Dérivabilité et convexité

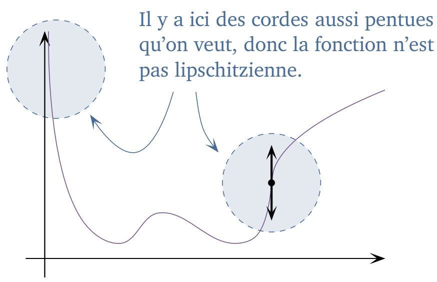
Dans ce chapitre, E et F sont des parties quelconques de $\mathbb{R}$, pas forcément des intervalles, et $\mathbb{K}$ est l'un des ensembles $\mathbb{R}$ ou $\mathbb{C}$. Quand on emploiera les notations $[a,b]$ ou $]a,b[$, il sera sous-entendu que a et b sont deux réels et que $a<b$.

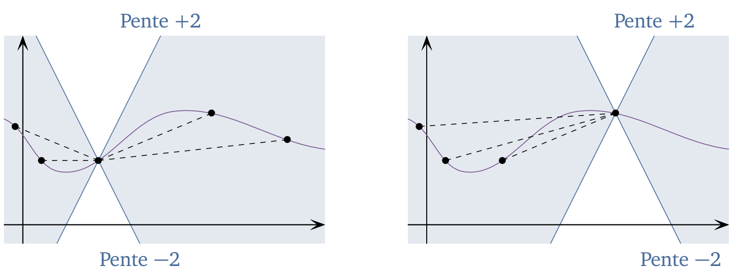
Ce chapitre est consacré en quelque sorte à la notion d'accroissement, i.e. à la comparaison des réels $x-y$ et $f(x)-f(y)$ pour toute fonction f ou encore aux taux d'accroissements $\dfrac{f(x)-f(y)}{x-y}$.

1 Lipschitzianité

Définition (Lipschitzianité) Soient $f : E \longrightarrow \mathbb{C}$ une fonction et $K \geqslant 0$. On dit que f est *K-lipschitzienne sur E* si :
$$\forall x,y \in E, \qquad \big|f(x)-f(y)\big| \leqslant K\,|x-y|.$$
On dit en outre que f est *contractante* si le réel K peut être choisi **strictement** inférieur à 1.

La fonction f est K-lipschitzienne sur E si $\left|\dfrac{f(x)-f(y)}{x-y}\right| \leqslant K$ pour tous $x,y \in E$ distincts, i.e. si les pentes des cordes de son graphe sont majorées par K en valeur absolue. Plus généralement, dire que f est lipschitzienne revient à dire que les pentes des cordes de son graphe sont bornées.





Les figures de gauche illustrent la 2-lipschitzianité de la fonction représentée. Les cordes issues de chaque point ont toutes un coefficient directeur compris entre -2 et 2.

Exemple La fonction $x \longmapsto |x|$ est 1-lipschitzienne sur $\mathbb{R}$ d'après l'inégalité triangulaire généralisée $\big||x|-|y|\big| \leqslant |x-y|$.

Exemple La fonction $x \overset{f}{\longmapsto} \dfrac{1}{x}$ est $\dfrac{1}{\varepsilon^2}$-lipschitzienne sur $[\varepsilon,+\infty[$ pour tout $\varepsilon>0$, mais elle n'est pas lipschitzienne sur $\mathbb{R}_+^*$.

Démonstration Pour tous $\varepsilon>0$ et $x,y \geqslant \varepsilon$: $\quad \big|f(x)-f(y)\big| = \left|\dfrac{1}{x}-\dfrac{1}{y}\right| = \dfrac{|x-y|}{xy} \leqslant \dfrac{|x-y|}{\varepsilon^2}$.

En revanche : $\quad \dfrac{f(2x)-f(x)}{2x-x} = -\dfrac{1}{2x^2} \underset{x\to 0}{\longrightarrow} -\infty, \quad$ donc l'ensemble des réels $\dfrac{f(x)-f(y)}{x-y}$ n'est pas borné, x et y décrivant $\mathbb{R}_+^*$ avec $x \neq y$, donc f n'est pas lipschitzienne sur $\mathbb{R}_+^*$.

Théorème (Lipschitzienne implique continue) Toute fonction lipschitzienne est continue.

Vous noterez en passant que la lipschitzianité est une propriété globale des fonctions, contrairement à la continuité qui est locale. La lipschitzianité est directement définie sur une partie de $\mathbb{R}$, il n'aurait pas grand sens de la définir en un point.

Démonstration Soit $f : E \longrightarrow \mathbb{C}$ une fonction K-lipschitzienne pour un certain $K \geqslant 0$. Pour tous $a,x \in E$: $\big|f(x)-f(a)\big| \leqslant K\,|x-a|, \quad$ donc $f(x) \underset{x\to a}{\longrightarrow} f(a)$ par encadrement, i.e. f est continue en a.

Exemple Pour toute partie non vide A de $\mathbb{R}$, la fonction $x \longmapsto d(x,A) = \inf\limits_{a\in A} |x-a|$ est continue sur $\mathbb{R}$.

Démonstration Pour tous $x,y \in \mathbb{R}$: $\quad \big|d(x,A)-d(y,A)\big| \leqslant |x-y| \quad$ comme nous l'avons déjà vu. Bref, la fonction $x \longmapsto d(x,A)$ est 1-lipschitzienne sur $\mathbb{R}$, donc continue.

2 Dérivabilité, point de vue local

2.1 Définitions

Définition (Dérivabilité en un point ou sur une partie de $\mathbb{R}$, tangente) Soient $f : E \longrightarrow \mathbb{C}$ une fonction et $a \in E$.

- **Dérivabilité :** On dit que f est *dérivable en a* si la fonction $x \longmapsto \dfrac{f(x)-f(a)}{x-a}$ possède une limite finie en a, notée $f'(a)$ le cas échéant et appelée le *nombre dérivé de f en a*.
- **Tangente :** Si f est dérivable en a, la droite d'équation $y = f(a) + f'(a)(x-a)$ est appelée la *tangente de f en a*. Et si $\dfrac{f(x)-f(a)}{x-a} \underset{x \to a}{\longrightarrow} \pm\infty$, la droite d'équation $x = a$ est appelée la *tangente de f en a*.

Si f est dérivable en a et si x est proche de a, alors $\dfrac{f(x)-f(a)}{x-a} \approx f'(a)$, donc $f(x) \approx f(a) + f'(a)(x-a)$. À défaut d'être rigoureux, c'est convaincant. La tangente de f en a est ainsi la droite la plus proche du graphe de f au voisinage de a.

Exemple Pour tout $n \in \mathbb{N}$, la fonction $x \longmapsto x^n$ est dérivable sur $\mathbb{R}$ de dérivée $x \longmapsto nx^{n-1}$.

Démonstration Pour tout $a \in \mathbb{R}$: $\dfrac{x^n - a^n}{x-a} = \displaystyle\sum_{k=0}^{n-1} a^k x^{n-k-1} \underset{x \to a}{\longrightarrow} \sum_{k=0}^{n-1} a^k a^{n-k-1} = na^{n-1}$.

Exemple La fonction valeur absolue $|\cdot|$ n'est pas dérivable en 0.

Démonstration Pour tout $x \in \mathbb{R}^*$: $\dfrac{|x|-|0|}{x-0} = \begin{cases} 1 & \text{si } x > 0 \\ -1 & \text{si } x < 0, \end{cases}$ donc la fonction $x \longmapsto \dfrac{|x|-|0|}{x-0}$ n'a pas de limite en 0 car $\dfrac{|x|-|0|}{x-0} \underset{x \to 0^+}{\longrightarrow} 1$ et $\dfrac{|x|-|0|}{x-0} \underset{x \to 0^-}{\longrightarrow} -1$.

Théorème (Dérivable implique continue) Soient $f : E \longrightarrow \mathbb{C}$ une fonction et $a \in E$. Si f est dérivable en a, alors f est continue en a.

Attention ! La réciproque est totalement fausse, pensez à la fonction valeur absolue en 0. C'est contre-intuitif, mais il existe même des fonctions qui sont continues sur $\mathbb{R}$ tout entier mais dérivables en aucun point.

Démonstration Si f est dérivable en a : $f(x) = \dfrac{f(x)-f(a)}{x-a} \times (x-a) + f(a) \underset{x \to a}{\longrightarrow} f'(a) \times 0 + f(a) = f(a)$, donc f est continue en a.

Le résultat suivant est la version dérivable d'un résultat analogue sur les limites.

Théorème (Caractérisation de la dérivabilité par les parties réelle et imaginaire) Soient $f : E \longrightarrow \mathbb{C}$ une fonction et $a \in E$.

f est dérivable en a si et seulement si $\mathrm{Re}(f)$ et $\mathrm{Im}(f)$ le sont.

De plus, le cas échéant : $f'(a) = \mathrm{Re}(f)'(a) + i\,\mathrm{Im}(f)'(a)$.

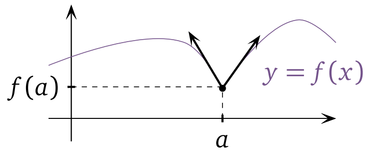
Définition-théorème (Dérivabilité à gauche/à droite en un point) Soient $f : E \longrightarrow \mathbb{C}$ une fonction et $a \in E$ un point au voisinage duquel f est définie à gauche et à droite.

- **Définition :** On dit que f est *dérivable à gauche en a* si $f|_{D \cap]-\infty,a]}$ est dérivable en a, autrement dit si la fonction $x \longmapsto \dfrac{f(x)-f(a)}{x-a}$ possède une limite finie à gauche en a, notée $f'_g(a)$ le cas échéant.

  On définit la dérivabilité à droite en a de façon analogue et on note $f'_d(a)$ le nombre dérivé à droite le cas échéant.
- **(Caractérisation de la dérivabilité à l'aide des dérivabilités à gauche/à droite :**

  f est dérivable en a si et seulement si f est dérivable à gauche et à droite en a avec de plus $f'_g(a) = f'_d(a)$.

Ci-contre, f est dérivable à gauche et à droite en a, mais pas en a car $f'_g(a) \neq f'_d(a)$.

