixé) on obtient la majoration de l'erreur :

$$|\alpha - x_p| \leq \frac{\lambda^p}{1-\lambda} |x_1 - x_0|$$

Si $\beta \in I$ est un autre point fixe de f dans I, on a alors :

$$|\alpha - \beta| = |f(\alpha) - f(\beta)| \leq \lambda |\alpha - \beta|$$

et nécessairement $\alpha = \beta$ si $\lambda \in [0,1[$. ∎

Corollaire 8.1 *Si $f : I \longrightarrow I$ est dérivable sur $\overset{\circ}{I}$ avec $\sup\limits_{x\in\overset{\circ}{I}} |f'(x)| = \lambda < 1$ alors f admet un unique point fixe $\alpha \in I$ et ce point fixe est limite de toute orbite suivant f.*

Démonstration. Résulte immédiatement du théorème des accroissements finis et du théorème du point fixe. ∎

Remarque 8.13 *Une fonction $f : I \longrightarrow I$ dérivable sur I avec $|f'(x)| < 1$ pour tout $x \in I$ n'a pas nécessairement de point fixe dans I. Par exemple la fonction f définie sur $I = [1,+\infty[$ par $f(x) = x + \frac{1}{x}$ est telle que $0 \leq f'(x) = 1 - \frac{1}{x^2} < 1$ pour tout $x \in I$ et l'équation $f(x) = x$ n'a pas de solution. Une telle fonction n'est pas strictement contractante.*

Exercice 8.6 *Montrer que pour $0 < r < 1$, la fonction f définie par $f(x) = rx(1-x)$ est strictement contractante sur $I = [0,1]$.*

Solution 8.6 *La fonction f est de classe $\mathcal{C}^1$ sur I avec :*

$$\forall x \in I,\ |f'(x)| = r|1-2x| \leq r = |f'(0)|$$

On a donc $\sup\limits_{x\in I} |f'(x)| = r < 1$ et f est strictement contractante sur I pour $0 < r < 1$.

Exercice 8.7 *Soit a un réel. Etudier la nature et le comportement de la suite définie par*

$$\begin{cases} x_0 = a \\ \forall n \in \mathbb{N},\ x_{n+1} = \cos(x_n) \end{cases}$$

Solution 8.7 *Comme :*

$$\cos(\mathbb{R}) = [-1;1] \text{ et } \cos([-1;1]) \subset [0,1]$$

le terme x_2 est dans $[0,1]$. On peut donc se limiter au segment $[0,1]$.
Avec :

$$\sup_{x\in[0,1]} |\cos'(x)| = \sup_{x\in[0,1]} \sin(x) = \sin(1) < 1$$

on déduit que la fonction cos *est contractante sur* $[0,1]$ *et la suite* $(x_n)_{n\in\mathbb{N}}$ *converge vers l'unique point fixe de cette fonction sur* $[0,1]$, *une majoration de l'erreur d'approximation étant :*

$$\forall n \in \mathbb{N},\ |\alpha - x_n| \leq \frac{(\sin(1))^n}{1-\sin(1)} |\cos(a) - a|$$

De plus, comme la fonction cos *est décroissante sur* $[0,1]$, *les suites extraites* $(x_{2n})_{n\in\mathbb{N}}$ *et* $(x_{2n+1})_{n\in\mathbb{N}}$ *sont adjacentes et on a :*

$$\forall n \geq 2,\ |\alpha - x_n| \leq |x_n - x_{n-1}|$$

Exercice 8.8 *Soit* $f \in \mathcal{C}^1(I)$ *telle que l'équation* $f(x) = 0$ *admette une unique racine* α *dans* $\overset{\circ}{I}$. *Cette équation est équivalente à l'équation :*

$$f_\lambda(x) = x - \lambda f(x) = x$$

où λ *est un réel non nul à choisir judicieusement. On suppose que* $f'(\alpha) > 0$.

1. *Montrer qu'il existe un réel* $\eta > 0$ *tel que l'intervalle* $J = [\alpha-\eta, \alpha+\eta]$ *soit contenu dans* I *et* $f'(x) > 0$ *pour tout* $x \in J$.
2. *On note* $m_1 = \inf_{x\in J} f'(x)$, $M_1 = \sup_{x\in J} f'(x)$. *Montrer que* $0 < m_1 \leq M_1$ *et qu'il existe un unique réel* $\lambda > 0$ *tel que :*

$$\forall x \in J,\ |f'_\lambda(x)| \leq \frac{M_1 - m_1}{M_1 + m_1}$$

3. *Montrer que pour tout réel* $x_0 \in J$ *on peut définir la suite* $(x_n)_{n\in\mathbb{N}}$ *de points de* J *par* $x_{n+1} = f_\lambda(x_n)$ *pour tout* $n \in \mathbb{N}$, *où* λ *est défini à la question précédente, que cette suite converge vers* α *et donner une majoration de l'erreur d'approximation.*
4. *On désigne par* f *la fonction définie sur* $\left]3\frac{\pi}{2}, 5\frac{\pi}{2}\right[$ *par* $f(x) = \tan(x) - x$.
   (a) *Montrer que l'équation* $f(x) = 0$ *admet une unique solution* α *dans l'intervalle* $J = [a,b] = [7.65, 7.75]$.
   (b) *Déterminer le coefficient* λ *qui intervient dans la question 2.*
   (c) *Proposer un procédé d'accélération de la convergence de la suite* $(x_n)_{n\in\mathbb{N}}$ *et donner une valeur approchée de* α *à* 10^{-6} *près.*

Solution 8.8

1. *Avec la continuité de* f' *et l'hypothèse* $f'(\alpha) > 0$ *on déduit que* $f'(x) > 0$ *dans un voisinage de* α *et considérant que* α *est intérieur à* I, *on déduit qu'il existe* $\eta > 0$ *tel que* $J = [\alpha-\eta, \alpha+\eta] \subset I$ *et* $f'(x) > 0$ *pour tout* $x \in J$.
2. *La fonction* f' *étant continue sur le compact* J *est bornée sur cet intervalle et atteint ses bornes, on peut donc poser* $m_1 = \inf_{x\in J} f'(x)$, $M_1 = \sup_{x\in J} f'(x)$ *et il existe* u, v *dans* J *tels que* $0 < m_1 = f'(u) \leq M_1 = f'(v)$.
   Pour tout réel non nul λ *et tout* x *dans* J, *on a* $f'_\lambda(x) = 1 - \lambda f'(x)$ *et :*

$$-\frac{M_1 - m_1}{M_1 + m_1} \leq f'_{\lambda_1}(x) \leq \frac{M_1 - m_1}{M_1 + m_1} \Leftrightarrow \frac{2m_1}{M_1 + m_1} \leq \lambda f'(x) \leq \frac{2M_1}{M_1 + m_1}$$

Pour un tel λ on a alors nécessairement :

$$\frac{2m_1}{M_1+m_1} \leq \lambda f'(u) = \lambda m_1 \text{ et } \lambda M_1 = \lambda f'(v) \leq \frac{2M_1}{M_1+m_1}$$

soit $\frac{2}{M_1+m_1} \leq \lambda \leq \frac{2}{M_1+m_1}$, *c'est-à-dire que* $\lambda = \frac{2}{M_1+m_1}$ *est la seule valeur possible. On vérifie facilement que cette valeur convient.*

3. *Pour* $\lambda = \frac{2}{M_1+m_1}$, *on a* $\sup_{x\in J} |f'_\lambda(x)| \leq \frac{M_1-m_1}{M_1+m_1} < 1$. *En utilisant le théorème des accroissements finis, on en déduit alors que l'intervalle* $J = [\alpha-\eta, \alpha+\eta]$ *est stable par* f_λ *et que l'application* f_λ *est strictement contractante sur cet intervalle. En effet, pour tout* x *dans* J, *on a :*

$$|f_\lambda(x) - \alpha| = |f_\lambda(x) - f_\lambda(\alpha)| \leq \frac{M_1-m_1}{M_1+m_1}|x-\alpha| < |x-\alpha| \leq \eta$$

ce qui signifie que $f_\lambda(x) \in J$. *Avec le même argument on vérifie que* f_λ *est strictement contractante. Il en résulte que pour tout* $x_0 \in J$ *la suite* $(x_n)_{n\in\mathbb{N}}$ *est bien définie et que cette suite converge vers* α. *Une majoration de l'erreur d'approximation étant donnée par :*

$$|x_n - \alpha| \leq \left(\frac{M_1-m_1}{M_1+m_1}\right)^n |x_0-\alpha| \leq \eta\left(\frac{M_1-m_1}{M_1+m_1}\right)^n$$

4.

(a) *La fonction* f *est continue strictement croissante sur* J ($f'(x) = \tan^2(x) > 0$ *pour tout* $x \in I \subset \left]2\pi, 5\frac{\pi}{2}\right[$) *avec* $f(a)f(b) < 0$, *en conséquence l'équation* $f(x) = 0$ *admet une unique solution* $\alpha \in J$.

(b) *On a :*

$$\begin{cases} m_1 = \inf_{x\in J} f'(x) = \tan^2(a) \backsimeq 23.36 \\ M_1 = \sup_{x\in J} f'(x) = \tan^2(b) \backsimeq 91.82 \end{cases}$$

ce qui donne $\lambda = \frac{2}{M_1+m_1} \backsimeq 7.73 \times 10^{-2}$. *Un algorithme de calcul approché de* α *est donc donné en considérant la suite* $(x_n)_{n\in\mathbb{N}}$ *définie par :*

$$\begin{cases} x_0 = 7.7 \in J \\ \forall n \geq 0,\ x_{n+1} = x_n - \lambda(\tan(x_n) - x_n) \end{cases}$$

(c) *On peut accélérer la convergence en procédant comme suit. On calcule tout d'abord* x_1 *et on a* $f(x_1) < 0$, $f(b) > 0$, *ce qui nous amène à remplacer l'intervalle* $[a,b]$ *par* $[a_1,b_1] = [x_1,b]$ *et le triplet* (m_1, M_1, λ) *par* (m_2, M_2, λ_1) *avec :*

$$\begin{cases} m_2 = \tan^2(a_1) \backsimeq 56.59 \\ M_2 = \tan^2(b_1) \backsimeq 91.82 \\ \lambda_1 = \frac{2}{M_2+m_2} \backsimeq 1.34 \times 10^{-2} \end{cases}$$

On calcule alors l'itéré correspondant, ce qui donne $x_2 = f_{\lambda_1}(x_1) \backsimeq 7.72$. *On regarde le signe de* $f(x_2)$ *et on recommence. Au bout de* 9 *itérations on obtient* $\alpha \backsimeq 7.725251$ *qui donne une approximation à* 10^{-6} *près de* α.

Une fonction $f : I \to I$ telle que $|f(x) - f(y)| < |x - y|$ pour tous $x \neq y$ dans I n'a pas nécessairement de point fixe dans I comme le montre l'exemple de $f(x) = x + \dfrac{1}{x}$ sur $I = [1, +\infty[$. Mais dans le cas où I est compact on a le résultat suivant.

Théorème 8.7 *Soient I un intervalle réel compact et $f : I \to I$ telle que :*

$$\forall (x, y) \in I^2,\ (x \neq y \Rightarrow |f(x) - f(y)| < |x - y|)$$

La fonction f admet un unique point fixe dans I et ce point fixe est limite de toute orbite suivant f.

Démonstration. La fonction f qui est lipschitzienne est en particulier uniformément continue.

La fonction $x \longmapsto |f(x) - x|$ étant continue sur le segment I est bornée et atteint ses bornes. En particulier, il existe un réel $\alpha \in I$ tel que :

$$|f(\alpha) - \alpha| = \inf_{x \in I} |f(x) - x|$$

Si $f(\alpha) \neq \alpha$ on a alors $|f(\alpha) - f(f(\alpha))| < |\alpha - f(\alpha)|$ avec $f(\alpha) \in I$, ce qui est contradictoire avec la définition de α. On a donc $f(\alpha) = \alpha$, c'est-à-dire que α est point fixe de f dans I.

Si $\beta \neq \alpha$ est un autre point fixe de f dans I, on a alors :

$$|\alpha - \beta| = |f(\alpha) - f(\beta)| < |\alpha - \beta|$$

ce qui est impossible. En conclusion, f admet un unique point fixe dans I.

Soit $(x_n)_{n \in \mathbb{N}}$ une orbite associée à f. Si il existe $n_0 \in \mathbb{N}$ tel que $x_{n_0} = \alpha$, la suite est alors stationnaire, donc convergente. On suppose donc que $x_n \neq \alpha$ pour tout $n \in \mathbb{N}$.

On a alors $0 < |x_{n+1} - \alpha| = |f(x_n) - f(\alpha)| < |x_n - \alpha|$ pour tout $n \in \mathbb{N}$, c'est-à-dire que la suite $(|x_n - \alpha|)_{n \in \mathbb{N}}$ est strictement décroissante minorée par 0, elle est donc convergente vers un réel $\delta \geq 0$. La suite $(x_n)_{n \in \mathbb{N}}$ étant à valeurs dans le compact I, on peut en extraire une sous-suite $\left(x_{\varphi(n)}\right)_{n \in \mathbb{N}}$ qui converge vers un réel $\beta \in I$. On a alors :

$$|f(\beta) - \alpha| = \lim_{n \to +\infty} \left|f\left(x_{\varphi(n)}\right) - \alpha\right| = \lim_{n \to +\infty} \left|x_{\varphi(n)+1} - \alpha\right| = \delta$$

et $\beta \neq \alpha$ entraîne :

$$\delta = |\alpha - f(\beta)| = |f(\alpha) - f(\beta)| < |\alpha - \beta| = \lim_{n \to +\infty} \left|x_{\varphi(n)} - \alpha\right| = d$$

ce qui est impossible. On a donc $\beta = \alpha$, soit $\delta = 0$ et $\lim\limits_{n \to +\infty} x_n = \alpha$. ■

Les intervalles réels étant convexes, on en déduit le résultat suivant.

Corollaire 8.2 *Soient I un intervalle réel compact et $f : I \to I$ telle que :*

$$\forall (x, y) \in I^2,\ |f(x) - f(y)| \leq |x - y|$$

La fonction f admet alors un point fixe dans K.

Démonstration. On se fixe un point a dans I et on définit la suite de fonctions $(f_n)_{n\geq 1}$ sur I par :

$$\forall n \geq 1,\ \forall x \in I,\ f_n(x) = f\left(\frac{1}{n}a + \left(1 - \frac{1}{n}\right)x\right)$$

(c'est la convexité de I qui permet de définir les f_n).

Chaque fonction f_n est à valeurs dans I et pour x, y dans I, on a :

$$|f_n(x) - f_n(y)| \leq \left(1 - \frac{1}{n}\right)|x - y|$$

Pour $n \geq 1$ la fonction f_n est strictement contractante de I dans I avec I fermé, elle admet donc un unique point fixe $\alpha_n \in I$. De la suite $(\alpha_n)_{n\geq 1}$ dans le compact I on peut extraire une sous-suite $\left(\alpha_{\varphi(n)}\right)_{n\in\mathbb{N}}$ qui converge vers un réel $\alpha \in I$. En notant $g_n = f_{\varphi(n)}$, et $\beta_n = \alpha_{\varphi(n)}$, on a pour tout $n \geq 1$:

$$|f(\alpha) - g_n(\beta_n)| \leq |f(\alpha) - g_n(\alpha)| + |g_n(\alpha) - g_n(\beta_n)|$$

avec :

$$|f(\alpha) - g_n(\alpha)| = \left|f(\alpha) - f\left(\frac{1}{\varphi(n)}a + \left(1 - \frac{1}{\varphi(n)}\right)\alpha\right)\right| \leq \frac{1}{\varphi(n)}|\alpha - a|$$

et :

$$|g_n(\alpha) - g_n(\beta_n)| = \left\|f_{\varphi(n)}(\alpha) - f_{\varphi(n)}\left(\alpha_{\varphi(n)}\right)\right\| \leq \left(1 - \frac{1}{\varphi(n)}\right)\left|\alpha - \alpha_{\varphi(n)}\right|$$

ce qui entraîne que $\lim\limits_{n\to+\infty} f_{\varphi(n)}\left(\alpha_{\varphi(n)}\right) = f(\alpha)$. Tenant compte du fait que $\alpha_{\varphi(n)}$ est point fixe de $f_{\varphi(n)}$, on a $f_{\varphi(n)}\left(\alpha_{\varphi(n)}\right) = \alpha_{\varphi(n)}$ et $f(\alpha) = \lim\limits_{n\to+\infty} \alpha_{\varphi(n)} = \alpha$, c'est-à-dire que α est point fixe de f dans I. ■

8.3.3 Points fixes attractifs ou répulsifs

Pour $f : I \longrightarrow I$ dérivable, c'est la pente de la tangente au voisinage de $\alpha \in I$, solution de $f(x) = x$, qui gouverne la convergence d'une orbite. Plus f' est proche de 0 au voisinage de α plus rapide sera la convergence. Le théorème suivant précise cette remarque.

Théorème 8.8 *Soit $f \in \mathcal{C}^1(I)$ admettant un unique point fixe $\alpha \in \overset{\circ}{I}$.*

(i) Si $|f'(\alpha)| < 1$ il existe alors un réel $\eta > 0$ tel que l'intervalle $[\alpha - \eta, \alpha + \eta]$ soit stable par f et pour tout $x_0 \in [\alpha - \eta, \alpha + \eta]$ l'orbite $(x_n)_{n\in\mathbb{N}}$ converge vers α.

(ii) Si $|f'(\alpha)| > 1$ et $f(I) \subset I$ alors pour tout $x_0 \in I$, l'orbite $(x_n)_{n\in\mathbb{N}}$ est soit stationnaire (sur α) à partir d'un certain rang, soit divergente.

Démonstration. Si on suppose que $|f'(\alpha)| < 1$, de la continuité de f', on déduit qu'il existe un réel $\eta > 0$ tel que $[\alpha - \eta, \alpha + \eta] \subset I$ (α est dans l'intérieur de I) et :

$$\forall x \in [\alpha - \eta, \alpha + \eta],\ |f'(x)| < 1$$

On a alors $\lambda = \sup\limits_{x\in I} |f'(x)| < 1$ (la borne supérieure d'une fonction continue sur un compact est atteinte) et en utilisant le théorème des accroissements finis :

$$|f(x) - \alpha| = |f(x) - f(\alpha)| \leq \lambda |x - \alpha| < |x - \alpha| \leq \eta$$

pour tout $x \in [\alpha - \eta, \alpha + \eta]$, c'est-à-dire que $f(x) \in [\alpha - \eta, \alpha + \eta]$.

L'intervalle $[\alpha - \eta, \alpha + \eta]$ est donc stable par f. La fonction f étant strictement contractante sur cet segment, on a le résultat annoncé.

Supposons que $|f'(\alpha)| > 1$. Avec la continuité de f', on déduit qu'il existe un réel $\eta > 0$ et un réel $\lambda > 1$ tels que $[\alpha - \eta, \alpha + \eta] \subset I$ et :

$$\forall x \in [\alpha - \eta, \alpha + \eta], \ |f'(x)| \geq \lambda > 1$$

Si on suppose de plus que I est stable par f on peut alors définir l'orbite $(x_n)_{n\in\mathbb{N}}$ de premier terme $x_0 \in I$.

S'il existe $n_0 \in \mathbb{N}$ tel que $x_{n_0} = \alpha$, on a alors $x_n = \alpha$ pour tout $n \geq n_0$ et la suite est stationnaire à partir du rang n_0.

On suppose donc que $x_n \neq \alpha$ pour tout $n \in \mathbb{N}$ et dans ce cas la suite $(x_n)_{n\in\mathbb{N}}$ est divergente. En effet si $\lim\limits_{n\to+\infty} x_n = \beta$, on a alors $f(\beta) = \beta$ puisque f est continue et $\beta = \alpha$ puisque α est l'unique point fixe de f dans I. Mais on peut alors trouver un entier $n_1 \in \mathbb{N}$ tel que $x_n \in [\alpha - \eta, \alpha + \eta]$ pour tout $n \geq n_1$ et avec le théorème des accroissements finis on déduit que :

$$\forall n \geq n_1, \ |x_{n+1} - \alpha| = |f(x_n) - f(\alpha)| \geq \lambda |x_n - \alpha|$$

puis par récurrence sur $n \geq n_1$:

$$\forall n \geq n_1, \ |x_n - \alpha| \geq \lambda^{n-n_1} |x_{n_1} - \alpha|$$

ce qui entraîne $\lim\limits_{n\to+\infty} |x_n - \alpha| = +\infty$ $(\lambda > 1)$ et contredit $\lim\limits_{n\to+\infty} x_n = \alpha$. ■

Remarque 8.14 *Avec les notations du théorème précédent, on dit que α est un point fixe attractif dans le premier cas et que c'est un point fixe répulsif dans le second (figures 8.4 et 8.3).Dans le cas où $|f'(\alpha)| = 1$ on ne peut rien dire à priori (figure (8.5)).*

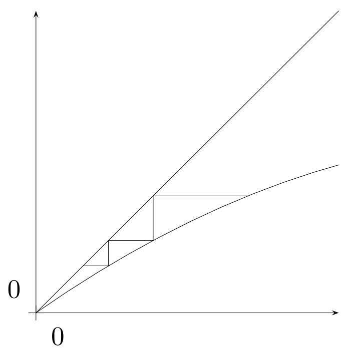

FIGURE 8.3 – Point fixe attractif

Exercice 8.9 *Soient r un réel strictement positif et f la fonction définie sur $I = [0, 1]$ par $f(x) = rx(1 - x)$.*

1. *Donner une condition nécessaire et suffisante sur r pour que I soit stable par f.*
   On suppose cette condition vérifiée pour la suite.
2. *Étudier les orbites de f pour $r \in]0, 1]$.*

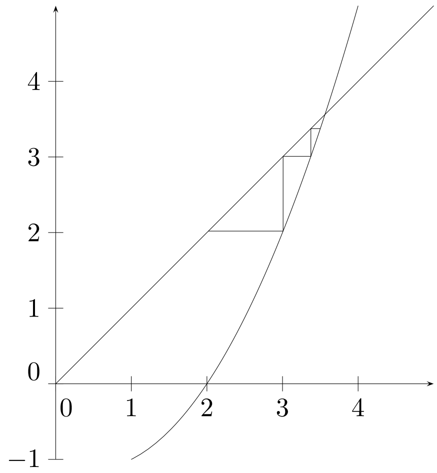

FIGURE 8.4 – Point fixe répulsif

Solution 8.9

1. *On a* $\inf\limits_{x \in I} f(x) = f(0) = 0$ *et* $\sup\limits_{x \in I} f(x) = f\left(\dfrac{1}{2}\right) = \dfrac{r}{4}$, *donc* $f(I) = \left[0, \dfrac{r}{4}\right]$ *et* I *est stable par* f *si, et seulement si,* $r \in \left]0, 4\right]$.
2. *Supposons que* $r \in \left]0, 1\right]$.
   L'équation $f(x) = x$ *équivaut à* $x(rx + (1 - r)) = 0$ *qui a pour unique solution* $\alpha = 0$ *dans* I *(pour* $r = 1$ *c'est clair et pour* $0 < r < 1$, *l'autre solution* $\dfrac{r-1}{r}$ *est strictement négative).*
   Comme :
   $$\sup_{x \in I} |f'(x)| = r \sup_{x \in I} |1 - 2x| = r$$
   ce point fixe $\alpha = 0$ *est attractif pour* $r \in \left]0, 1\right[$ *et toute orbite suivant* f *est convergente.*
   Pour $r = 1$, *on a* $\sup\limits_{x \in I} |f'(x)| = 1$.
   Dans ce cas, une orbite est définie par $x_0 \in I$ *et* $x_{n+1} = x_n - x_n^2$ *pour* $n \geq 0$.
   Avec $x_{n+1} - x_n = -x_n^2 \leq 0$, *on déduit que la suite* $(x_n)_{n \in \mathbb{N}}$ *est décroissante. Comme elle est minorée par* 0, *elle converge et c'est nécessairement vers* 0.
   En fait, pour $0 < r \leq 1$, *on a :*
   $$x_{n+1} - x_n = (r - 1) x_n - r x_n^2 \leq 0$$
   puisque $x_n \geq 0$ *pour tout* n. *Là encore la suite est décroissante minorée.*

Exercice 8.10 *Soit* f *la fonction définie sur* $I = [-1, 1]$ *par* $f(x) = 2x^2 - 1$.