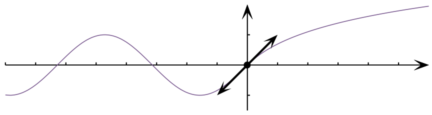


Démonstration

$$\begin{aligned} f \text{ est dérivable en } a \quad &\iff \quad x \longmapsto \frac{f(x)-f(a)}{x-a} \text{ possède une limite finie en } a \\ &\iff \quad x \longmapsto \frac{f(x)-f(a)}{x-a} \text{ possède une limite finie à gauche et à droite en } a \\ &\iff \quad f \text{ est dérivable à gauche et à droite en } a \text{ et } f'_g(a) = f'_d(a). \end{aligned}$$

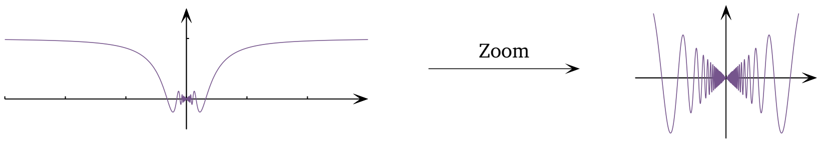
Exemple

La fonction $x \overset{f}{\longmapsto} \begin{cases} \ln(1+x) & \text{si } x \geqslant 0 \\ \sin x & \text{si } x < 0 \end{cases}$ est dérivable en 0 car $f'_g(0) = f'_d(0) = 1$, donc $f'(0) = 1$.

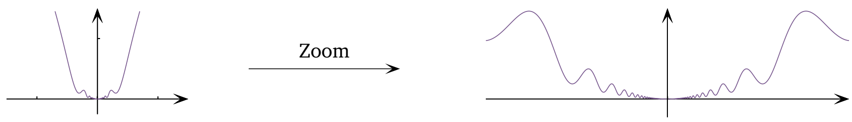


2.2 Idées fausses et contre-exemples

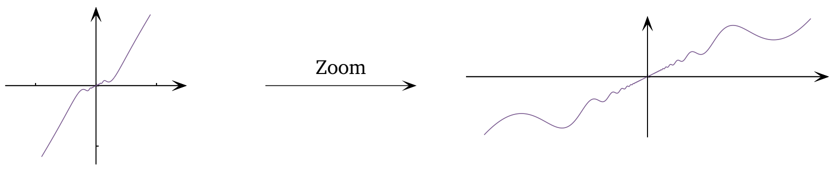
✖ Attention ! Une fonction peut n'être ni dérivable à gauche ni dérivable à droite en un point. C'est le cas de la fonction $x \overset{f}{\longmapsto} x \sin \frac{1}{x}$ en 0 prolongée par continuité en 0 par $f(0) = 0$ car la fonction $x \longmapsto \frac{f(x)-f(0)}{x-0} = \sin \frac{1}{x}$ n'a pas de limite en 0, ni à gauche ni à droite.



✖ Attention ! Une fonction peut avoir un minimum en un point sans être décroissante à gauche et croissante à droite au voisinage de ce point. C'est le cas de la fonction $x \longmapsto x^2 + 2x^2 \sin^2 \frac{1}{x}$ prolongée par 0 en 0.



✖ Attention ! On peut avoir $f'(a) > 0$ en un point a sans que f soit strictement croissante au voisinage de a. C'est le cas de la fonction $x \overset{f}{\longmapsto} x^2 \sin \frac{1}{x} + \frac{x}{2}$ prolongée par 0 en 0. Il est clair que f est dérivable sur $\mathbb{R}^*$ et en revenant au taux d'accroissement de f, on montre aisément que $f'(0) = \frac{1}{2} > 0$. Pourtant, f' change de signe régulièrement au voisinage de 0, donc f change de sens de variation régulièrement au voisinage de 0 car pour tout $x \in \mathbb{R}^*$: $f'(x) = \underbrace{\underbrace{2x \sin \frac{1}{x} - \cos \frac{1}{x}}_{\text{À peu près entre } -1 \text{ et } 1} + \frac{1}{2}}_{\text{À peu près entre } -\frac{1}{2} \text{ et } \frac{3}{2}}$.



2.3 Opérations sur la dérivabilité

Théorème (Opérations sur la dérivabilité)

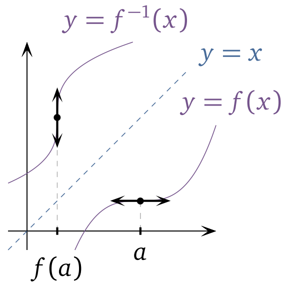
(i) **Combinaison linéaire, produit, quotient :** Pour toutes fonctions $f, g \in \mathscr{D}(E, \mathbb{C})$ et $\lambda, \mu \in \mathbb{C}$, les fonctions $\lambda f + \mu g$ et fg sont dérivables sur E, ainsi que $\frac{f}{g}$ si g ne s'annule pas. En outre :

$$(\lambda f + \mu g)' = \lambda f' + \mu g', \qquad (fg)' = f'g + fg' \qquad \text{et} \qquad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.$$

(ii) **Composition :** Pour toutes fonctions $f \in \mathscr{D}(E, \mathbb{R})$ et $g \in \mathscr{D}(F, \mathbb{C})$, si $f(E) \subset F$, alors $g \circ f$ est dérivable sur E et :

$$(g \circ f)' = f' \times g' \circ f.$$

(iii) **Réciproque :** Soit I un intervalle. Pour toute fonction $f \in \mathscr{D}(I, \mathbb{R})$ bijective de I sur $J = f(I)$, si f' ne s'annule pas sur I, alors f^{-1} est dérivable sur J et $\left(f^{-1}\right)' = \frac{1}{f' \circ f^{-1}}$.



On aurait pu énoncer ce résultat dans le cadre de la dérivabilité en un seul point. Dans le cas de la composition, le théorème énoncerait que si f est dérivable en $a \in E$ et si g est dérivable en $g(a) \in E$, alors $g \circ f$ est dérivable en a.

Attention ! Pour la dérivabilité de f^{-1}, l'hypothèse de non-annulation de f' est cruciale ! Sur la figure ci-contre, f' s'annule en a, donc f possède une tangente horizontale en a. Il en découle que f^{-1} possède une tangente verticale en $f(a)$, donc n'est pas dérivable en $f(a)$.

Démonstration

(i) Fixons $a \in E$. La dérivabilité de g en a implique sa continuité, donc $g(x) \underset{x \to a}{\longrightarrow} g(a)$.

Tout d'abord : $\dfrac{(\lambda f + \mu g)(x) - (\lambda f + \mu g)(a)}{x-a} = \lambda \dfrac{f(x)-f(a)}{x-a} + \mu \dfrac{g(x)-g(a)}{x-a} \underset{x \to a}{\longrightarrow} \lambda f'(a) + \mu g'(a)$.

Ensuite : $\dfrac{(fg)(x) - (fg)(a)}{x-a} = \dfrac{f(x)-f(a)}{x-a} \times g(x) + f(a) \times \dfrac{g(x)-g(a)}{x-a} \underset{x \to a}{\longrightarrow} f'(a)g(a) + f(a)g'(a)$.

Enfin, si $g(a) \neq 0$: $\dfrac{\dfrac{1}{g(x)} - \dfrac{1}{g(a)}}{x-a} = -\dfrac{1}{g(x)g(a)} \times \dfrac{g(x)-g(a)}{x-a} \underset{x \to a}{\longrightarrow} -\dfrac{g'(a)}{g(a)^2}$.

(ii) Fixons $a \in E$. Pour tout $y \in E$, posons $\tau(y) = \begin{cases} \dfrac{g(y) - g\big(f(a)\big)}{y - f(a)} & \text{si } y \neq f(a) \\ g'\big(f(a)\big) & \text{si } y = f(a). \end{cases}$ Par dérivabilité de g en $f(a)$: $\tau(y) \underset{y \to f(a)}{\longrightarrow} g'\big(f(a)\big)$, et pour tout $x \in E$: $\tau\big(f(x)\big)\big(f(x) - f(a)\big) = g \circ f(x) - g \circ f(a)$, y compris pour $x = a$, donc $\dfrac{g \circ f(x) - g \circ f(a)}{x-a} = \dfrac{f(x)-f(a)}{x-a} \times \tau\big(f(x)\big) \underset{x \to a}{\longrightarrow} f'(a)g'\big(f(a)\big)$.

(iii) Fixons $b \in J$. Par continuité de f en $f^{-1}(b)$, f^{-1} est continue en b, donc $f^{-1}(y) \underset{y \to b}{\longrightarrow} f^{-1}(b)$ ♣.

Ensuite : $\dfrac{f(x) - f\big(f^{-1}(b)\big)}{x - f^{-1}(b)} \underset{x \to f^{-1}(b)}{\longrightarrow} f'\big(f^{-1}(b)\big)$ par dérivabilité de f en $f^{-1}(b)$, donc comme f' ne s'annule pas en $f^{-1}(b)$: $\dfrac{x - f^{-1}(b)}{f(x) - f\big(f^{-1}(b)\big)} \underset{x \to f^{-1}(b)}{\longrightarrow} \dfrac{1}{f'\big(f^{-1}(b)\big)}$ ♠.

Composons enfin ♣ et ♠ : $\displaystyle \lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{y-b} = \lim_{y \to b} \frac{f^{-1}(y) - f^{-1}(b)}{f\big(f^{-1}(y)\big) - f\big(f^{-1}(b)\big)} = \frac{1}{f'\big(f^{-1}(b)\big)}$.

Exemple La fonction $x \longmapsto \sqrt{x^3 \operatorname{Arcsin} x}$ est dérivable sur $]-1,1[$.

Démonstration

- La fonction $x \longmapsto x^3 \operatorname{Arcsin} x$ est dérivable sur $]-1,1[$ par produit. Positive sur $[-1,1]$ et nulle seulement en 0, cette fonction est donc dérivable sur $]-1,1[\setminus \{0\}$ à valeurs dans $\mathbb{R}_+^*$. La fonction racine carrée étant dérivable sur $\mathbb{R}_+^*$ SEULEMENT, $x \longmapsto \sqrt{x^3 \operatorname{Arcsin} x}$ est dérivable sur $]-1,1[\setminus \{0\}$ par composition.
- Ce raisonnement ne nous apprend rien sur la fonction $x \longmapsto \sqrt{x^3 \operatorname{Arcsin} x}$ en 0 car nos théorèmes d'opérations sur la dérivabilité nous parlent de dérivabilité mais pas de NON-dérivabilité. De fait, la fonction $x \longmapsto \sqrt{x^3 \operatorname{Arcsin} x}$ est quand même dérivable en 0 car :
$$\frac{\sqrt{x^3 \operatorname{Arcsin} x} - 0}{x - 0} = \frac{\sqrt{x^4}}{x} \times \sqrt{\frac{\operatorname{Arcsin} x}{x}} = x\sqrt{\frac{\operatorname{Arcsin} x - \operatorname{Arcsin} 0}{x - 0}} \underset{x \to 0}{\longrightarrow} 0 \times \sqrt{\operatorname{Arcsin}'(0)} = 0.$$
On pourrait montrer en revanche que $x \longmapsto \sqrt{x^3 \operatorname{Arcsin} x}$ n'est dérivable ni en -1 ni en 1.

Théorème (Opérations sur les dérivées successives) Soit $k \in \mathbb{N} \cup \{\infty\}$.

- **Combinaison linéaire, produit, quotient :** $\mathscr{C}^k(E,\mathbb{C})$ est stable par combinaison linéaire et produit, ainsi que par quotient tant que le dénominateur ne s'annule pas.
En outre, pour toutes fonctions $f, g \in \mathscr{C}^k(E,\mathbb{C})$: $\displaystyle (fg)^{(k)} = \sum_{p=0}^{k} \binom{k}{p} f^{(p)} g^{(k-p)}$ *(formule de Leibniz)*.
- **Composition :** Pour toutes fonctions $f \in \mathscr{C}^k(E,\mathbb{R})$ et $g \in \mathscr{C}^k(F,\mathbb{C})$, si $f(E) \subset F$, alors $g \circ f$ est de classe $\mathscr{C}^k$ sur E.
- **Réciproque :** Soit I un intervalle. Pour toute fonction $f \in \mathscr{C}^k(I,\mathbb{R})$ bijective de I sur $J = f(I)$, SI f' NE S'ANNULE PAS SUR I, alors f^{-1} est de classe $\mathscr{C}^k$ sur J.

On peut remplacer dans chacune de ces assertions la classe $\mathscr{C}^k$ par une hypothèse de dérivabilité k fois.

On n'a pas besoin de dériver 99 fois une fonction pour savoir qu'elle est 100 fois dérivable !

Démonstration On raisonne par récurrence sur k avec des initialisations toutes triviales car être de classe $\mathscr{C}^0$, c'est être continue. Le cas des combinaisons linéaires tombe sous le sens.

- **Produit (hérédité) :** Soit $k \in \mathbb{N}$. On suppose que $\mathscr{C}^k(E,\mathbb{C})$ est stable par produit et que la formule de Leibniz y est vraie pour les dérivées $k^{\text{èmes}}$. Soient $f,g \in \mathscr{C}^{k+1}(E,\mathbb{C})$. Les fonctions f et g sont dérivables car $k+1 \geqslant 1$, donc fg aussi et $(fg)' = f'g + fg'$. Or $f'g$ et fg' sont de classe $\mathscr{C}^k$ par hypothèse de récurrence, donc $(fg)'$ aussi par addition. Conclusion : fg est de classe $\mathscr{C}^{k+1}$. Ensuite :

$$(fg)^{(k+1)} = \left(f'g\right)^{(k)} + \left(fg'\right)^{(k)} \overset{\text{HDR}}{=} \sum_{p=0}^{k} \binom{k}{p} f^{(p+1)} g^{(k-p)} + \sum_{p=0}^{k} \binom{k}{p} f^{(p)} g^{(k-p+1)} = \sum_{p=1}^{k+1} \binom{k}{p-1} f^{(p)} g^{(k-p+1)} + \sum_{p=0}^{k} \binom{k}{p} f^{(p)} g^{(k-p+1)}$$
$$= \sum_{p=0}^{k+1} \binom{k}{p-1} f^{(p)} g^{(k-p+1)} + \sum_{p=0}^{k+1} \binom{k}{p} f^{(p)} g^{(k-p+1)} \overset{\text{Pascal}}{=} \sum_{p=0}^{k+1} \binom{k+1}{p} f^{(p)} g^{(k+1-p)}.$$

- **Quotient (hérédité) :** Soit $k \in \mathbb{N}$. On suppose que $\mathscr{C}^k(E,\mathbb{C})$ est stable par quotient tant que le dénominateur ne s'annule pas. Soient $f,g \in \mathscr{C}^{k+1}(E,\mathbb{C})$. On suppose que g ne s'annule pas sur E. Les fonctions f et g sont dérivables car $k+1 \geqslant 1$, donc $\frac{f}{g}$ aussi et $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$. Or $f'g - fg'$ et g^2 sont de classe $\mathscr{C}^k$ par produit et différence, donc $\left(\frac{f}{g}\right)'$ aussi par hypothèse de récurrence. Conclusion : $\frac{f}{g}$ est de classe $\mathscr{C}^{k+1}$.
- **Composition (hérédité) :** Soit $k \in \mathbb{N}$. On suppose que $g \circ f \in \mathscr{C}^k(E,\mathbb{C})$ pour toutes fonctions $f \in \mathscr{C}^k(D,E)$ et $g \in \mathscr{C}^k(E,\mathbb{C})$. Soient $f \in \mathscr{C}^{k+1}(D,E)$ et $g \in \mathscr{C}^{k+1}(E,\mathbb{C})$. Les fonctions f et g sont dérivables car $k+1 \geqslant 1$, donc $g \circ f$ aussi et $(g \circ f)' = f' \times g' \circ f$. Or $g' \circ f$ est de classe $\mathscr{C}^k$ par hypothèse de récurrence, donc $(g \circ f)'$ aussi par produit. Conclusion : $g \circ f$ est de classe $\mathscr{C}^{k+1}$.
- **Réciproque :** Imiter les preuves précédentes. ■

Exemple La fonction $x \overset{f}{\longmapsto} x^2 \sin \frac{1}{x}$ prolongée par 0 en 0 est dérivable sur tout $\mathbb{R}$, mais de classe $\mathscr{C}^1$ seulement sur $\mathbb{R}^*$.

Démonstration Pour commencer, f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^*$. Ensuite, $\frac{f(x)-f(0)}{x-0} = x \sin \frac{1}{x} \underset{x \to 0}{\longrightarrow} 0$ car la fonction $x \longmapsto \sin \frac{1}{x}$ est bornée, donc f est dérivable en 0. Cela dit, pour tout $x \in \mathbb{R}^*$: $f'(x) = 2x \sin \frac{1}{x} - \cos \frac{1}{x}$ et $x \sin \frac{1}{x} \underset{x \to 0}{\longrightarrow} 0$ alors que $x \longmapsto \cos \frac{1}{x}$ n'a pas de limite en 0, donc f' n'a pas non plus de limite en 0, donc f' n'est pas continue en 0.

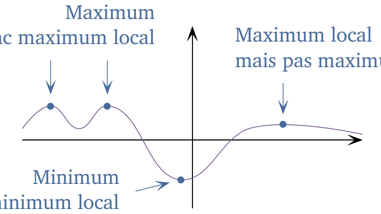
3 Dérivabilité, point de vue global

3.1 Extrema locaux et points critiques

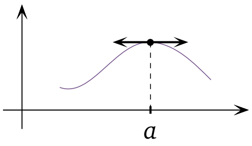
Définition (Extremum local, point critique) Soient $f : E \longrightarrow \mathbb{R}$ une fonction et $a \in E$.

- **Extremum local :** On dit que f possède un *maximum* (resp. *minimum*) *local en* a si f est majorée (resp. minorée) par $f(a)$ au voisinage de a.
- **Point critique :** On dit que a est un *point critique de* f si f est dérivable en a avec $f'(a) = 0$.

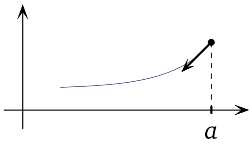
Un maximum local n'est pas forcément un maximum de la fonction sur tout son domaine de définition.



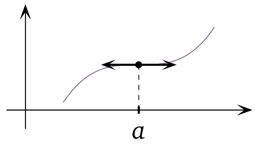
Théorème (Condition nécessaire d'existence d'un extremum local en un point intérieur) Soient $f : E \longrightarrow \mathbb{R}$ une fonction et $a \in \mathring{E}$ — il est très important que a soit intérieur à E. Si f est dérivable en a et possède un extremum local en a, a est un point critique de f.



Situation standard du théorème (cas d'un maximum local).



Il est très important que a soit intérieur à E.



Réciproque fausse : Tout point critique n'est pas forcément le signe d'un extremum local.

Démonstration Étudions le cas d'un maximum local en a. Pour commencer, $a \in \mathring{E}$ donc $]a-\varepsilon, a+\varepsilon[\subset E$ de a pour un certain $\varepsilon > 0$. Ensuite, f possède un maximum local en a, donc $f(x) \leqslant f(a)$ pour tout $x \in]a-\varepsilon, a+\varepsilon[$ — avec le même ε quitte à le diminuer. Par conséquent :

$$\forall x \in]a-\varepsilon, a[, \quad \frac{f(x)-f(a)}{x-a} \geqslant 0 \qquad \text{et} \qquad \forall x \in]a, a+\varepsilon[, \quad \frac{f(x)-f(a)}{x-a} \leqslant 0.$$

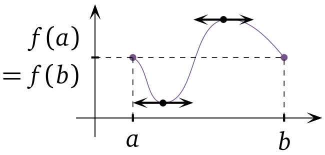
Cela dit, f est dérivable en a, donc $f'(a) \geqslant 0$ par passage à la limite dans l'inégalité de gauche et $f'(a) \leqslant 0$ pour la même raison dans l'inégalité de droite. Conclusion : $f'(a) = 0$.

3.2 Théorème de Rolle et accroissements finis

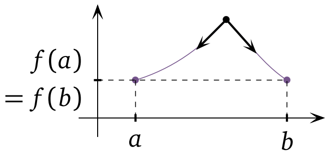
Le théorème de Rolle est le premier d'une longue série de théorèmes globaux sur la dérivabilité. En réalité, cela dit, la globalité du résultat va découler surtout du théorème des bornes atteintes.

Théorème (Théorème de Rolle) Soit $f : [a,b] \longrightarrow \mathbb{R}$ une fonction continue sur $[a,b]$ et dérivable sur $]a,b[$ pour laquelle $f(a) = f(b)$. Il existe un réel $c \in]a,b[$ pour lequel $f'(c) = 0$.

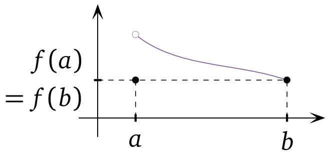
Chaque hypothèse du théorème a son importance.



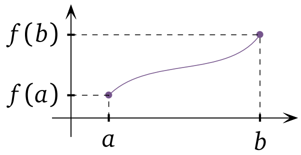
Situation standard du théorème de Rolle.