1. *Montrer que l'intervalle* I *est stable par* f *et que* f *a deux points fixes* $\alpha > \beta$ *dans* I.
2. *Soit* $(x_n)_{n \in \mathbb{N}}$ *une orbite suivant* f.
   (a) *Justifier l'existence d'un unique réel* $\theta \in [0, \pi]$ *tel que* $x_0 = \cos(\theta)$.
   (b) *Exprimer, pour tout entier naturel* n, *le réel* x_n *en fonction de* n *et* θ.
   (c) *Déterminer l'ensemble* A_α *[resp.* A_β*] des réels* $x_0 \in [-1, 1]$ *pour lesquels la suite* $(x_n)_{n \in \mathbb{N}}$ *est convergente vers* α *[resp* β*]. Montrer que* A_α *[resp.* A_β*] est une partie dénombrable et dense de* $[-1, 1]$.

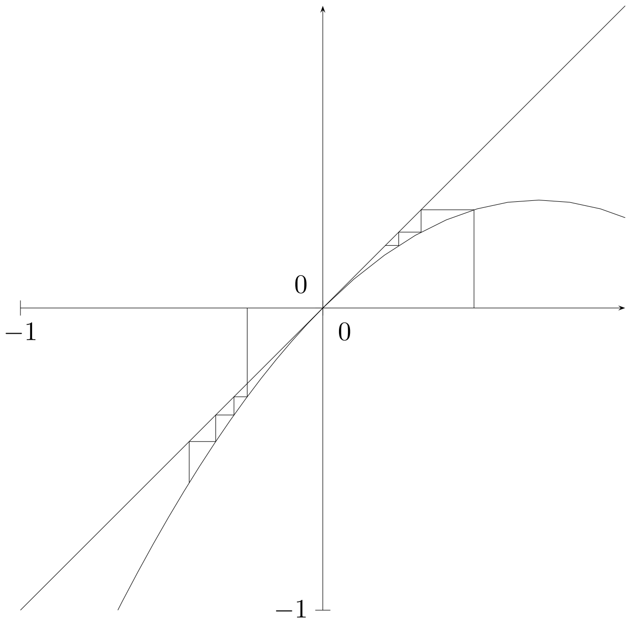

FIGURE 8.5 – Point fixe ni attractif, ni répulsif

(d) *En déduire l'ensemble A des réels $x_0 \in [-1, 1]$ pour lesquels la suite $(x_n)_{n\in\mathbb{N}}$ est convergente.*

(e) *Montrer que si θ est rationnel, la suite $(x_n)_{n\in\mathbb{N}}$ est alors périodique à partir d'un certain rang.*

Solution 8.10

1. *On a* $f(I) = I$ *et les points fixes de* f *sont* $\alpha = 1$ *et* $\beta = -0.5$.
2.
   (a) *La fonction* $\cos$ *étant bijective de* $[0, \pi]$ *sur* $[-1, 1]$, *il existe un unique réel* $\theta \in [0, \pi]$ *tel que* $x_0 = \cos(\theta)$.

   (b) *On vérifie par récurrence sur* $n \geq 0$, *que* $x_n = \cos(2^n \theta)$.
   C'est vrai pour $n = 0$ *et supposant le résultat acquis pour* $n \geq 0$, *on a :*

$$x_{n+1} = f(x) = 2\cos^2(2^n\theta) - 1 = \cos(2 \cdot 2^n \theta) = \cos\left(2^{n+1}\theta\right)$$

   (c) *Les deux points fixes sont répulsifs (*$f'(-0.5) = -2$ *et* $f'(1) = 4$*), donc la suite* $(x_n)_{n\in\mathbb{N}}$ *est convergente si, et seulement si, elle est stationnaire à partir d'un certain rang, ce qui revient à dire qu'il existe un entier* $p \geq 0$ *tel que* $\cos(2^p\theta) = \alpha = 1$ *[resp.* $\cos(2^p\theta) = \beta = -0.5 = \cos\left(\frac{4\pi}{3}\right)$*], encore équivalent à dire qu'il existe deux entiers naturels* p *et* q *tels que* $\theta = \frac{q}{2^{p-1}}\pi$ *[resp.* $\theta = \left(\frac{1}{3} + q\right)\frac{\pi}{2^{p-1}}$*] et* $\theta \in [0, \pi]$. *On a donc :*

$$A_\alpha = \{\cos(r\pi) \mid r \in \mathcal{D} \cap [0, 1]\}$$
$$\textit{resp. } A_\beta = \{\cos(r\pi) \mid r \in \mathcal{D}' \cap [0, 1]\}$$

   où $\mathcal{D}$ *est l'ensemble des nombres dyadiques (i. e. l'ensemble des nombres rationnels de la forme* $\frac{a}{2^m}$ *avec* a *et* m *dans* $\mathbb{Z}$*) et* $\mathcal{D}'$ *l'ensemble des nombres rationnels de la forme* $\frac{a}{2^m}$ *avec* $a \in \frac{1}{3} + \mathbb{Z}$ *et* $m \in \mathbb{Z}$.
   Les ensemble $\mathcal{D}$ *et* $\mathcal{D}'$ *qui sont contenus dans* $\mathbb{Q}$ *sont dénombrables, donc aussi les ensembles* A_α *et* A_β. *De la densité de* $\mathcal{D}$ *et* $\mathcal{D}'$ *dans* $\mathbb{R}$ *(celle de* $\mathcal{D}$ *est connue et si* $(r_n)_{n\in\mathbb{N}}$ *est une suite de* $\mathcal{D}$ *qui converge vers* x, *alors la suite* $\left(\frac{1}{3 \cdot 2^n} + r_n\right)_{n\in\mathbb{N}}$ *est dans* $\mathcal{D}'$ *et converge vers* x*) et de la continuité de* $\cos$, *on déduit que* A_α *[resp.* A_β*] dense dans* $[-1, 1]$.

   (d) *Comme la suite ne peut converger que vers* α *ou* β, *on a* $A = A_\alpha \cup A_\beta$. *Cette partie est dénombrable et dense dans* $[-1, 1]$.

3. *Le réel* $\frac{\theta}{\pi}$ *a un développement dyadique de la forme :*

$$\frac{\theta}{\pi} = \sum_{k=0}^{+\infty} \frac{\varepsilon_k}{2^k}$$

où $\varepsilon_k \in \{0, 1\}$ *pour tout entier naturel* k. *Ce qui nous donne, pour tout entier* $n \geq 1$ *:*

$$\begin{aligned} x_n = \cos(2^n\theta) &= \cos\left(\left(\sum_{k=0}^{n-1} \varepsilon_k \cdot 2^{n-k}\right)\pi + \left(\sum_{k=n}^{+\infty} \frac{\varepsilon_k}{2^{k-n}}\right)\pi\right) \\ &= \cos\left(\left(\sum_{k=0}^{+\infty} \frac{\varepsilon_{k+n}}{2^k}\right)\pi\right) \end{aligned}$$

Dans le cas où θ est rationnel, son développement dyadique est périodique à partir d'un certain rang, c'est à dire qu'il existe un entier $n_0 \geq 0$ et un entier $p \geq 1$ tels que :

$$\forall n \geq n_0,\ \varepsilon_{n+p} = \varepsilon_n$$

Il en résulte que :

$$\forall n \geq n_0,\ x_{n+p} = \cos\left(\left(\sum_{k=0}^{+\infty} \frac{\varepsilon_{k+n+p}}{2^k}\right)\pi\right) = \cos\left(\left(\sum_{k=0}^{+\infty} \frac{\varepsilon_{k+n}}{2^k}\right)\pi\right) = x_n$$

c'est-à-dire que $(x_n)_{n\in\mathbb{N}}$ est périodique à partir d'un certain rang.

8.3.4 Le théorème itéré du point fixe

Il peut arriver que la fonction f ne soit pas strictement contractante mais que l'une de ces itérées le soit. Dans ce cas on a encore existence et unicité du point fixe de f qui est limite de toute suite d'approximations successives.

Lemme 8.4 *Soient $p \in \mathbb{N}-\{0\}$ et $(x_n)_{n\in\mathbb{N}}$ une suite de réels telle que pour tout $r \in \{0, \cdots, p-1\}$ la suite extraite $(x_{jp+r})_{j\in\mathbb{N}}$ converge vers x. La suite $(x_n)_{n\in\mathbb{N}}$ converge alors vers x.*

Démonstration. Soit $\varepsilon > 0$. Pour tout $r \in \{0, \cdots, p-1\}$, on peut trouver un entier j_r tel que :

$$\forall j \geq j_r,\ |x_{jp+r} - x| < \varepsilon$$

On note $j_\varepsilon = \max\limits_{0\leq r\leq p-1} j_r$. Tout entier $n \in \mathbb{N}$ s'écrit de manière unique $n = jp + r$ avec $j \in \mathbb{N}$, $r \in \{0, \cdots, p-1\}$ et pour $j \geq j_\varepsilon$ on a $|x_n - x| < \varepsilon$. On déduit donc que :

$$\forall n \geq (j_\varepsilon + 1)\,p,\ |x_n - x| < \varepsilon$$

On a donc ainsi prouvé que la suite $(x_n)_{n\in\mathbb{N}}$ converge vers x. ■

Théorème 8.9 *Si $f : I \to I$ est une application telle que l'une des itérées $f^p = \underbrace{f \circ \cdots \circ f}_{p \text{ fois}}$ soit strictement contractante, elle admet alors un unique point fixe dans I limite de toute suite d'approximations successives.*

Démonstration. L'application f^p étant strictement contractante sur l'intervalle fermé I, elle admet un unique point fixe $\alpha \in I$. De $f^p(\alpha) = \alpha$ on déduit que $f^p(f(\alpha)) = f(f^p(\alpha)) = f(\alpha)$ et $f(\alpha)$ est aussi point fixe de f^p. On a donc $f(\alpha) = \alpha$ du fait de l'unicité du point fixe de f^p. L'application f admet donc un point fixe dans I.

En considérant que tout point fixe de f est aussi point fixe de f^p, on déduit que ce point fixe est unique dans I.

Si $(x_n)_{n\in\mathbb{N}}$ est définie par $x_0 \in I$ et $x_{n+1} = f(x_n)$ pour tout $n \in \mathbb{N}$, on a alors pour tout $r \in \{0, \cdots, p-1\}$ et tout $j \in \mathbb{N}$:

$$x_{jp+r} = f^{jp+r}(x_0) = (f^p)^j\,(f^r(x_0))$$

c'est-à-dire que $(x_{jp+r})_{j\in\mathbb{N}}$ est une suite d'approximations successives pour f^p de valeur initiale $f^r(x_0)$. On a donc $\lim\limits_{j\to+\infty} x_{jp+r} = \alpha$ pour tout $r \in \{0, \cdots, p-1\}$ ce qui équivaut à dire que $\lim\limits_{n\to+\infty} x_n = \alpha$. ■

8.4 La méthode de Lagrange (ou de la sécante)

L'idée de la méthode de la sécante pour résoudre l'équation $f(x) = 0$ est de remplacer dans un voisinage $]a, b[$ de la solution α de cette équation (en supposant que cette solution existe et est unique dans $]a, b[$) le graphe de f par celui de la sécante passant par les points $M = (x, f(x))$ et $B = (b, f(b))$ (interpolation de Lagrange).