Si on enlève la dérivabilité même en un point, rien ne va plus.



Si on enlève la continuité, même sur les bords, ce n'est pas mieux.



Si $f(a) \neq f(b)$, c'est toujours la cata.

Démonstration Continue sur le SEGMENT $[a,b]$, f y est bornée et possède un minimum m et un maximum M d'après le théorème des bornes atteintes.

- Si $f(a) = f(b) \neq M$, alors comme f atteint ses bornes, $f(c) = M$ pour un certain $c \in]a,b[$. Par hypothèse, c n'est alors pas une borne de $[a,b]$, donc $f'(c) = 0$ d'après le théorème précédent.
- Si $f(a) = f(b) \neq m$, même raisonnement.
- Pour finir, si $f(a) = f(b) = m = M$, alors f est constante de valeur $M = m$ sur tout $[a,b]$ par définition de m et M, donc f' est nulle sur $[a,b]$ tout entier !

Exemple Pour tous $\alpha, \beta, \gamma \in \mathbb{R}$, le polynôme $P = 4\alpha X^3 + 3\beta X^2 + 2\gamma X - (\alpha+\beta+\gamma)$ possède une racine dans $]0,1[$.

Démonstration Fixons $\alpha, \beta, \gamma \in \mathbb{R}$.

- On a bien sûr envie d'utiliser le TVI. Hélas : $P(0) = -(\alpha+\beta+\gamma)$ et $P(1) = 3\alpha + 2\beta + \gamma$ et le signe de ces deux nombres n'est pas du tout clair.
- Changeons de point de vue. Le polynôme $Q = \alpha X^4 + \beta X^3 + \gamma X^2 - (\alpha+\beta+\gamma)X$ est une primitive de P et $Q(0) = Q(1) = 0$, donc d'après le théorème de Rolle, $P = Q'$ s'annule au moins une fois entre 0 et 1.

Bref, nous connaissons maintenant deux théorèmes d'existence d'un antécédent, le TVI et le théorème de Rolle, dans un cas avec f et dans l'autre avec f'.

Attention ! Le théorème de Rolle est faux pour les fonctions complexes. Par exemple, la fonction $t \longmapsto e^{it}$ est continue sur $[0, 2\pi]$, dérivable sur $]0, 2\pi[$, et $e^{i0} = e^{2i\pi} = 1$, mais sa dérivée $t \longmapsto i\,e^{it}$ ne s'annule pas.

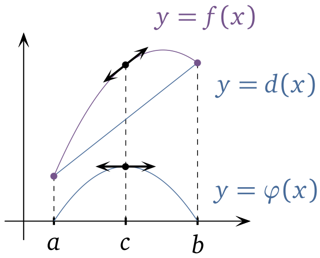
Théorème (Théorème des accroissements finis) Soit $f : [a,b] \longrightarrow \mathbb{R}$ continue sur $[a,b]$ et dérivable sur $]a,b[$.

Il existe un réel $c \in]a,b[$ pour lequel $f'(c) = \dfrac{f(b)-f(a)}{b-a}$, ou encore $f(b) - f(a) = f'(c)(b-a)$.

Le théorème des accroissements finis généralise le théorème de Rolle. La morale de l'histoire ?

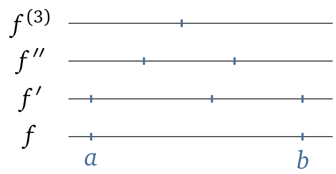
Si j'ai des informations sur f', j'en ai aussi sur les accroissements f.

Typiquement, toute majoration/minoration de f' peut être convertie en une majoration/minoration sur f.



Démonstration Notons d la fonction affine $x \longmapsto \dfrac{f(b)-f(a)}{b-a}(x-a)+f(a)$ et φ la fonction $f-d$. La fonction φ est continue sur $[a,b]$, dérivable sur $]a,b[$ et $\varphi(a)=\varphi(b)=0$, donc $\varphi'(c)=0$ pour un certain $c \in]a,b[$ d'après le théorème de Rolle, ou encore $\dfrac{f(b)-f(a)}{b-a}=f'(c)$. ■

Exemple Soit $f \in \mathscr{C}^3\big([a,b],\mathbb{R}\big)$. On suppose que $f(a)=f'(a)=f(b)=f'(b)=0$.
Alors pour un certain $c \in]a,b[$: $\quad f^{(3)}(c)=0$.

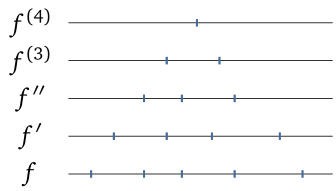


Démonstration D'abord f est continue sur $[a,b]$ et dérivable sur $]a,b[$ avec $f(a)=f(b)=0$, donc $f'(u)=0$ pour un certain $u \in]a,b[$ d'après le théorème de Rolle.

Ensuite, f' est continue sur $[a,u]$ et $[u,b]$ et dérivable sur $]a,u[$ et $]u,b[$ avec $f'(a)=f'(u)=f'(b)=0$, donc $f''(v)=f''(w)=0$ pour certains $v \in]a,u[$ et $w \in]u,b[$ d'après le théorème de Rolle.

Enfin, f'' est continue sur $[v,w]$ et dérivable sur $]v,w[$ avec $f''(v)=f''(w)$, donc $f^{(3)}(c)=0$ pour un certain $c \in]v,w[\subset]a,b[$ d'après le théorème de Rolle.

Exemple Soient I un intervalle et $f \in \mathscr{C}^k(I,\mathbb{R})$. Si f s'annule en au moins $k+1$ points distincts, alors $f^{(k)}$ s'annule en au moins un point. Ce résultat est un grand classique, étudiez-le bien !



Démonstration La figure ci-contre illustre l'idée de la preuve pour $k=4$. On part de 5 zéros (au moins) pour f. En appliquant le théorème de Rolle entre ces zéros on parvient à exhiber 4 zéros (au moins) de f', puis 3 (au moins) pour f''... et enfin (au moins) un zéro de $f^{(4)}$.

Montrons proprement par récurrence que $f^{(i)}$ s'annule en au moins $k-i+1$ points distincts pour tout $i \in [\![0,k]\!]$. Pour $i=k$, c'est le résultat souhaité.

Initialisation : L'assertion pour $i=0$ n'est autre que l'hypothèse du théorème.

Hérédité : Soit $i \in [\![0,k-1]\!]$. Faisons l'hypothèse que $f^{(i)}$ s'annule en au moins $k-i+1$ points $x_1,\ldots,x_{k-i+1}$ rangés dans l'ordre $x_1<\ldots<x_{k-i+1}$. Pour tout $j \in [\![1,k-i]\!]$, $f^{(i)}$ est continue sur $[x_j,x_{j+1}]$, dérivable sur $]x_j,x_{j+1}[$ et $f^{(i)}(x_j)=f^{(i)}(x_{j+1})=0$. Le théorème de Rolle affirme donc l'existence d'un zéro x'_j de $f^{(i+1)}$ compris strictement entre x_j et x_{j+1}. Nous voilà donc à la tête de $k-i=k-(i+1)+1$ zéros de $f^{(i+1)}$, lesquels sont bien distincts car $x'_1<x_2<x'_2<x_3<\ldots<x'_{k-i}<x_{k-i+1}$.

Théorème (Inégalité des accroissements finis) Soient I un intervalle et $f \in \mathscr{D}(I,\mathbb{C})$. Si f' est bornée sur I, alors f y est $\|f'\|_\infty$-lipschitzienne sur I.

L'inégalité des accroissements finis est plus robuste que le théorème de Rolle et le théorème des accroissements finis car elle est énoncée pour des fonctions complexes.

Démonstration Soient $x,y \in I$ deux réels pour lesquels $x<y$.

- **Cas où f est réelle :** Comme f est continue sur $[x,y]$ et dérivable sur $]x,y[$, alors $\dfrac{f(x)-f(y)}{x-y}=f'(c)$ pour un certain $c \in]x,y[$ d'après le théorème des accroissements finis, et donc :
$$\big|f(x)-f(y)\big|=\big|f'(c)\big|\times|x-y| \leqslant \|f'\|_\infty\,|x-y|.$$
- **Cas général :** Pour majorer le nombre complexe $f(x)-f(y)$ en module, ramenons-nous au cas précédent. En notant θ un réel pour lequel $\mathrm{e}^{-\mathrm{i}\theta}\big(f(x)-f(y)\big)$ est réel et φ la fonction $\mathrm{Re}\big(\mathrm{e}^{-\mathrm{i}\theta}f\big)$:
$$\big|f(x)-f(y)\big|=\big|\mathrm{e}^{-\mathrm{i}\theta}\big(f(x)-f(y)\big)\big|=\Big|\mathrm{Re}\Big(\mathrm{e}^{-\mathrm{i}\theta}\big(f(x)-f(y)\big)\Big)\Big|=\big|\varphi(x)-\varphi(y)\big|.$$
Or φ est dérivable sur I et $|\varphi'|=\big|\mathrm{Re}\big(\mathrm{e}^{-\mathrm{i}\theta}f'\big)\big| \leqslant \big|\mathrm{e}^{-\mathrm{i}\theta}f'\big|=|f'| \leqslant \|f'\|_\infty$, donc d'après le cas réel :
$$\big|f(x)-f(y)\big| \leqslant \|f'\|_\infty|x-y|.$$ ■

Exemple Pour tout $x \in \mathbb{R}$: $|\sin x| \leqslant |x|$.

Démonstration La fonction sinus est dérivable sur $\mathbb{R}$ et $|\sin'| = |\cos| \leqslant 1$, donc sinus est 1-lipschitzienne d'après l'inégalité des accroissements finis, donc pour tout $x \in \mathbb{R}$: $|\sin x| = |\sin x - \sin 0| \leqslant |x - 0| = |x|$.

Sur la figure ci-dessous, chaque flèche décrit une implication.

D'après l'inégalité des accroissements finis, l'implication ★ est valable sous réserve que la dérivée de la fonction considérée soit bornée.

$$\begin{tikzcd}
 & & & & & \text{Lipschitzianité} \arrow[dr] & \\
\text{Classe } \mathscr{C}^\infty \arrow[r] & \cdots \arrow[r] & \text{Classe } \mathscr{C}^2 \arrow[r] & \text{Dérivabilité deux fois} \arrow[r] & \text{Classe } \mathscr{C}^1 \arrow[r] & \text{Dérivabilité} \arrow[r] \arrow[u, dashed, "\bigstar"] & \text{Continuité (Classe } \mathscr{C}^0)
\end{tikzcd}$$

En réalité, nous avons déjà rencontré rapidement la version $\mathscr{C}^1$ de l'inégalité des accroissements finis au chapitre « Techniques élémentaires de calcul intégral ». En effet, pour toute fonction $f \in \mathscr{C}^1\big([a,b],\mathbb{C}\big)$:

$$\big|f(b) - f(a)\big| = \left|\int_a^b f'(t)\,\mathrm{d}t\right| \leqslant \int_a^b \big|f'(t)\big|\,\mathrm{d}t \leqslant \int_a^b \|f'\|_\infty\,\mathrm{d}t = \|f'\|_\infty\,(b-a), \qquad \text{ou encore} \quad \left|\frac{f(b)-f(a)}{b-a}\right| \leqslant \|f'\|_\infty.$$

Au moins, c'est rapide et efficace, mais nous n'avons pas encore défini proprement la notion d'intégrale et l'inégalité des accroissements finis de ce chapitre est vraie pour des fonctions seulement dérivables, pas forcément de classe $\mathscr{C}^1$.

3.3 Constance, monotonie et dérivabilité

Théorème (Caractérisation des fonctions constantes dérivables) Soient I un **intervalle** et $f \in \mathscr{D}(I,\mathbb{C})$.

f est constante sur I si et seulement si f' est nulle sur I.

Démonstration Une preuve du cas réel suffit via la caractérisation de la dérivabilité à partir des parties réelle et imaginaire. Soit donc $f \in \mathscr{D}(I,\mathbb{R})$.

- Si f est constante sur I, alors pour tout $a \in I$: $\dfrac{f(x)-f(a)}{x-a} = 0 \underset{x \to a}{\longrightarrow} 0$, donc $f'(a) = 0$.
- Réciproquement, supposons f' nulle sur I et donnons-nous deux réels $x, y \in I$ pour lesquels $x < y$. Comme f est continue sur $[x,y]$ et dérivable sur $]x,y[$, $\dfrac{f(y)-f(x)}{y-x} = f'(c) = 0$ pour un certain $c \in]x,y[$ d'après le théorème des accroissements finis. Ainsi, $f(x) = f(y)$ donc f est constante.

Théorème (Caractérisation des fonctions monotones dérivables) Soient I un **intervalle** et $f \in \mathscr{D}(I,\mathbb{R})$.

(i) f est croissante sur I si et seulement si f' est positive (ou nulle) sur I.

(ii) f est strictement croissante sur I si et seulement si f' est positive (ou nulle) sur I et n'est identiquement nulle sur aucun intervalle $[a,b]$ avec $a < b$.

En particulier, si f' est strictement positive sur I, alors f est strictement croissante sur I.

Démonstration

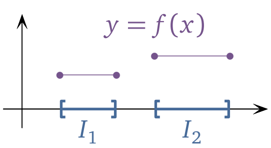
(i) Remplacez simplement les inégalités par des égalités dans la preuve du théorème précédent.

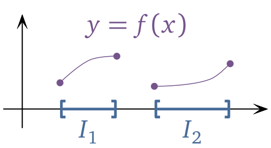
(ii) Supposons d'abord f strictement croissante sur I. D'après (i), f' est positive sur I. Ensuite, si f' est identiquement nulle sur un intervalle $[a,b]$ avec $a \leqslant b$, alors f y est constante donc $f(a) = f(b)$, donc $a = b$ par stricte monotonie.

Réciproquement, supposons que f' est positive sur I et qu'elle n'est identiquement nulle sur aucun intervalle $[a,b]$ avec $a < b$. D'après (i), f est croissante. Soient $x, y \in I$ deux réels pour lesquels $x < y$. Aussitôt, $f(x) \leqslant f(y)$ par croissance. Et si $f(x) = f(y)$, la croissance de f montre que f est constante sur $[x,y]$, donc que f' y est identiquement nulle, ce qui est faux. Forcément, $f(x) < f(y)$.

✖ **Attention !** Mine de rien, il est indispensable que I soit un **intervalle**.

f est constante sur I_1 et sur I_2, donc $f' = 0$ sur I_1 et sur I_2, mais f n'est pas constante sur $I_1 \cup I_2$.





f est croissante sur I_1 et sur I_2, donc $f' \geqslant 0$ sur I_1 et sur I_2, mais f n'est pas croissante sur $I_1 \cup I_2$.

3.4 Points fixes attractifs/répulsifs

Nous avons jusqu'ici étudié les suites récurrentes $u_{n+1} = f(u_n)$ à l'aide du théorème de la limite monotone. En cas de convergence vers un réel ℓ et sous l'hypothèse que f est continue en ℓ, nous savons que ℓ est un point fixe de f. En revanche, le théorème de la limite monotone ne précise pas la vitesse à laquelle $(u_n)_{n\in\mathbb{N}}$ s'approche de ℓ. Pour une approximation de ℓ à 10^{-3} près, suffit-il de calculer u_{10} ou bien faut-il pousser jusqu'à u_{1000}, voire $u_{10^{10^{10}}}$? En d'autres termes, l'écart $|u_n - \ell|$ converge-t-il plutôt vite ou plutôt lentement vers 0 ? Car $|u_n - \ell| \approx \dfrac{1}{2^n}$ et $|u_n - \ell| \approx \dfrac{1}{\ln n}$, ce n'est pas la même chose.

Comme nous allons le voir, l'inégalité des accroissements finis garantit souvent à peu de frais une convergence rapide de $(u_n)_{n\in\mathbb{N}}$ vers ℓ. Le principe est simple, **vraiment important**, et on peut l'étendre à bien d'autres contextes. Soient I un intervalle, $f : I \longrightarrow \mathbb{R}$ une fonction et $u_0 \in I$. Formulons quelques hypothèses raisonnables :

— I est stable par f,

— f possède un point fixe ℓ sur I,

Ouvert en 1, attention !

— f est contractante, i.e. K-lipschitzienne pour un certain $K \in [0, 1[$, typiquement grâce à l'inégalité des accroissements finis.

Nous pouvons ainsi noter $(u_n)_{n\in\mathbb{N}}$ l'unique suite pour laquelle $u_{n+1} = f(u_n)$ pour tout $n \in \mathbb{N}$. Les hypothèses précédentes permettent alors de montrer :

— d'abord que pour tout $n \in \mathbb{N}$: $\quad |u_{n+1} - u_n| = \big|f(u_n) - f(\ell)\big| \leqslant K\,|u_n - \ell|$,

— puis par récurrence que pour tout $n \in \mathbb{N}$: $\quad |u_n - \ell| \leqslant K^n\,|u_0 - \ell|$,

— et enfin par encadrement que $u_n \underset{n\to+\infty}{\longrightarrow} \ell$ car $0 \leqslant K < 1$.

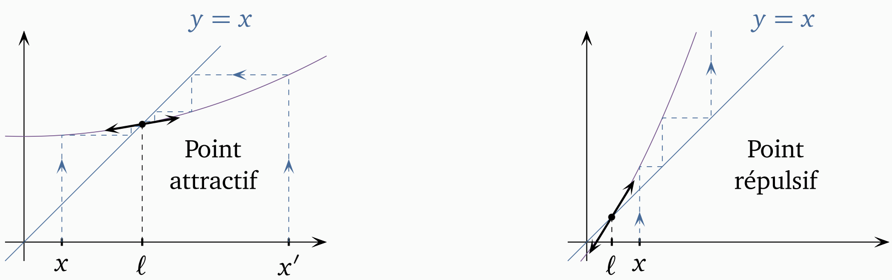
Bref, $(u_n)_{n\in\mathbb{N}}$ converge vers ℓ, mais nous n'avons pas eu besoin de connaître sa monotonie pour le prouver et nous avons obtenu mieux que la convergence en réalité. L'inégalité $|u_n - \ell| \leqslant K^n\,|u_0 - \ell|$ montre que $(u_n)_{n\in\mathbb{N}}$ converge rapidement vers ℓ, au moins aussi vite que la suite géométrique $(K^n)_{n\in\mathbb{N}}$ tend vers 0.

En principe, il est maintenant clair que l'inégalité stricte $K < 1$ est cruciale. La méthode qui précède ne s'applique qu'à des fonctions dont les cordes ont des pentes majorées par un réel strictement inférieur à 1 en valeur absolue, i.e. dont les variations sont faibles. En particulier, l'inégalité $\left|\dfrac{f(x) - f(\ell)}{x - \ell}\right| \leqslant K$ pour tout $x \in I \setminus \{\ell\}$ montre que si f est dérivable sur I, alors $\big|f'(\ell)\big| \leqslant K < 1$ après passage à la limite. On dit que ℓ est un point fixe *attractif* de f.

Définition-théorème (Point fixe attractif/répulsif) Soient I un intervalle, $f \in \mathscr{C}^1(I, \mathbb{R})$ et $\ell \in I$ un point fixe de f. On suppose I stable par f et on note $\big(u_n(x)\big)_{n\in\mathbb{N}}$ pour tout $x \in I$ la suite définie par $u_0(x) = x$ et $u_{n+1}(x) = f\big(u_n(x)\big)$ pour tout $n \in \mathbb{N}$.

(i) **Point fixe attractif :** On dit que ℓ est *attractif (pour f)* si $\big|f'(\ell)\big| < 1$. Le cas échéant, il existe un réel $\alpha > 0$ pour lequel $I \cap]\ell-\alpha, \ell+\alpha[$ est stable par f et la suite $\big(u_n(x)\big)_{n\in\mathbb{N}}$ converge pour toute valeur initiale $x \in I \cap]\ell-\alpha, \ell+\alpha[$.

(ii) **Point fixe répulsif :** On dit que ℓ est *répulsif (pour f)* si $\big|f'(\ell)\big| > 1$. Le cas échéant, il existe un réel $\alpha > 0$ pour lequel pour tout $x \in I \cap]\ell-\alpha, \ell+\alpha[$, soit $\big(u_n(x)\big)_{n\in\mathbb{N}}$ ne converge pas vers ℓ, soit $\big(u_n(x)\big)_{n\in\mathbb{N}}$ est stationnaire de valeur ℓ.



On ne peut rien dire de général si $\big|f'(\ell)\big| = 1$.