L'équation de cette sécante est, en supposant $x < b$:

$$Y = f(x) + \frac{f(b) - f(x)}{b - x}(X - x)$$

Si $f(b) \neq f(x)$ alors le point d'intersection de cette sécante avec l'axe des x est donné par :

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

La résolution de l'équation $f(x) = 0$ est ainsi ramenée à la résolution du problème de point fixe $g(x) = x$, ce qui revient à considérer la suite $(x_n)_{n \in \mathbb{N}}$ définie par :

$$\begin{cases} x_0 \in I \\ x_{n+1} = x_n - \dfrac{b - x_n}{f(b) - f(x_n)} f(x_n) \quad (n \geq 0) \end{cases}$$

quand cette dernière peut être définie.

Dans le cas où la fonction $f : I \to \mathbb{R}$ est de classe $\mathcal{C}^2$, admet une racine $\alpha \in \overset{\circ}{I}$ et est telle que $f'f''$ garde un signe constant au voisinage de α (f est strictement monotone et à convexité constante sur ce voisinage), la suite $(x_n)_{n \in \mathbb{N}}$ est bien définie, strictement monotone ou stationnaire sur α, converge vers α et on peut donner une majoration de l'erreur d'approximation. Pour démontrer ce résultat on aura besoin du lemme suivant.

Lemme 8.5 *Soit f une fonction de classe $\mathcal{C}^2$ sur un intervalle $[a, b]$ telle que $f(a) f(b) < 0$ et $f'(x) \neq 0$ pour tout $x \in [a, b]$. L'équation $f(x) = 0$ admet une unique solution $\alpha \in]a, b[$, on peut définir la fonction $g : [a, b[\to \mathbb{R}$ par :*

$$\forall x \in [a, b[, \ g(x) = x - \frac{b - x}{f(b) - f(x)} f(x)$$

et :

$$\forall x \in [a, b[, \ |g(x) - \alpha| \leq \frac{b - \alpha}{2} \frac{M_2}{m_1} |x - \alpha|$$

où $M_2 = \sup\limits_{x \in [a,b]} |f''(x)|$ et $m_1 = \inf\limits_{x \in [a,b]} |f'(x)|$.

Démonstration. L'existence d'une racine α de l'équation $f(x) = 0$ se déduit du théorème des valeurs intermédiaires avec l'hypothèse $f(a) f(b) < 0$ et l'unicité est conséquence de la stricte monotonie de f.

En utilisant le théorème des accroissements finis, on peut écrire, pour tout $x \in [a, b[$:

$$f(b) - f(x) = (b - x) f'(c_x)$$

avec $c_x \in]x, b[$ et $f'(c_x) \neq 0$. On a donc $f(b) - f(x) \neq 0$. On peut donc définir la fonction $g : [a, b[\to \mathbb{R}$ par :

$$\forall x \in [a, b[, \ g(x) = x - \frac{b - x}{f(b) - f(x)} f(x)$$

Pour tout $x \in [a, b[$, on a :

$$g(x) - \alpha = \frac{f(b)(x-\alpha) - f(x)(b-\alpha)}{f(b) - f(x)} = \frac{h(x)}{f(b) - f(x)}$$

la fonction h étant de classe $\mathcal{C}^2$ sur $[a, b]$ avec $h(\alpha) = h(b) = 0$.

Pour x fixé dans $[a, b]$, on définit la fonction φ_x par :

$$\forall t \in [a, b[, \ \varphi_x(t) = h(t) - k_x(t-\alpha)(t-b)$$

la constante k_x étant choisie de telle sorte que $\varphi_x(x) = 0$, ce qui équivaut à dire que :

$$k_x = \frac{h(x)}{(x-\alpha)(x-b)}$$

Cette fonction est de classe $\mathcal{C}^2$ sur $[a, b]$ avec $\varphi_x(\alpha) = \varphi_x(b) = \varphi_x(x) = 0$. En utilisant le théorème de Rolle itéré, on déduit qu'il existe c_x dans $]\alpha, b[$ tel que $\varphi''_x(c_x) = 0$, ce qui équivaut à $h''(c_x) - 2k_x = 0$, soit :

$$k_x = \frac{h(x)}{(x-\alpha)(x-b)} = \frac{h''(c_x)}{2} = \frac{-(b-\alpha) f''(c_x)}{2}$$

En notant $M_2 = \sup\limits_{x \in [a,b]} |f''(x)|$, on aboutit à :

$$|h(x)| \leq \frac{b-\alpha}{2} |x-\alpha| (b-x) M_2$$

inégalité encore valable pour $x = \alpha$ et $x = b$. On en déduit alors que pour tout $x \in [a, b[$ on a :

$$|g(x) - \alpha| \leq \frac{b-\alpha}{2} |x-\alpha| \frac{b-x}{|f(b) - f(x)|} M_2$$

En notant $m_1 = \inf\limits_{x \in [a,b]} |f'(x)|$ et en utilisant le théorème des accroissements finis, on a :

$$|f(b) - f(x)| \geq m_1 (b-x)$$

avec $m_1 > 0$ (la borne inférieure de f' est atteinte et $|f'| > 0$), ce qui donne :

$$|g(x) - \alpha| \leq \frac{b-\alpha}{2} \frac{M_2}{m_1} |x-\alpha|$$

■

Théorème 8.10 *Soit f une fonction de classe $\mathcal{C}^2$ sur un intervalle $[a, b]$ telle que $f(a) f(b) < 0$ et $f'(x) f''(x) \neq 0$ pour tout $x \in [a, b]$. Dans ces conditions, l'équation $f(x) = 0$ admet une unique solution $\alpha \in]a, b[$, pour tout x_0 dans $[a, b[$, on peut définir la suite $(x_n)_{n \in \mathbb{N}}$ de points de $[a, b[$ par :*

$$\forall n \in \mathbb{N}, \ x_{n+1} = x_n - \frac{b - x_n}{f(b) - f(x_n)} f(x_n)$$

cette suite est monotone, converge vers α et on a la majoration de l'erreur d'approximation :

$$\forall n \in \mathbb{N}, \ |x_n - \alpha| \leq (b-a) \left(\frac{b-a}{2} \frac{M_2}{m_1} \right)^n$$

où $M_2 = \sup\limits_{x \in [a,b]} |f''(x)|$ et $m_1 = \inf\limits_{x \in [a,b]} |f'(x)|$.

Démonstration. On peut supposer que $f'(x) > 0$ et $f''(x) > 0$ pour tout $x \in [a,b]$. En effet, si $f' < 0$ et $f'' < 0$, on remplace f par $-f$, si $f' < 0$ et $f'' > 0$, on remplace f par $x \mapsto f(-x)$ et si $f' > 0$ et $f'' < 0$, on remplace f par $x \mapsto -f(-x)$.

Dans la démonstration du lemme précédent, on a montré l'existence et l'unicité de la racine α de l'équation $f(x) = 0$ et on a vu qu'on peut définir la fonction $g : [a,b[\to \mathbb{R}$ par :

$$\forall x \in [a,b[,\ g(x) = x - \frac{b-x}{f(b) - f(x)} f(x)$$

Cette fonction est de classe $\mathcal{C}^2$ sur $[a,b[$ et peut se prolonger par continuité en b en posant :

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

Pour tout x dans $[a,b[$, on a :

$$\begin{aligned} g'(x) &= \frac{f(b)}{(f(b) - f(x))^2} (f(b) - f(x) - (b-x) f'(x)) \\ &= \frac{f(b)}{(f(b) - f(x))^2} \frac{(b-x)^2}{2} f''(c_x) \end{aligned}$$

où $c_x \in]x,b[$ (formule de Taylor à l'ordre 2 pour f entre x et b). Tenant compte de $f(b) > f(\alpha) = 0$ et $f''(c_x) > 0$, on déduit que $g'(x) > 0$ pour tout $x \in [a,b[$ et donc que g est strictement croissante sur $[a,b]$.

Pour tout $x \in [a,\alpha[$ on a $f(x) < f(\alpha) = 0$ et $f(b) - f(x) > 0$ (f est strictement croissante) et donc $g(x) > x$. De plus, avec la stricte croissance de g on obtient $g(x) < g(\alpha) = \alpha$. On a donc :

$$\forall x \in [a,\alpha[,\ a \leq x < g(x) < \alpha$$

De même pour tout $x \in]\alpha,b]$, on a $g(x) > g(\alpha) = \alpha$ et $f(x) > f(\alpha) = 0$, donc $g(x) < x$. D'où :

$$\forall x \in]\alpha,b],\ \alpha < g(x) < x \leq b$$

En définitive les intervalles $[a,\alpha]$, $[\alpha,b]$ et donc $[a,b]$ sont stables par g. On peut donc définir la suite $(x_n)_{n\in\mathbb{N}}$ par $x_0 \in [a,b]$ et $x_{n+1} = g(x_n)$ pour tout $n \geq 0$. Pour $x_0 = \alpha$, cette suite est stationnaire sur α.

Des encadrements précédents, on déduit que si $x_0 \in [a,\alpha[$ on a alors :

$$\forall n \in \mathbb{N},\ a \leq x_n < x_{n+1} < \alpha.$$

Cette suite est donc strictement croissante majorée et en conséquence elle converge, sa limite étant α (continuité de g). Pour $x_0 \in]\alpha,b[$, on a :

$$\forall n \in \mathbb{N},\ \alpha \leq x_{n+1} < x_n < b$$

Cette suite est donc strictement décroissante minorée et en conséquence elle converge, sa limite étant α.

En utilisant le lemme précédent, on déduit que pour tout $n \in \mathbb{N}$ on a :

$$|x_{n+1} - \alpha| = |g(x_n) - \alpha| \leq \frac{b-\alpha}{2}\frac{M_2}{m_1} |x_n - \alpha|$$

et par récurrence on aboutit à :

$$\forall n \in \mathbb{N},\ |x_n - \alpha| \leq \left(\frac{b-\alpha}{2}\frac{M_2}{m_1}\right)^n |x_0 - \alpha| \leq (b-a)\left(\frac{b-a}{2}\frac{M_2}{m_1}\right)^n$$

■

8.5 La méthode de Newton-Raphson

L'idée de la méthode de Newton-Raphson pour résoudre l'équation $f(x) = 0$ est de remplacer au voisinage de la solution α de cette équation (en supposant que cette solution existe et est unique) le graphe de f par celui de la tangente à ce graphe en un point. Ce qui peut être illustré par la figure 8.6.

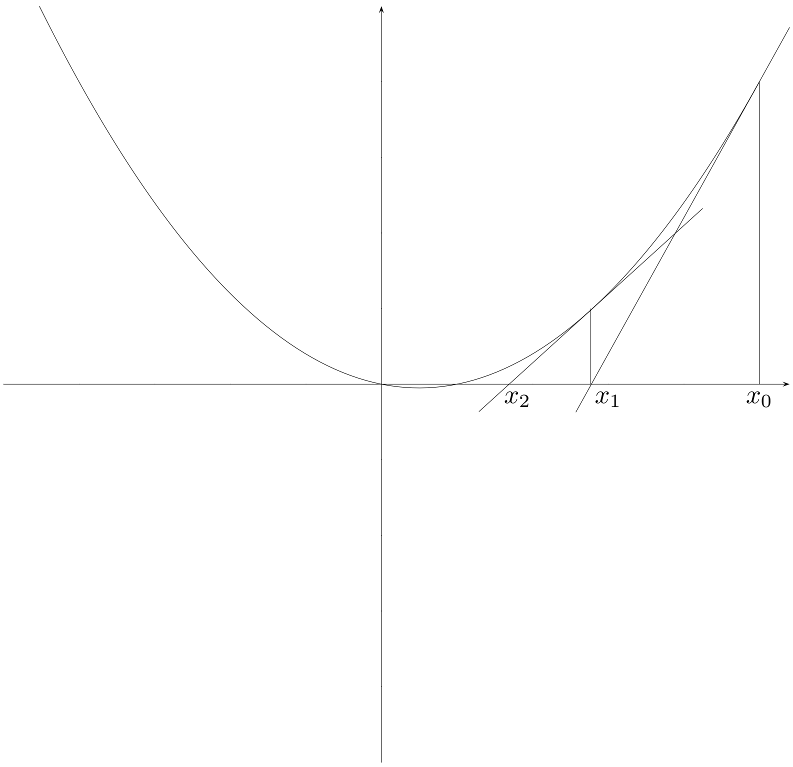

FIGURE 8.6 –

L'équation de la tangente au graphe de f en $(x, f(x))$ est :

$$Y = f(x) + f'(x)(X - x)$$

Si $f'(x) \neq 0$, le point d'intersection de cette tangente avec l'axe des x est alors donné par :

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

La résolution de l'équation $f(x) = 0$ est ainsi ramené à la résolution du problème de point fixe $g(x) = x$, ce qui revient à considérer la suite définie par :

$$\begin{cases} x_0 \in I \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)} \quad (n \geq 0) \end{cases}$$

Cette méthode s'applique si α est racine simple de f, c'est-à-dire si $f(\alpha) = 0$ et $f'(\alpha) \neq 0$. Dans ce cas on aura $f'(x) \neq 0$ pour x voisin de la racine α. De manière plus précise, une condition suffisante de convergence est donnée par le théorème qui suit.

Théorème 8.11 *Soient $f \in \mathcal{C}^2(I, \mathbb{R})$ et $\alpha \in \overset{\circ}{I}$ tel que $f(\alpha) = 0$ et $f'(\alpha) \neq 0$. Il existe un réel $\eta > 0$ tel que pour tout $x_0 \in [\alpha - \eta, \alpha + \eta]$ la suite $(x_n)_{n\in\mathbb{N}}$ définie par $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)}$ converge vers α.*

Démonstration. La fonction f' étant continue sur I avec $f'(\alpha) \neq 0$, on peut trouver un réel $\delta > 0$ tel que $J = [\alpha - \delta, \alpha + \delta] \subset I$ et $f'(x) \neq 0$ pour tout $x \in J$. On peut donc définir la fonction g sur J par :

$$\forall x \in J,\ g(x) = x - \frac{f(x)}{f'(x)}$$

Si on suppose de plus que f est de classe $\mathcal{C}^2$ alors g est de classe $\mathcal{C}^1$ avec :

$$\forall x \in J,\ g'(x) = \frac{f(x) f''(x)}{(f'(x))^2}$$

De $f(\alpha) = 0$ on déduit que $g'(\alpha) = 0$ et le théorème 8.8 appliqué à la fonction g permet alors de conclure. ∎

Le théorème qui suit donne une majoration de l'erreur d'approximation.

Théorème 8.12 *Soit $f \in \mathcal{C}^2([a, b], \mathbb{R})$ telle que :*

$$f(a) f(b) < 0,\ \forall x \in [a, b],\ f'(x) \neq 0,\ f''(x) \neq 0$$

Pour tout x_0 dans $[a, b]$ tel que $f(x_0) f''(x_0) > 0$ (règle de Fourier), on peut définir la suite $(x_n)_{n\in\mathbb{N}}$ de points de $[a, b]$ par :

$$\forall n \geq 0,\ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

cette suite est monotone et converge vers l'unique solution $\alpha \in]a, b[$ de $f(x) = 0$.

Une majoration de l'erreur est donnée par :

$$|x_n - \alpha| \leq |x_0 - \alpha|^{2^n} \left(\frac{M_2}{2m_1}\right)^{2^n - 1}$$

où :

$$m_1 = \inf_{x\in[a,b]} |f'(x)|,\ M_2 = \sup_{x\in[a,b]} |f''(x)|$$

Démonstration. On peut supposer que $f''(x) > 0$ pour tout $x \in [a, b]$. L'hypothèse $f(x_0) f''(x_0) > 0$ se traduit alors par $f(x_0) > 0$.

Le théorème des valeurs intermédiaires nous garantit l'existence et l'unicité de la solution $\alpha \in]a, b[$ de $f(x) = 0$.

L'hypothèse $f'(x) \neq 0$ pour tout $x \in [a, b]$ nous permet de définir la fonction g par :

$$\forall x \in [a, b],\ g(x) = x - \frac{f(x)}{f'(x)}$$

Avec la formule de Taylor à l'ordre 2 appliquée à f entre x et α on a :

$$\forall x \in [a, b],\ g(x) - \alpha = (x - \alpha)^2 \frac{f''(c)}{2f'(x)}$$

où c est un réel compris entre x et α.

On distingue alors deux cas.

- *Premier cas* : $f'(x) > 0$ pour tout $x \in [a,b]$. C'est-à-dire que f est strictement croissante. Dans ce cas on a $g(x) \geq \alpha$ pour tout $x \in [a,b]$.
  De plus pour $x \in [\alpha, b]$, on a $f(x) \geq f(\alpha) = 0$ et $g(x) = x - \dfrac{f(x)}{f'(x)} \leq x$.
  On a donc $g([\alpha,b]) \subset [\alpha,b]$ et la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie si on choisit $x_0 \in [\alpha,b]$, ce qui équivaut à dire que $f(x_0) > 0$, soit $f(x_0)f''(x_0) > 0$.
  Dans ces conditions, on a :
  $$\forall n \geq 0,\ x_n \geq \alpha \text{ et } x_{n+1} = g(x_n) \leq x_n$$
  C'est-à-dire que la suite $(x_n)_{n\in\mathbb{N}}$ est décroissante et minorée, elle est donc convergente et c'est nécessairement vers la solution $\alpha \in]a,b[$ de $f(x) = 0$.
- *Deuxième cas* : $f'(x) < 0$ pour tout $x \in [a,b]$. C'est-à-dire que f est strictement décroissante.
  Dans ce cas on a alors $g(x) \leq \alpha$ pour tout $x \in [a,b]$.
  De plus pour $x \in [a,\alpha]$, on a $f(x) \geq f(\alpha) = 0$ et $g(x) = x - \dfrac{f(x)}{f'(x)} \geq x$.
  On a donc $g([a,\alpha]) \subset [a,\alpha]$ et la suite $(x_n)_{n\in\mathbb{N}}$ est bien définie si on choisit $x_0 \in [a,\alpha]$, ce qui équivaut à dire que $f(x_0) > 0$, soit $f(x_0)f''(x_0) > 0$.
  Dans ces conditions, on a :
  $$\forall n \geq 0,\ x_n \leq \alpha \text{ et } x_{n+1} = g(x_n) \geq x_n$$
  C'est-à-dire que la suite $(x_n)_{n\in\mathbb{N}}$ est croissante et majorée, elle est donc convergente et c'est nécessairement vers la solution $\alpha \in]a,b[$ de $f(x) = 0$.
  La majoration de l'erreur d'approximation se déduit immédiatement des inégalités :
  $$\forall n \geq 0,\ |x_{n+1} - \alpha| = |x_n - \alpha|^2 \left| \frac{f''(c_n)}{2f'(x_n)} \right| \leq |x_n - \alpha|^2 \frac{M_2}{2m_1}$$

■

Exemple 8.1 *Le calcul approché de la racine $p^{\text{ème}}$ ($p \in \mathbb{N} - \{0,1\}$) d'un réel $a > 0$ peut se faire en utilisant la méthode de Newton. En effet $\alpha = \sqrt[p]{a}$ est racine de $f(x) = x^p - a$ et la méthode de Newton consiste à utiliser la suite définie par :*

$$\begin{cases} x_0 > 0 \\ x_{n+1} = \dfrac{1}{p}\left\{(p-1)x_n + \dfrac{a}{x_n^{p-1}}\right\} & (n \geq 0) \end{cases}$$

et on a $\sqrt[p]{a} = \lim\limits_{n\to+\infty} x_n$.
Pour $n \geq 0$ on a :

$$\begin{aligned} x_{n+1} - \alpha &= (x_n - \alpha)^2 \frac{f''(c_n)}{2f'(x_n)} \\ &= (x_n - \alpha)^2 \frac{(p-1)c_n^{p-2}}{2x_n^{p-1}} \end{aligned}$$

et par récurrence $x_n > \alpha$ pour $n \geq 1$. De plus $x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)} < x_n$. La suite $(x_n)_{n\in\mathbb{N}}$ est donc décroissante minorée par α.

Remarque 8.15 *Dans la pratique, la méthode de Newton est très efficace si le point de départ est choisi proche de la solution α. Dans le cas contraire la suite peut diverger ou converger vers une autre racine. Ce dernier cas peut se produire si $f''(x)$ change de signe au voisinage de α.*

Remarque 8.16 *Dans le cas où la racine α est multiple d'ordre p, on applique la méthode de Newton à la fonction $\sqrt[p]{f(x)}$, ce qui revient à considérer la suite des approximations successives définie par :*

$$\forall n \geq 0,\ x_{n+1} = x_n - p\frac{f(x_n)}{f'(x_n)}$$

Remarque 8.17 *Cette méthode de Newton se généralise au cas des systèmes d'équations non linéaires.*

La programmation de la méthode de Newton peut se faire en utilisant les approximations des nombres dérivés par différence finie centrée, soit :

$$f'(x) \backsimeq \frac{f(x+h) - f(x-h)}{2h}$$

l'erreur d'approximation étant de l'ordre de h^2. Comme test d'arrêt, on peut prendre $|f(x)| \leq \varepsilon$, où ε est une précision donnée.

On peut écrire la programmation structurée qui suit.

Entrée f : *Fonction ;* x_0, ε : *Réel ;* $MaxIter$: Entier ; *Sortie* x : *Réel ;*

Début

$x = x_0$; $k = 0$; $h = 10^{-3}$;

Répéter

$k = k + 1$;

$df = \dfrac{f(x+h) - f(x-h)}{2h}$;

$\delta = \dfrac{f(x)}{df}$;

$x = x - \delta$;

Jusqu'à $(k = MaxIter)$ *ou* $(|\delta| < \varepsilon)$

Fin ;

8.6 Résolution approchée des équations algébriques

8.6.1 Méthode de Bernoulli

Soit $P(x) = x^p + \sum_{k=0}^{p-1} a_k x^k$ un polynôme à coefficients réels admettant $p \geq 2$ racines réelles ou complexes $\lambda_1, ..., \lambda_p$ telles que $|\lambda_1| > |\lambda_2| > ... > |\lambda_n|$. On définit la suite $(x_n)_{n\in\mathbb{N}}$ par :

$$x_n + \sum_{j=1}^{p} a_{p-j} x_{n-j} = 0 \ \ (n \geq p)$$

les valeurs initiales $x_0, ..., x_{p-1}$ étant des réels arbitraires.

Dans notre situation, il existe des coefficients complexes $\alpha_1, \cdots, \alpha_p$ uniquement déterminés par les valeurs initiales tels que, pour tout $n \in \mathbb{N}$, on ait $x_n = \sum_{j=1}^{p} \alpha_j \lambda_j^n$.