En résumé, si ℓ est attractif, toute suite dont la valeur initiale est assez proche de ℓ est piégée autour de ℓ et irrémédiablement aspirée vers lui. Nous montrerons plus tard que si $f'(\ell) \neq 0$, alors pour tout $x \in I \cap]\ell-\alpha, \ell+\alpha[$: $u_n(x) - \ell \underset{n\to+\infty}{\sim} \lambda f'(\ell)^n$ pour un certain $\lambda \in \mathbb{R}$ qui dépend de x. Au contraire, si ℓ est répulsif, aucune suite dont la valeur initiale est assez proche de ℓ ne peut converger vers ℓ, sauf à être carrément stationnaire.

Qu'il soit attractif ou répulsif, ℓ n'exerce qu'une action locale sur les points qui l'entourent et son rayon d'action dépend de la fonction f. Au-delà, ℓ n'a plus d'influence.

Démonstration

(i) Supposons ℓ attractif. Continue en ℓ, $|f'|$ est donc majorée par $K = \dfrac{\left|f'(\ell)\right|+1}{2} < 1$ au voisinage de ℓ, disons sur $I \cap \left]\ell-\alpha, \ell+\alpha\right[$ pour un certain $\alpha > 0$. Ainsi, f est K-lipschitzienne sur $I \cap \left]\ell-\alpha, \ell+\alpha\right[$ d'après l'inégalité des accroissements finis.

Pour tout $x \in I \cap \left]\ell-\alpha, \ell+\alpha\right[$: $\quad \left|f(x)-\ell\right| = \left|f(x)-f(\ell)\right| \leqslant K\,|x-\ell| < K\alpha < \alpha \quad$ et $f(x) \in I$ par stabilité de I par f, donc $f(x) \in I \cap \left]\ell-\alpha, \ell+\alpha\right[$. En d'autres termes, $I \cap \left]\ell-\alpha, \ell+\alpha\right[$ est stable par f.

Fixons enfin $x \in I \cap \left]\ell-\alpha, \ell+\alpha\right[$. La suite $\big(u_n(x)\big)_{n\in\mathbb{N}}$ est à valeurs dans $I \cap \left]\ell-\alpha, \ell+\alpha\right[$ par stabilité, donc $|u_{n+1}-\ell| = \left|f(u_n)-f(\ell)\right| \leqslant K\,|u_n-\ell|$ pour tout $n \in \mathbb{N}$ par K-lipschitzianité, puis $|u_n-\ell| \leqslant K^n\,|u_0-\ell|$ par récurrence, et enfin $u_n \underset{n\to+\infty}{\longrightarrow} \ell$ par encadrement.

(ii) Supposons ℓ répulsif. Continue en ℓ, $|f'|$ est donc minorée par $K = \dfrac{\left|f'(\ell)\right|+1}{2} > 1$ au voisinage de ℓ, disons sur $I \cap \left]\ell-\alpha, \ell+\alpha\right[$ pour un certain $\alpha > 0$. D'après le théorème des accroissements finis, pour tout $x \in I \cap \left]\ell-\alpha, \ell+\alpha\right[$ distinct de ℓ, $f(x)-f(\ell) = f'(c)(x-\ell)$ pour un certain c compris strictement entre ℓ et x, donc $\left|f(x)-\ell\right| = \left|f(x)-f(\ell)\right| = \left|f'(c)\right| \times |x-\ell| \geqslant K\,|x-\ell| > |x-\ell|$.

À présent, soit $x \in I \cap \left]\ell-\alpha, \ell+\alpha\right[$. Faisons l'hypothèse que $\big(u_n(x)\big)_{n\in\mathbb{N}}$ converge vers ℓ. Dans ce cas, $u_n \in I \cap \left]\ell-\alpha, \ell+\alpha\right[$ à partir d'un certain rang N, donc si $u_n \neq \ell$ pour tout $n \geqslant N$, alors $|u_{n+1}-\ell| > |u_n-\ell|$ comme nous l'avons vu à l'instant, donc la suite $\big(|u_n-\ell|\big)_{n\geqslant N}$ est positive et strictement croissante, mais cela contredit le fait qu'elle converge vers 0. Conclusion : $u_{N'} = \ell$ pour un certain $N' \geqslant N$, donc $u_n = \ell$ pour tout $n \geqslant N'$, autrement dit $(u_n)_{n\in\mathbb{N}}$ est stationnaire.

Exemple Le polynôme X^3+3X-1 possède une unique racine réelle, égale à 0,32218 à 10^{-5} près.

Démonstration

- Le TVI strictement monotone montre assez vite que la fonction $x \longmapsto x^3+3x-1$ est bijective de $\mathbb{R}$ sur $\mathbb{R}$. Le polynôme $P = X^3+3X-1$ possède donc une et une seule racine réelle ρ. Plus précisément, $\rho \in [0,1]$ car $P(0) = -1 \leqslant 0$ et $P(1) = 3 \geqslant 0$, et comme : $\quad P(\rho) = 0 \quad\Longleftrightarrow\quad \dfrac{1}{\rho^2+3} = \rho, \quad$ le réel ρ est aussi le seul point fixe de la fonction $x \overset{f}{\longmapsto} \dfrac{1}{x^2+3}$ sur $[0,1]$. En outre, $[0,1]$ est stable par f.
- Nous pouvons ainsi noter $(u_n)_{n\in\mathbb{N}}$ la suite définie par $u_0 = 0$ et $u_{n+1} = f(u_n)$ pour tout $n \in \mathbb{N}$. Avec un peu de chance, $(u_n)_{n\in\mathbb{N}}$ converge vers ρ, mais cela dépend de l'amplitude des variations de f. Majorons donc $|f'|$ pour pouvoir utiliser l'inégalité des accroissements finis. Pour tout $x \in [0,1]$:
$$f'(x) = -\frac{2x}{(x^2+3)^2} \qquad \text{et} \qquad f''(x) = \frac{6(x^2-1)}{(x^2+3)^3} \leqslant 0,$$
donc f' est décroissante sur $[0,1]$, donc pour tout $x \in [0,1]$: $\quad -\dfrac{1}{8} = f'(1) \leqslant f'(x) \leqslant f'(0) = 0. \quad$ Ainsi, f est contractante sur $[0,1]$, $\dfrac{1}{8}$-lipschitzienne en l'occurrence. En particulier, $\left|f'(\rho)\right| \leqslant \dfrac{1}{8} < 1$ donc ρ est attractif.
- Finalement, $|u_{n+1}-\rho| = \left|f(u_n)-f(\rho)\right| \leqslant \dfrac{1}{8}\,|u_n-\rho|$ pour tout $n \in \mathbb{N}$, donc $|u_n-\rho| \leqslant \dfrac{|u_0-\rho|}{8^n} \leqslant \dfrac{1}{8^n}$ par récurrence et il en découle que $u_n \underset{n\to+\infty}{\longrightarrow} \rho$ par encadrement. Nous pouvons ainsi utiliser $(u_n)_{n\in\mathbb{N}}$ pour calculer une valeur approchée de ρ à 10^{-5} près, mais combien de termes devons-nous calculer ? Pour tout $n \in \mathbb{N}$: $\quad \dfrac{1}{8^n} \leqslant 10^{-5} \quad\Longleftrightarrow\quad 8^n \geqslant 100\,000, \quad$ et cette inégalité est vraie à partir de $n = 6$. On obtient donc une valeur approchée de ρ à 10^{-5} près en calculant u_6 : $\quad u_6 \approx 0{,}32218$.

3.5 Limite de la dérivée

Attention ! Les limites $\displaystyle\lim_{x\to a} \frac{f(x)-f(a)}{x-a}$ et $\displaystyle\lim_{x\to a} f'(x)$ sont conceptuellement très différentes. La première est liée à la dérivabilité de f et la deuxième à la continuité de f', mais l'expérience montre que vous confondez souvent les deux gestes :

$$\boxed{\text{Prolonger } f' \text{ par continuité en } a \quad\neq\quad \text{Montrer que } f \text{ est dérivable en } a.}$$

Par exemple, la fonction $\mathbb{1}_{\mathbb{R}_+}$ est dérivable sur $\mathbb{R}^*$ et sa dérivée $\left(\mathbb{1}_{\mathbb{R}_+}\right)'$ est nulle sur $\mathbb{R}^*$, donc prolongeable par continuité en 0. Faut-il en déduire que $\mathbb{1}_{\mathbb{R}_+}$ est dérivable en 0 ? Certainement pas, elle n'est même pas continue en 0. Cet exemple illustre l'importance de la continuité en a des deux prochains théorèmes.

Théorème **(Théorème de la limite de la dérivée)** Soient I un intervalle, $a \in I$, $\ell \in \overline{\mathbb{R}}$ et $f : I \longrightarrow \mathbb{R}$ une fonction dérivable sur $I \setminus \{a\}$ et continue en a. Si $f'(x) \underset{\substack{x \to a \\ x \neq a}}{\longrightarrow} \ell$, alors $\dfrac{f(x)-f(a)}{x-a} \underset{x \to a}{\longrightarrow} \ell$.

La précision $x \neq a$ en indice signifie qu'on s'intéresse à la limite en a de la restriction $f'_{|I\setminus\{a\}}$.

Démonstration Pour tout $x \in I \cap]a,+\infty[$, f est continue sur $[a,x]$ et dérivable sur $]a,x[$, donc d'après le théorème des accroissements finis, $\dfrac{f(x)-f(a)}{x-a} = f'(c_x)$ pour un certain $c_x \in]a,x[$ — idem si $x \in I \cap]-\infty,a[$. Nous héritons ainsi d'une fonction $c : I \setminus \{a\} \longrightarrow I \setminus \{a\}$ pour laquelle $|c_x - a| < |x-a|$ pour tout $x \in I \setminus \{a\}$. En particulier, $c_x \underset{x \to a}{\longrightarrow} a$ par encadrement, donc $\dfrac{f(x)-f(a)}{x-a} = f'(c_x) \underset{x \to a}{\longrightarrow} \ell$ par composition.

Théorème **(Version $\mathscr{C}^1$ du théorème de la limite de la dérivée)** Soient I un intervalle, $a \in I$ et $f : I \longrightarrow \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ sur $I \setminus \{a\}$ et continue en a. Si $f'_{|I\setminus\{a\}}$ possède une limite finie en a, alors f est de classe $\mathscr{C}^1$ sur I tout entier.

Démonstration En notant ℓ la limite finie de l'hypothèse, $\dfrac{f(x)-f(a)}{x-a} \underset{x \to a}{\longrightarrow} \ell$ d'après le théorème de la limite de la dérivée, donc f est dérivable en a et $f'(a) = \ell$. Il en découle en retour que $f'(x) \underset{\substack{x \to a \\ x \neq a}}{\longrightarrow} \ell = f'(a)$, donc f' est continue en a.