Théorème 8.13 *Dans le cas où $\alpha_1 \neq 0$, il existe un entier n_0 tel que $x_n \neq 0$ pour tout $n \geq n_0$ et la suite $\left(\frac{x_{n+1}}{x_n}\right)_{n\in\mathbb{N}}$ converge vers λ_1.*

Démonstration. Si $\alpha_1 \neq 0$, on peut alors écrire, pour tout $n \in \mathbb{N}$:

$$x_n = \alpha_1 \lambda_1^n \left(1 + \sum_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^n\right)$$

avec $\lim\limits_{n\to+\infty} \left(\frac{\lambda_j}{\lambda_1}\right)^n = 0$ pour tout j compris entre 2 et p puisque dans ce cas, on a $\left|\frac{\lambda_j}{\lambda_1}\right| < 1$. On a donc $x_n \underset{n\to+\infty}{\backsim} \alpha_1\lambda_1^n$ avec $\alpha_1\lambda_1^n \neq 0$ pour tout $n \in \mathbb{N}$, il existe donc un entier n_0 tel que $x_n \neq 0$ pour tout $n \geq n_0$ et :

$$\frac{x_{n+1}}{x_n} \underset{n\to+\infty}{\backsim} \frac{\alpha_1\lambda_1^{n+1}}{\alpha_1\lambda_1^n} = \lambda_1$$

■

Théorème 8.14 *Avec les hypothèses du théorème précédent, en supposant qu'il existe un entier $n_1 \geq n_0$ tel que $\frac{x_{n+1}}{x_n} \neq \lambda_1$ pour tout $n \geq n_1$, on a $\lim\limits_{n\to+\infty}\left(\dfrac{\frac{x_{n+1}}{x_n} - \lambda_1}{\frac{x_n}{x_{n-1}} - \lambda_1}\right) = \frac{\lambda_2}{\lambda_1}$.*

Démonstration. Pour $n \geq n_1$, on a :

$$y_{n+1} = \frac{x_{n+1}}{x_n} - \lambda_1 = \lambda_1 \left(\frac{1 + \sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n+1}}{1 + \sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n}} - 1\right) = \lambda_1 \left(\frac{\sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n}\left(\frac{\lambda_j}{\lambda_1} - 1\right)}{1 + \sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n}}\right)$$

et :

$$\frac{y_{n+1}}{y_n} = \frac{\sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n}\left(\frac{\lambda_j}{\lambda_1} - 1\right)}{\sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n-1}\left(\frac{\lambda_j}{\lambda_1} - 1\right)} \frac{1 + \sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n-1}}{1 + \sum\limits_{j=2}^{p} \alpha_j \left(\frac{\lambda_j}{\lambda_1}\right)^{n}} \underset{n\to+\infty}{\backsim} \frac{\lambda_2}{\lambda_1} = \mu_1$$

■

Le théorème précédent, nous dit que l'erreur d'approximation et $\left|\frac{x_{k+1}}{x_k} - \lambda_1\right|$ est de l'ordre de $\left|\frac{\lambda_2}{\lambda_1}\right|^n$.

Pour déterminer les autres racines, on remplace le polynôme P par le polynôme $P_1(x) = \frac{P(x)}{x - \lambda_1}$ (méthode de déflation).

8.6.2 Méthode de Newton-Maehly

Dans le cas où f est une fonction polynomiale à coefficients réels dont toutes les racines sont réelles et distinctes, on est presque toujours assuré de la convergence de la méthode de Newton. On a plus précisément le résultat suivant due à C. Masse en 1984 que nous admettons.

Théorème 8.15 *Soit f une fonction polynomiale de degré $n \geq 1$ à coefficients réels dont toutes les racines sont réelles et distinctes et $\mathcal{E}$ l'ensemble des réels x_0 pour lesquels la méthode de Newton associée à f de premier terme x_0 diverge.*
Si $n \leq 2$, alors $\mathcal{E}$ est vide ou réduit à un point.
Si $n = 3$, alors $\mathcal{E}$ est dénombrable.
Si $n \geq 4$ alors $\mathcal{E}$ est non dénombrable de mesure de Lebesgue nulle.

Dans ce paragraphe, on considère le cas où f est un polynôme de degré $p \geq 2$. On le notera :

$$P(x) = \sum_{k=0}^{p} a_k x^k$$

avec $a_p \neq 0$.

Le théorème de D'Alembert-Gauss nous dit que l'équation $P(x) = 0$ admet au plus p solutions distinctes dans $\mathbb{C}$. Mais ce théorème n'est pas constructif.

On suppose que toutes les racines du polynôme P sont réelles. C'est le cas, par exemple, pour les polynômes orthogonaux classiques ou pour le polynôme caractéristique d'une matrice symétrique.

La méthode de Newton-Maehly permet d'obtenir toutes les racines de P avec leur multiplicité.

On note $\lambda_1 > \lambda_2 > \cdots > \lambda_r$ les racines distinctes de P. La racine λ_j étant de multiplicité $m_j \geq 1$ pour tout $j = 1, 2, \cdots, r$ avec $\sum\limits_{j=1}^{r} m_j = p$.

Lemme 8.6 *Le polynôme dérivé P' admet les $\lambda_1, \lambda_2, ..., \lambda_r$ pour racines de multiplicités respectives $m_1 - 1, m_2 - 1, \cdots, m_p - 1$ (une multiplicité nulle signifie que λ_j n'est pas racine de P') et des racines simples $\mu_j \in]\lambda_{j+1}, \lambda_j[$ avec $1 \leq j \leq r-1$. En particulier les racines de P' sont réelles et contenues dans l'intervalle $[\lambda_r, \lambda_1]$.*

Démonstration. Si, pour $1 \leq j \leq r$, on a $m_j \geq 2$, alors λ_j est racine d'ordre $m_j - 1$ de P'. On a donc ainsi $\sum\limits_{j=1}^{r} (m_j - 1) = p - r$ racines réelles pour P'.

D'autre part, le théorème de Rolle nous dit que pour $1 \leq j \leq r-1$ il existe $\mu_j \in]\lambda_{j+1}, \lambda_j[$ tel que $P'(\mu_j) = 0$, ce qui donne $r - 1$ racines réelles supplémentaires pour P'. On a donc un total de $p - 1$ racines réelles pour P' et les μ_j sont nécessairement simples. ■

Lemme 8.7 *Si $x > \lambda_1$, on a alors $a_p P^{(k)}(x) > 0$ pour tout $k = 0, 1, \cdots, p$.*

Démonstration. Supposons que $a_p > 0$.

Le lemme 8.6 nous dit que, pour $0 \leq k \leq p-1$, les racines de $P^{(k)}$ sont dans l'intervalle $[\lambda_r, \lambda_1]$. On en déduit donc que $P^{(k)}$ est de signe constant sur $]\lambda_1, +\infty[$. Soit avec $a_p > 0$, $P^{(k)}(x) > 0$ pour $x > \lambda_1$ et $k = 0, 1, \cdots, p-1$.

Pour $k = p$, on a $P^{(p)}(x) = p! a_p > 0$. ■

La méthode de Newton nous donne un algorithme de calcul de la plus grande racine λ_1.

Théorème 8.16 *Pour tout $x_0 > \lambda_1$, on peut définir la suite $(x_n)_{n \in \mathbb{N}}$ par :*

$$\forall n \geq 0,\ x_{n+1} = x_n - \frac{P(x_n)}{P'(x_n)}$$

et elle converge en décroissant vers λ_1. Une majoration de l'erreur est donnée par :

$$\forall n \geq 0,\ 0 \leq x_n - \lambda_1 \leq \left(1 - \frac{1}{p}\right)^n (x_0 - \lambda_1)$$

Démonstration. La fonction g définie sur $[\lambda_1, +\infty[$ par :

$$\forall x \geq \lambda_1,\ g(x) = \begin{cases} \lambda_1 \text{ si } x = \lambda_1 \\ x - \dfrac{P(x)}{P'(x)} \text{ si } x > \lambda_1 \end{cases}$$

est indéfiniment dérivable.

Pour $x > \lambda_1$, on a $\dfrac{P(x)}{P'(x)} > 0$ (lemme 8.7) et $g(x) < x$.

Avec la formule de Taylor à l'ordre 2, on a :

$$\forall x > \lambda_1,\ g(x) - \lambda_1 = (x - \lambda_1)^2 \frac{P''(c_x)}{2P'(x)}$$

où c_x est un réel strictement compris entre x et λ_1. Et avec $\dfrac{P''(c_x)}{P'(x)} > 0$ (lemme 8.7), on déduit que $g(x) > \lambda_1$ pour tout $x > \lambda_1$, c'est-à-dire que $g([\lambda_1, +\infty[) \subset [\lambda_1, +\infty[$ et l'orbite $(x_n)_{n\in\mathbb{N}}$ est bien définie.

On a donc :

$$\forall n \geq 0,\ \lambda_1 < x_{n+1} < x_n$$

c'est-à-dire que la suite $(x_n)_{n\in\mathbb{N}}$ est strictement décroissante et minorée. Elle converge donc nécessairement vers λ_1.

Pour tout $x > \lambda_1$, on a :

$$1 - g'(x) = \frac{P'(x)^2 - P(x)P''(x)}{P'(x)^2} = \frac{-\left(\dfrac{P'}{P}\right)'(x)}{\left(\dfrac{P'}{P}\right)^2(x)}$$

c'est-à-dire avec $P(x) = a_p \prod\limits_{j=1}^{r} (x - \lambda_j)^{m_j}$ et $P'(x) = a_p \sum\limits_{j=1}^{r} m_j \dfrac{P(x)}{x - \lambda_j}$:

$$1 - g'(x) = \frac{\sum\limits_{j=1}^{r} \dfrac{m_j}{(x-\lambda_j)^2}}{\left(\sum\limits_{j=1}^{r} \dfrac{m_j}{x - \lambda_j}\right)^2}$$

Avec l'inégalité de Cauchy-Schwarz, on peut écrire :

$$\begin{aligned} \left(\sum_{j=1}^{r} \frac{m_j}{x-\lambda_j}\right)^2 &= \left(\sum_{j=1}^{r} \frac{\sqrt{m_j}}{x-\lambda_j}\sqrt{m_j}\right)^2 \\ &\leq \sum_{j=1}^{r} \frac{m_j}{(x-\lambda_j)^2} \sum_{j=1}^{r} m_j = p \sum_{j=1}^{r} \frac{m_j}{(x-\lambda_j)^2} \end{aligned}$$

Il en résulte que :

$$\forall x > \lambda_1,\ 0 < g'(x) \leq 1 - \frac{1}{p}$$

Et avec le théorème des accroissements finis, on a :

$$\forall n \geq 0,\ 0 \leq x_{n+1} - \lambda_1 = g(x_n) - g(\lambda_1) \leq \left(1 - \frac{1}{p}\right)(x_n - \lambda_1)$$

ce qui entraîne par récurrence :

$$\forall n \geq 0,\ 0 \leq x_n - \lambda_1 \leq \left(1 - \frac{1}{p}\right)^n (x_0 - \lambda_1)$$

■

Remarque 8.18 *On reprend les notations de la démonstration du théorème précédent et on note :*

$$M_1 = \sup_{x \in [\lambda_1, x_0]} |g'(x)|$$

En utilisant le théorème des accroissements finis, on a :

$$\forall n \geq 0,\ 0 \leq x_n - \lambda_1 \leq (M_1)^n (x_0 - \lambda_1)$$

En tenant compte de $g'(x) = \dfrac{P(x)P''(x)}{P'(x)^2}$*, on en déduit que, dans le cas où* λ_1 *est une racine simple de* P*, on a* $g'(\lambda_1) = 0$ *et la convergence est d'autant plus rapide que* x_0 *est proche de* λ_1*.*

De manière plus générale, on a $g'(\lambda_1) = \lim\limits_{x \to \lambda_1^+} g'(x) = 1 - \dfrac{1}{m_1}$*, ce qui résulte de :*

$$1 - g'(x) = \frac{\sum\limits_{j=1}^{r} \dfrac{m_j}{(x - \lambda_j)^2}}{\left(\sum\limits_{j=1}^{r} \dfrac{m_j}{x - \lambda_j}\right)^2} = \frac{1}{m_1} \frac{1 + (x-\lambda_1)^2 \sum\limits_{j=2}^{r} \dfrac{m_j}{m_1} \dfrac{1}{(x-\lambda_j)^2}}{\left(1 + (x - \lambda_1) \sum\limits_{j=2}^{r} \dfrac{m_j}{m_1} \dfrac{1}{x - \lambda_j}\right)^2}$$

pour $x > \lambda_1$*.*

Remarque 8.19 *Pour appliquer ce théorème, il faut disposer de critères pratiques permettant de choisir* $x_0 > \lambda_1$*. De tels critères sont donnés par le lemme qui suit.*

Lemme 8.8 *Soit* $P(x) = \sum\limits_{k=0}^{p} a_k x^k \in \mathbb{C}[x]$ *avec* $a_p \neq 0$*. Pour toute racine* λ *de* P*, on a :*

$$|\lambda| \leq \max \left\{1, \sum_{i=0}^{p-1} \frac{|a_i|}{|a_p|}\right\} \tag{8.2}$$

$$|\lambda| \leq \max_{1 \leq i \leq p-1} \left\{\frac{|a_0|}{|a_p|}, 1 + \frac{|a_i|}{|a_p|}\right\} \tag{8.3}$$

Démonstration. Pour toute racine λ de P, on a $P(\lambda) = 0$ et :

$$|a_p| |\lambda|^p = \left|\sum_{i=0}^{p-1} a_i \lambda^i\right| \leq \sum_{i=0}^{p-1} |a_i| |\lambda|^i$$

et pour $|\lambda| > 1$, on obtient :

$$|a_p|\,|\lambda| \leq \sum_{i=0}^{p-1} \frac{|a_i|}{|\lambda|^{p-i}} \leq \sum_{i=0}^{p-1} |a_i|$$

Pour la deuxième inégalité, c'est un peu plus délicat.

Le polynôme $Q = \dfrac{1}{a_p} P = X^p + \sum\limits_{k=0}^{p-1} \dfrac{a_k}{a_p} x^k$ est le polynôme caractéristique de sa matrice compagnon :

$$C(Q) = \begin{pmatrix} 0 & \cdots & 0 & b_0 \\ 1 & \ddots & \vdots & b_1 \\ \vdots & \ddots & 0 & \vdots \\ 0 & \cdots & 1 & b_{p-1} \end{pmatrix}$$

où on a noté $b_k = -\dfrac{a_k}{a_p}$ pour k compris entre 0 et $p-1$, le polynôme caractéristique d'une matrice étant choisi unitaire et les λ_k sont les valeurs propres de $C(Q)$. En désignant par $\rho(C(Q)) = \max\limits_{1 \leq k \leq p} |\lambda_k|$ le rayon spectral de $C(Q)$, on sait que :

$$\rho(C(Q)) \leq \|C(Q)\|_\infty = \max_{1 \leq i \leq p-1} \{|b_0|, 1 + |b_i|\} = \max_{1 \leq i \leq p-1} \left\{ \frac{|a_0|}{|a_p|}, 1 + \frac{|a_i|}{|a_p|} \right\}$$

■

Le calcul des autres racines de P peut se faire en appliquant ce qui précède à la fonction $P_1(x) = \dfrac{P(x)}{(x - \lambda_1)}$ et en remarquant que :

$$\frac{P_1(x)}{P_1'(x)} = \frac{P(x)}{P'(x) - \dfrac{P(x)}{x - \lambda_1}}$$

La limite de la suite $(x_n)_{n \in \mathbb{N}}$ correspondante sera encore λ_1 si $m_1 > 1$ et λ_2 si $m_1 = 1$.

De manière plus générale, on désigne par $\xi_1 \geq \xi_2 \geq ... \geq \xi_p$ les p racines réelles distinctes ou confondues de P et on suppose qu'on a obtenu des valeurs approchées des $m-1$ premières racines $\xi_1 \geq ... \geq \xi_{m-1}$, où m est compris entre 2 et p. On se propose alors d'obtenir des valeurs approchées de ξ_m. Pour ce faire on remplace le polynôme P par le polynôme :

$$P_{m-1}(x) = \frac{P(x)}{(x - \xi_1) ... (x - \xi_{m-1})}$$

et on lui associe la fonction $g_{m-1} :]\xi_m, +\infty[\to R$ définie par :

$$\forall x > \xi_m,\ g_{m-1}(x) = x - \frac{P_{m-1}(x)}{P_{m-1}'(x)}$$

Lemme 8.9 *Avec les notations qui précèdent, la suite* $\left(x_k^{(m)}\right)_{k \in \mathbb{N}}$ *définie par :*

$$\begin{cases} x_0^{(m)} = b_m \\ \forall k \geq 0,\ x_{k+1}^{(m)} = g_{m-1}\left(x_k^{(m)}\right) \end{cases}$$

où b_m est un réel strictement supérieur à ξ_m converge vers ξ_m.

Démonstration. On a :

$$P_{m-1}(x) = a_p \prod_{i=m}^{p} (x - \xi_i)$$

Pour $m = p$, on a $P_{p-1}(x) = a_p (x - \xi_p)$, $g_{p-1}(x) = \xi_p$ et la suite $\left(x_k^{(p)}\right)_{k \in \mathbb{N}}$ est stationnaire sur ξ_p.

Pour $m < n$, comme pour le théorème 8.2, on montre que la suite $\left(x_k^{(m)}\right)_{k \in \mathbb{N}}$ converge en décroissant strictement vers ξ_m. ■

Lemme 8.10 *Pour tout $x > \xi_m$, on a :*

$$\frac{P_{m-1}(x)}{P'_{m-1}(x)} = \frac{P(x)}{P'(x) - P(x) \sum_{j=1}^{m-1} \frac{1}{x - \xi_j}}$$

Démonstration. On a :

$$\forall x \in \mathbb{R},\ P_{m-1}(x) \prod_{k=1}^{m-1} (x - \xi_k) = P(x)$$

et en dérivant :

$$\forall x \in \mathbb{R},\ P'_{m-1}(x) \prod_{k=1}^{m-1} (x - \xi_k) + P_{m-1}(x) \sum_{k=1}^{m-1} \prod_{\substack{j=1 \\ j \neq k}}^{m-1} (x - \xi_j) = P'(x)$$

En écrivant, pour tout réel x :

$$\prod_{k=1}^{m-1} (x - \xi_k) = \frac{P(x)}{P_{m-1}(x)},\ P_{m-1}(x) \prod_{\substack{j=1 \\ j \neq k}}^{m-1} (x - \xi_j) = \frac{P(x)}{x - \xi_k}$$

(les fonctions de droite dans ces égalités se prolonge par continuité à $\mathbb{R}$), on obtient :

$$\forall x \in \mathbb{R},\ P'_{m-1}(x) \frac{P(x)}{P_{m-1}(x)} + \sum_{k=1}^{m-1} \frac{P(x)}{x - \xi_k} = P'(x)$$

Les fonctions P_{m-1} et P'_{m-1} ne s'annulant jamais sur $]\rho_m, +\infty[$, on peut écrire :

$$\forall x > \xi_m,\ \frac{P'_{m-1}(x)}{P_{m-1}(x)} P(x) = P'(x) - \sum_{k=1}^{m-1} \frac{P(x)}{x - \xi_k}$$

ou encore :

$$\forall x > \rho_m,\ \frac{P_{m-1}(x)}{P'_{m-1}(x)} = \frac{P(x)}{P'(x) - \sum_{j=1}^{m-1} \frac{P(x)}{x - \xi_j}}.$$

■

L'intérêt de cette écriture est qu'elle permet la programmation de la suite $\left(x_k^{(m)}\right)_{k \in \mathbb{N}}$ en utilisant la donnée des coefficients de P et les approximations obtenues aux étapes précédentes des $m - 1$ premières racines.

Remarque 8.20 *Pour accélérer la convergence, on considère tout d'abord la suite* $\left(y_n^{(m)}\right)_{n\in\mathbb{N}}$ *définie par :*

$$\begin{cases} y_0^{(m)} = x_0 \\ y_{n+1}^{(m)} = y_n^{(m)} - 2\dfrac{P_{m-1}\left(y_n^{(m)}\right)}{P'_{m-1}\left(y_n^{(m)}\right)} \quad (n \geq 0) \end{cases}$$

On utilise cette suite tant que $y_{n+1}^{(m)} < y_n^{(m)}$ *(la suite est théoriquement croissante, mais dès qu'on est très proche de la solution cela peut ne plus être vrai à cause des erreurs d'arrondis), puis dès que* $y_{n+1}^{(m)} \geq y_n^{(m)}$ *pour un indice* n_0*, on utilise la suite* $\left(x_n^{(m)}\right)_{n\in\mathbb{N}}$ *avec* $y_{n_0}^{(m)}$ *comme point de départ.*

On a donc en définitive la programmation structurée suivante.

Entrée p : Entier ; P : Polynôme ; Sortie ρ : Vecteur ;

Début

$\alpha_0 = \max\limits_{1\leq i\leq p-1}\left\{\dfrac{|a_0|}{|a_p|}, 1 + \dfrac{|a_i|}{|a_p|}\right\}$;

Pour m Allant de 1 à p Faire

Début

$y_0 = \alpha_0$; $U = P(y_0)$; $y_1 = y_0 - 1$; $y = y_0$;

Tant que ($|U| > \varepsilon$) *et* ($y_1 < y$) *faire*

Début

$S = 0$;

Pour j Allant de 1 à m − 1 Faire

Début

$S = S + \dfrac{1}{y_0 - \rho_j}$;

Fin ;

$V = P'(y_0)$;

$y_1 = y_0 - 2\dfrac{U}{V - US}$;

$y = y_0$; $y_0 = y_1$; $U = P(y_0)$;

Fin ;

$x_0 = y_0$; $U = P(y_0)$;

Tant que ($|U| > \varepsilon$) *faire*

Début

$S = 0$;

Pour j Allant de 1 à m − 1 Faire

Début

$S = S + \dfrac{1}{x_0 - \rho_j}$;

Fin ;

$V = P'(x_0)$;

$x_1 = x_0 - \dfrac{U}{V - US}$;

$x_0 = x_1$; $U = P(x_0)$;

Fin ;

$\rho_m = x_0$;

Fin ;

Fin ;

8.6.3 Méthode de Bairstow

Cette méthode utilise la méthode de Newton-Raphson (supposée connue) pour résoudre un système non linéaire de deux équations à deux inconnues.

L'idée est d'écrire le polynôme $P(x) = \sum_{k=0}^{p} a_k x^k$ comme un produit de polynômes de degré au plus égal à 2 (on suppose que $p \geq 2$).

La première étape consiste à trouver des réels p et q tels que :

$$P(x) = Q(x)\left(x^2 + px + q\right)$$

On itère ensuite le procédé.

Pour p, q réels donnés, on peut effectuer la division euclidienne :

$$P(x) = B(x)\left(x^2 + px + q\right) + rx + s$$

les coefficients $r = r(p,q)$ et $s = s(p,q)$ dépendants de p et q. Il s'agit donc de déterminer p et q solutions du système de deux équations à deux inconnues :

$$\left\{\begin{array}{l} r(p,q) = 0 \\ s(p,q) = 0 \end{array}\right. \tag{8.4}$$

Le premier problème est donc de calculer s et s en fonction de p, q et des coefficients du polynôme P. Ensuite le système (8.4) sera résolu par la méthode de Newton-Raphson.

Pour ce faire, on pose :

$$\left\{\begin{array}{l} B(x) = \sum_{k=0}^{p-2} b_k x^k \\ R(x) = rx + s = b_{-1}(x+p) + b_{-2} \end{array}\right.$$

On va d'abord calculer les b_i pour $i = -2, -1, ..., p-2$.

Par identification, on a :

$$\left\{\begin{array}{l} a_p = b_{p-2} \\ a_{p-1} = b_{p-3} + pb_{p-2} \\ a_i = b_{i-2} + pb_{i-1} + qb_{i-2} \end{array}\right. \quad (i = p-2, \cdots, 0)$$

d'où :

$$\left\{\begin{array}{l} b_p = b_{p-1} = 0 \\ b_{i-2} = a_i - pb_{i-1} - qb_i \end{array}\right. \quad (i = p, \cdots, 0) \tag{8.5}$$

En particulier on a b_{-1} etb_{-2}, soit r et s.