Exemple Pour tout $\alpha > 2$, la fonction $x \overset{f}{\longmapsto} x^\alpha \sin \dfrac{1}{x}$ prolongée par 0 en 0 est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+$ tout entier.

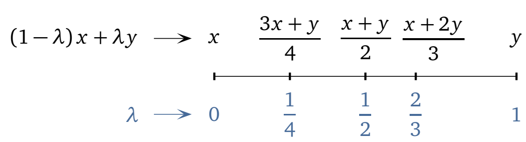
Démonstration Pour commencer, f est de classe $\mathscr{C}^1$ sur $\mathbb{R}_+^*$ et $f(x) \underset{x \to 0^+}{\longrightarrow} 0 = f(0)$ car la fonction $x \longmapsto \sin \dfrac{1}{x}$ est bornée, donc f est continue sur $\mathbb{R}$ tout entier.

Ensuite : $f'(x) = \alpha x^{\alpha-1} \sin \dfrac{1}{x} - x^{\alpha-2} \cos \dfrac{1}{x} \underset{x \to 0^+}{\longrightarrow} 0$ car $\alpha > 2$, donc d'après le théorème de la limite de la dérivée, f est dérivable en 0 et f' y est continue.

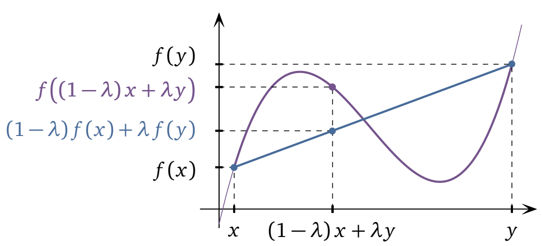
4 CONVEXITÉ

Dans cette partie, I est un intervalle et non pas une partie quelconque de $\mathbb{R}$. En d'autres termes, $[x,y] \subset I$ pour tous $x,y \in I$ pour lesquels $x \leqslant y$, où l'on rappelle que par définition $[x,y] = \left\{ t \in \mathbb{R} \mid \ x \leqslant t \leqslant y \right\}$.

Le segment $[x,y]$ peut cela dit être décrit autrement. Quand λ décrit $[0,1]$, $\lambda(y-x)$ décrit $[0,y-x]$, donc $(1-\lambda)x + \lambda y = x + \lambda(y-x)$ décrit $[x,y]$.



Conclusion : $\boxed{[x,y] = \left\{ (1-\lambda)x + \lambda y \mid \ \lambda \in [0,1] \right\}.}$



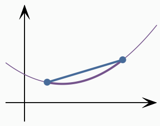
Donnons-nous maintenant une fonction $f : I \longrightarrow \mathbb{R}$. Quand λ décrit $[0,1]$, $(1-\lambda)f(x) + \lambda f(y)$ décrit le segment d'extrémités $f(x)$ et $f(y)$, et dans le plan, le point de coordonnées $\big((1-\lambda)x+\lambda y\ ,\ (1-\lambda)f(x)+\lambda f(y)\big)$ décrit le segment d'extrémités $\big(x,f(x)\big)$ et $\big(y,f(y)\big)$, appelé une *corde de f*.

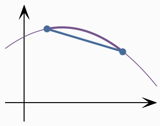
Si $x \neq y$, la droite complète qui passe par $\big(x,f(x)\big)$ et $\big(y,f(y)\big)$ est quant à elle appelée une *sécante de f*.

Définition (Fonction convexe/concave) Soit $f : I \longrightarrow \mathbb{R}$ une fonction.

- **Fonction convexe :** On dit que f est *convexe* si son graphe est situé en-dessous de toutes ses cordes, i.e. si :
$$\forall x, y \in I, \quad \forall \lambda \in [0,1], \quad f\big((1-\lambda)x + \lambda y\big) \leqslant (1-\lambda) f(x) + \lambda f(y).$$
- **Fonction concave :** On dit que f est *concave* si son graphe est situé au-dessus de toutes ses cordes, i.e. si :
$$\forall x, y \in I, \quad \forall \lambda \in [0,1], \quad f\big((1-\lambda)x + \lambda y\big) \geqslant (1-\lambda) f(x) + \lambda f(y).$$
Il est équivalent de dire que $-f$ est convexe.

Fonction convexe,
perpétuel virage à gauche





Fonction concave,
perpétuel virage à droite

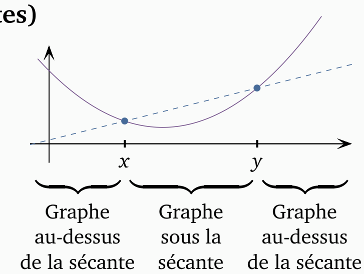
Nous n'énoncerons et ne démontrerons ci-après que des théorèmes sur les fonctions convexes, mais le cas des fonctions concaves s'en déduit par multiplication par -1. Remplacez simplement dans chaque résultat convexe le plus petit par le plus grand, le positif par le négatif, la croissance par la décroissance et le dessus par le dessous.

Exemple La fonction valeur absolue est convexe sur $\mathbb{R}$ car pour tous $x, y \in \mathbb{R}$ et $\lambda \in [0,1]$, d'après l'inégalité triangulaire :
$$\big|(1-\lambda)x + \lambda y\big| \leqslant |1-\lambda| \times |x| + |\lambda| \times |y| = (1-\lambda)|x| + \lambda |y|.$$

Théorème (Position du graphe d'une fonction convexe par rapport à ses sécantes)

Soient $f : I \longrightarrow \mathbb{R}$ une fonction convexe et $x, y \in I$ avec $x < y$.

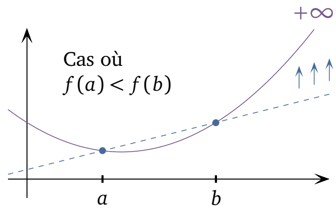
Le graphe de f est situé sous sa sécante sur $[x, y]$ et au-dessus à l'extérieur de $[x, y]$.



Démonstration Montrons que pour tout $t \in I \cap [y, +\infty[$: $f(t) \geqslant \dfrac{f(y) - f(x)}{y - x}(t - y) + f(y)$. On travaillerait de même à gauche de x. Fixons $t \in I \cap [y, +\infty[$. On veut montrer que :
$$(t-x) f(y) \leqslant (t-y) f(x) + (y-x) f(t), \qquad \text{i.e. que } f(y) \leqslant \left(1 - \frac{y-x}{t-x}\right) f(x) + \frac{y-x}{t-x} f(t).$$
Or $x < y \leqslant t$, donc le réel $\lambda = \dfrac{y-x}{t-x}$ appartient à $[0,1]$, et il se trouve que $(1-\lambda)x + \lambda t = \dfrac{t-y}{t-x}\,x + \dfrac{y-x}{t-x}\,t = y$. On conclut par convexité de f : $\quad f(y) = f\big((1-\lambda)x + \lambda t\big) \leqslant (1-\lambda) f(x) + \lambda f(t)$.

Exemple Soient $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction convexe et $a, b \in \mathbb{R}$ avec $a < b$. Si $f(a) < f(b)$, alors $f(x) \underset{x \to +\infty}{\longrightarrow} +\infty$, et si $f(a) > f(b)$, alors $f(x) \underset{x \to -\infty}{\longrightarrow} +\infty$.

Démonstration Le graphe de f est au-dessus de sa sécante à l'extérieur de $[a, b]$, donc pour tout $t < a$ ou $t > b$: $\quad f(t) \geqslant \dfrac{f(b) - f(a)}{b - a}(t - a) + f(a)$. Le résultat en découle par minoration en distinguant bien le cas $f(a) < f(b)$ où l'on fait tendre t vers $+\infty$ du cas $f(a) > f(b)$ où on le fait tendre vers $-\infty$.

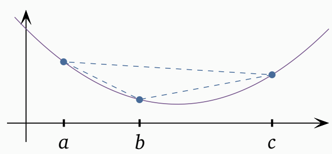


Théorème (Caractérisation de la convexité par les pentes des sécantes, inégalité des pentes) Soit $f : I \longrightarrow \mathbb{R}$ une fonction.

(i) **Caractérisation de la convexité par les pentes des sécantes :**

f est convexe sur I si et seulement si la fonction $x \longmapsto \dfrac{f(x) - f(a)}{x - a}$ est croissante sur $I \setminus \{a\}$ pour tout $a \in I$.

(ii) **Inégalité des pentes :** Si f est convexe sur I, alors pour tous $a, b, c \in I$ pour lesquels $a < b < c$:
$$\frac{f(b) - f(a)}{b - a} \leqslant \frac{f(c) - f(a)}{c - a} \leqslant \frac{f(c) - f(b)}{c - b}.$$



Bref, les pentes des sécantes croissent en cas de convexité !

Démonstration

(i) Supposons d'abord f convexe sur I. Fixons $a \in I$ et montrons que la fonction $t \longmapsto \dfrac{f(t)-f(a)}{t-a}$ est croissante sur $I \setminus \{a\}$. Soient $x, y \in I \setminus \{a\}$ deux réels pour lesquels $x < y$. Nous allons distinguer deux cas et les traiter chacun en comparant le graphe de f à sa sécante passant par $\big(a, f(a)\big)$ et $\big(x, f(x)\big)$, i.e. en comparant les fonctions f et $t \longmapsto \dfrac{f(x)-f(a)}{x-a}(t-a)+f(a)$.

— Si $y > a$, y est situé à l'extérieur du segment $[a, x]$ ou $[x, a]$, donc $f(y) \geqslant \dfrac{f(x)-f(a)}{x-a}(y-a)+f(a)$, ou encore $\dfrac{f(x)-f(a)}{x-a} \leqslant \dfrac{f(y)-f(a)}{y-a}$ puisque $y-a>0$.

— Si $y < a$, y appartient au segment $[a, x]$ ou $[x, a]$, donc $f(y) \leqslant \dfrac{f(x)-f(a)}{x-a}(y-a)+f(a)$, ou encore $\dfrac{f(x)-f(a)}{x-a} \leqslant \dfrac{f(y)-f(a)}{y-a}$ puisque $y-a<0$.

C'est l'inégalité souhaitée dans les deux cas.