Le deuxième problème est de calculer les dérivées partielles de r et s par rapport à p et q.

On a :

$$\left\{\begin{array}{l} r(p,q) = b_{-1}(p,q) \\ s(p,q) = pb_{-1}(p,q) + b_{-2}(p,q) \end{array}\right.$$

donc :

$$\left\{\begin{array}{l} \dfrac{\partial r}{\partial p} = \dfrac{\partial b_{-1}}{\partial p} \\ \dfrac{\partial r}{\partial q} = \dfrac{\partial b_{-1}}{\partial q} \\ \dfrac{\partial s}{\partial p} = b_{-1} + p\dfrac{\partial b_{-1}}{\partial p} + \dfrac{\partial b_{-2}}{\partial p} \\ \dfrac{\partial s}{\partial q} = p\dfrac{\partial b_{-1}}{\partial q} + \dfrac{\partial b_{-2}}{\partial q} \end{array}\right.$$

et tout revient à calculer les $\dfrac{\partial b_i}{\partial p}$ et $\dfrac{\partial b_i}{\partial q}$ pour $i=-2$ et $i=-1$. Ce qui se fait en dérivant les relations (8.5).

On obtient alors les relations de récurrence :

$$\left\{\begin{array}{l} \dfrac{\partial b_p}{\partial p}=\dfrac{\partial b_{p-1}}{\partial p}=\dfrac{\partial b_p}{\partial q}=\dfrac{\partial b_{p-1}}{\partial q}=0 \\ \dfrac{\partial b_{i-2}}{\partial p}=-b_{i-1}-p\dfrac{\partial b_{i-1}}{\partial p}-q\dfrac{\partial b_i}{\partial p} \\ \dfrac{\partial b_{i-2}}{\partial q}=-p\dfrac{\partial b_{i-1}}{\partial q}-b_i-q\dfrac{\partial b_i}{\partial q} \end{array}\right. \quad (i=p,\cdots,0)$$

En notant $c_i=-\dfrac{\partial b_{i-1}}{\partial p}$ et $d_i=-\dfrac{\partial b_{i-2}}{\partial q}$, on constate que ces deux suites vérifient la même relation de récurrence :

$$\left\{\begin{array}{l} u_{p+1}=u_p=0 \\ u_{i-1}=b_{i-1}(p,q)+pu_i+qu_{i+1} \end{array}\right. \quad (i=p,\cdots,0) \tag{8.6}$$

et en particulier $c_i=d_i$ pour tout i.

Le troisième problème est alors de résoudre le système linéaire intervenant dans la méthode de Newton-Raphson.

À l'étape $k+1$ du calcul, on a :

$$(p_{k+1},q_{k1})=(p_k,q_k)+(\delta p,\delta q)$$

où $(\delta p,\delta q)$ est solution du système linéaire de deux équations à deux inconnues :

$$d(r,s)(p_k,q_k)(\delta p,\delta q)=-(r(p_k,q_k),s(p_k,q_k))$$

soit :

$$\left\{\begin{array}{l} \dfrac{\partial b_{-1}}{\partial p}(p_k,q_k)\,\delta p+\dfrac{\partial b_{-1}}{\partial q}(p_k,q_k)\,\delta q=-b_{-1}(p_k,q_k) \\ \left(b_{-1}(p_k,q_k)+p_k\dfrac{\partial b_{-1}}{\partial p}(p_k,q_k)+\dfrac{\partial b_{-2}}{\partial p}(p_k,q_k)\right)\delta p+\left(p_k\dfrac{\partial b_{-1}}{\partial q}(p_k,q_k)+\dfrac{\partial b_{-2}}{\partial q}(p_k,q_k)\right)\delta q \\ =-\left(b_{-1}(p_k,q_k)\,p_k+b_{-2}(p_k,q_k)\right) \end{array}\right.$$

Avec les notations précédentes, ce système va s'écrire, en omettant les variables p_k,q_k :

$$\left\{\begin{array}{l} c_0\delta p+c_1\delta q=-b_{-1} \\ (b_{-1}+p_kc_0+c_{-1})\,\delta p+(p_kc_1+c_0)\,\delta q=-(b_{-1}p_k+b_{-2}) \end{array}\right.$$

et sa solution est :

$$\left\{\begin{array}{l} \delta p=\dfrac{c_0b_{-1}-c_1b_{-2}}{c_0^2-c_1c_{-1}+c_1b_{-1}} \\ \delta q=\dfrac{c_0b_{-2}+b_{-1}^2-b_{-1}c_{-1}}{c_0^2-c_1c_{-1}+c_1b_{-1}} \end{array}\right.$$

Pour la programmation, on écrit tout d'abord la procédure de division qui donne les coefficients p et q et le polynôme B tels que $P(x)=B(x)(x^2+px+q)+rx+s$.

PROCEDURE Division(Entrée p : Entier ; P : Polynôme ; ε : Réel ; Sortie B : Vecteur ; p,q : Réels) ;Début

Se donner p_0 et q_0 ;

$k = 0$;
$p = p_0$; $q = q_0$;
Répéter
Calcule_Coeff_bk(p, P, p, q, B) ; (Récurrence (8.5))
Calcule_Coeff_ck(p, B, p, q, C) ; (Récurrence (8.6)}
$D = c_0^2 - c_1c_{-1} + c_1b_{-1}$;
$U = c_0b_{-1} - c_1b_{-2}$;
$V = c_0b_{-2} + b_{-1}^2 - b_{-1}c_{-1}$;
$\delta p = \frac{U}{D}$; $\delta q = \delta p = \frac{V}{D}$;
$p = p + \delta p$; $q = q + \delta q$;
Jusqu'à $(|\delta p| < \varepsilon)$ et $(|\delta q| < \varepsilon)$;
Fin ;
Les procédures de calcul des coefficients b_k et c_k s'écrivant :
PROCEDURE Calcule_Coeff_bk(Entrée p : Entier ; P : Polynôme ; p, q : Réel ; Sortie B Vecteur) ;
Début
$b_p = 0$; $b_{p--1} = 0$;
Pour k allant de p à 0 faire$b_{k-2} = a_k - pb_{k--1} - qb_k$;
Fin ;
PROCEDURE Calcule_Coeff_ck(Entrée p : Entier ; B : Vecteur ; p, q : Réel ; Sortie C : Vecteur) ;
Début
$c_{p+1} = 0$; $c_p = 0$;
Pour k allant de p à 0 faire $c_{k-1} = b_{k-1} + pc_k + qc_{k+1}$;
Fin ;
La procédure principale peut alors s'écrire :
PROCEDURE Bairstow(Entrée p : Entier ; P : Polynôme ; Sortie x : VecteurComplexe) ;
Début
Pour i allant de 0 à p faire $a_i = \frac{a_i}{a_n}$;
$k = 0$;
Répéter
Division(p, P, B, p, q) ;
$p = p - 2$;
Pour i allant de 0 à p faire $a_i = b_i$;
RésolEquDegre2(p, q, x_1, x_2) ;
$k = k + 1$; $x_k = x_1$;
$k = k + 1$; $x_k = x_2$;
Jusqu'à $(p \leq 2)$;
Si $p = 2$ Alors
Début
RésolEquDegre2(a_1, a_0, x_1, x_2) ;
$k = k + 1$; $x_k = x_1$;
$k = k + 1$; $x_k = x_2$;
Fin ;
Si $p = 1$ Alors
Début
$k = k + 1$; $x_k = -\frac{a_0}{a_1}$;

        Fin ;
    Fin ;

Bibliographie

[1] H. Alzer. *A proof of the arithmetic mean-geometric mean inequality.* American Mathematical Monthly. (1996).

[2] J. M. Arnaudies, J. Lelong-Ferrand. *Cours de Mathématiques. Tomes 1 à 4.* Dunod (1974).

[3] E. Artin. *The Gamma function.* Holt, Rinehart and Winston, Inc. New York (1964).

[4] G. Auliac, J. Y. Caby. *Analyse pour le Capes et l'Agrégation interne.* Ellipses (2002).

[5] B. Balaguer. *La leçon d'analyse au Capes de Mathématiques.* Ellipses (1999).

[6] R. P. Boas. *A primer of real functions.* Carus mathematical monograph 13. Wiley (1960).

[7] R. Burlisch, J. Stoer — *Introduction to numerical analysis.* Springer-Verlag (1980).

[8] H. Cartan. *Théorie élémentaire des fonctions analytiques d'une ou plusieurs variables complexes.* Hermann (1961).

[9] A. Chambert-Loir, S. Fermigier. *Analyse 2.* Masson (1995).

[10] C. Deschamps, A. Warusfel. *Mathématiques tout en un. Série E. Ramis. Volumes 1 et 2.* Dunod. (1999).

[11] P. Doukhan, J. C. Sifre — *Cours d'analyse. Analyse réelle et intégration.* Dunod. (2001).

[12] A. Dufetel. *Analyse. Capes externe, agrégation interne de Mathématiques.* Vuibert (2011).

[13] J. M. Ferrard, H. Lemberg. *Mathématiques concrètes, illustrées par la TI-92 et la TI-89.* Springer (1998).

[14] B. Flannery, W. Press, S. Teukolsky, W. Vetterling — *Numerical recipes.* Cambridge University Press (1988).

[15] B. Gostiaux. *Cours de Mathématiques Spéciales. Volumes 1 à 4.* P. U. F. (1995).

[16] X. Gourdon. *Les Maths en tête. Analyse.* Ellipses (1994).

[17] A. Gramain. *Géométrie élémentaire.* Hermann (1997).

[18] G. H. Hardy. *A course of pure Mathematics.* Cambridge University Press (1952).

[19] Hardy, Littlewood, Polya. *Inequalities.* Cambridge Mathematical Library. (1999).

[20] W. J. Kaczor, M. T. Nowak. *Problems in mathematical analysis. Vol. I, II et III.* American Mathematical Society (2001).

[21] S. Lang. *Real analysis.* Addison-Wesley (1969).

[22] E. Landau. *Foundations of analysis.* Chelsea.

[23] E. Landau. *Differential and integral calculus.* Chelsea (1960).

[24] E. Leichtnam, X Schauer. *Exercices corrigés de Mathématiques, volume 3.* Ellipses (1982).

[25] F. Liret, D. Martinais. *Cours de mathématiques. Analyse 1-ère et 2-ème année. Algèbre 1-ère et 2-ème année.* Dunod.

[26] K. Madere. *Préparation à l'oral de l'agrégation. Développements d'analyse.* Ellipses (1997).

[27] K. Madere. *Préparation à l'oral de l'agrégation. Leçons d'analyse.* Ellipses (1997).

[28] D. Monasse. *La méthode de Newton et les racines des polynômes.* R M. S. (janvier 1987).

[29] I. Nourdin. *Agrégation de mathématiques. Leçons d'analyse, probabilités, algèbre et géométrie.* Dunod (2001).

[30] J. P. Nougier. *Méthodes de calcul numérique.* Masson.

[31] H. Queffelec. *Topologie.* Masson (1998).

[32] H. Queffelec, C. Zuily. *Éléments d'analyse pour l'agrégation.* Dunod (2002).

[33] Ramis, Deschamps, Odoux. *Analyse 2, exercices avec solutions.* Masson. (1972).

[34] J. P. Ramis, A. Warusfel. *Mathématiques tout en un pour la licence. Niveau L2.* Dunod. (2007).

[35] A.W. Roberts, D.E. Varberg. *Convex functions.* Academic Press (1973).

[36] J. E. Rombaldi. *Problèmes corrigés d'analyse numérique.* Masson (1996).

[37] J. E. Rombaldi. *Éléments d'analyse réelle.* EDP Sciences (2004).

[38] J. E. Rombaldi. *Interpolation et approximation.* Vuibert (2005).

[39] F. Rouviere. *Calcul différentiel.* Cassini (1999).

[40] W. Rudin. *Principes d'analyse mathématique.* Edisciences (1995).

Index