Réciproquement, supposons $t \longmapsto \dfrac{f(t)-f(a)}{t-a}$ croissante sur $I \setminus \{a\}$ pour tout $a \in I$ et montrons que f est convexe. Fixons $x, y \in I$ et $\lambda \in [0, 1]$. Pour montrer que $f\big((1-\lambda)x+\lambda y\big) \leqslant (1-\lambda)f(x)+\lambda f(y)$, on peut supposer sans perte de généralité que $\lambda \in]0, 1[$ et $x \neq y$, et même $x < y$ quitte à remplacer λ par $1-\lambda$.

Posons maintenant $a = (1-\lambda)x+\lambda y$, de sorte que $x < a < y$. La fonction $t \longmapsto \dfrac{f(t)-f(a)}{t-a}$ étant croissante par hypothèse : $\dfrac{f(x)-f(a)}{x-a} \leqslant \dfrac{f(y)-f(a)}{y-a}$, donc sachant que $x-a<0$ et $y-a>0$:

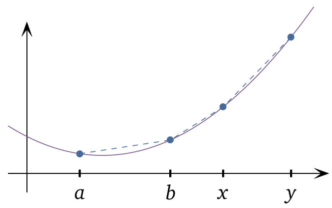
$$(y-a)\big(f(x)-f(a)\big) \geqslant (x-a)\big(f(y)-f(a)\big), \qquad \text{puis } f(a) \leqslant \frac{y-a}{y-x}f(x)+\frac{a-x}{y-x}f(y),$$

or cette inégalité s'écrit aussi $f(a) \leqslant (1-\lambda)f(x)+\lambda f(y)$ par définition de a et c'est le résultat voulu.

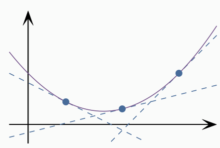
(ii) La fonction $t \longmapsto \dfrac{f(t)-f(a)}{t-a}$ est croissante sur $I \setminus \{a\}$ d'après (i), donc $\dfrac{f(b)-f(a)}{b-a} \leqslant \dfrac{f(c)-f(a)}{c-a}$. De même, la fonction $t \longmapsto \dfrac{f(t)-f(c)}{t-c}$ est croissante sur $I \setminus \{c\}$, donc $\dfrac{f(c)-f(a)}{c-a} \leqslant \dfrac{f(c)-f(b)}{c-b}$. ■

Exemple Soient $f : I \longrightarrow \mathbb{R}$ une fonction convexe et $a, b \in I$ deux réels pour lesquels $a < b$ et $f(a) \leqslant f(b)$. Alors f est croissante sur $I \cap [b, +\infty[$. Ce résultat complète bien sûr notre dernier exemple.

Démonstration Soient $x, y \in I \cap [b, +\infty[$ deux réels pour lesquels $x < y$. Pour montrer que $f(x) \leqslant f(y)$, une simple figure nous suggère d'appliquer deux fois l'inégalité des pentes : $\dfrac{f(y)-f(x)}{y-x} \geqslant \dfrac{f(x)-f(b)}{x-b} \geqslant \dfrac{f(b)-f(a)}{b-a} \geqslant 0$.



Théorème (Caractérisation des fonctions convexes dérivables) Soit $f \in \mathscr{D}(I, \mathbb{R})$. Les assertions suivantes sont équivalentes :



(i) f est convexe sur I.

(ii) f' est croissante sur I — ou bien $f'' \geqslant 0$ si f est deux fois dérivable sur I.

(iii) Le graphe de f est situé au-dessus de toutes ses tangentes.

Démonstration

(i) $\Longrightarrow$ (ii) Soient $x, y \in I$ deux réels pour lesquels $x < y$. Pour tout $t \in]x, y[$:

$$\frac{f(t)-f(x)}{t-x} \leqslant \frac{f(y)-f(x)}{y-x} \leqslant \frac{f(y)-f(t)}{y-t}$$

par convexité de f d'après l'inégalité des pentes. Faisons tendre t vers x à droite, puis indépendamment vers y à gauche : $f'(x) \leqslant \dfrac{f(y)-f(x)}{y-x} \leqslant f'(y)$. Comme voulu, f' est croissante sur I.

(ii) $\Longrightarrow$ (iii) Fixons $a \in I$ et montrons que le graphe de f est situé au-dessus de sa tangente en a. Notons pour cela φ la fonction $t \longmapsto f(t)-f'(a)(t-a)-f(a)$. Cette fonction est dérivable sur I et $\varphi'(x) = f'(x)-f'(a)$ pour tout $x \in I$. Par croissance de f', φ' est négative à gauche de a et positive à droite. Ainsi, φ est décroissante à gauche de a et croissante à droite, donc positive sur I tout entier puisque $\varphi(a) = 0$.

(iii) $\Longrightarrow$ (i) Montrons que f est convexe sur I. Soient $x,y\in I$ et $\lambda\in[0,1]$. Posons $a=(1-\lambda)x+\lambda y$. Pour tout $t\in I$: $f(t)\geqslant f'(a)(t-a)+f(a)$, donc :

$$\begin{aligned}(1-\lambda)f(x)+\lambda f(y)&\geqslant(1-\lambda)\big(f'(a)(x-a)+f(a)\big)+\lambda\big(f'(a)(y-a)+f(a)\big)\\&=f'(a)\Big((1-\lambda)x+\lambda y-a\Big)+f(a)=f\big((1-\lambda)x+\lambda y\big).\end{aligned}$$

Exemple La fonction logarithme est concave sur $\mathbb{R}_+^*$ car sa dérivée seconde $x\longmapsto -\dfrac{1}{x^2}$ est négative, donc pour tous $x,y>0$ et $\lambda=\dfrac{1}{2}$ dans la définition de la concavité : $\ln\dfrac{x+y}{2}\geqslant\dfrac{1}{2}\ln x+\dfrac{1}{2}\ln y$.

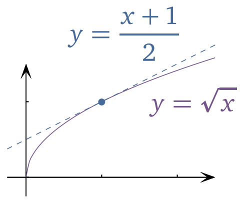
De nombreuses inégalités de convexité découlent de la définition ou des résultats qui précèdent. Rappelons-en trois qu'il faut connaître et comprendre graphiquement. À l'exception de la minoration du sinus, ces inégalités illustrent toutes l'idée que le graphe d'une fonction convexe (resp. concave) est au-dessus (resp. en dessous) de ses tangentes. La minoration du sinus exprime quant à elle l'idée que la corde du sinus joignant les points d'abscisses 0 et $\dfrac{\pi}{2}$ a pour équation $y=\dfrac{2x}{\pi}$.

$$\forall x\in\mathbb{R},\quad \mathrm{e}^x\geqslant 1+x.$$

$$\forall x>-1,\quad \ln(1+x)\leqslant x.\qquad \forall x>0,\quad \ln x\leqslant x-1.$$

$$\forall x\in\left[0,\frac{\pi}{2}\right],\quad \frac{2x}{\pi}\leqslant\sin x\leqslant x.$$

Exemple Pour tout $x>0$: $\sqrt{x}\leqslant\dfrac{x+1}{2}$.



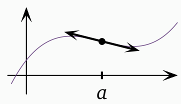
Démonstration La fonction $x\longmapsto\sqrt{x}$ est concave sur $\mathbb{R}_+^*$ car sa dérivée $x\longmapsto\dfrac{1}{2\sqrt{x}}$ y est décroissante. Son graphe est situé sous sa tangente en 1, d'équation $y=\dfrac{x+1}{2}$.

Avant de poursuivre, revenons rapidement sur la définition de la convexité. On dit qu'une fonction $f:I\longrightarrow\mathbb{R}$ est *strictement convexe sur I* si pour tous $x,y\in I$ distincts et pour tout $\lambda\in\,]0,1[$: $f\big((1-\lambda)x+\lambda y\big)<(1-\lambda)f(x)+\lambda f(y)$. L'intuition géométrique est la même qu'en cas de convexité simple à ceci près qu'une fonction strictement convexe ne peut pas être affine sur un intervalle $[a,b]$ avec $a<b$. Son graphe est toujours strictement au-dessus de ses cordes.

Les résultats précédents s'étendent bien au cadre de la convexité stricte. L'inégalité des pentes devient simplement stricte et une fonction $f:I\longrightarrow\mathbb{R}$ est strictement convexe si et seulement si pour tout $a\in I$, la fonction $x\longmapsto\dfrac{f(x)-f(a)}{x-a}$ est strictement croissante sur $I\setminus\{a\}$. Enfin, une fonction dérivable $f\in\mathscr{D}(I,\mathbb{R})$ est strictement convexe si et seulement si sa dérivée f' est strictement croissante, si et seulement si pour tout $a\in I$, son graphe est situé au-dessus de sa tangente en a avec $\big(a,f(a)\big)$ pour seul point de contact. Par exemple, la fonction exponentielle est strictement convexe sur $\mathbb{R}$ car sa dérivée est strictement croissante, donc dans l'inégalité de convexité $\mathrm{e}^x\geqslant 1+x$, l'égalité n'est atteinte que pour $x=0$.

Définition-théorème (Point d'inflexion) Soient $f:I\longrightarrow\mathbb{R}$ une fonction et $a\in\mathring{I}$. On dit que f *possède un point d'inflexion en a* si f est convexe au voisinage de a à gauche et concave au voisinage de a à droite — ou l'inverse.

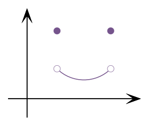
Si f est deux fois dérivable sur I, f possède un point d'inflexion en a si et seulement si f'' s'annule en a et est positive au voisinage de a à gauche et négative au voisinage de a à droite — ou l'inverse.



Exemple Soient $a\in\mathbb{R}^*$ et $b,c\in\mathbb{R}$. La fonction $x\longmapsto ax^3+bx^2+cx+d$ possède un et un seul point d'inflexion, à savoir en $-\dfrac{b}{3a}$, car sa dérivée seconde $x\longmapsto 6ax+2b$ s'annule en changeant de signe en $-\dfrac{b}{3a}$ et ne le fait nulle part ailleurs.

Théorème (Convexe implique continue) Pour une fois, l'intervalle I est supposé ouvert. Soit $f:I\longrightarrow\mathbb{R}$ une fonction. Si f est convexe sur I, alors f est continue sur I.

Attention ! Il est essentiel que l'intervalle I soit ouvert. La fonction smiley ci-contre est convexe mais pas continue aux bornes. Vous noterez en passant qu'une fonction convexe n'est pas forcément dérivable sur son ensemble de définition, pensez à la fonction valeur absolue !



Démonstration Fixons $a\in I$ et contentons-nous de montrer que f est continue à gauche en a. L'intervalle I étant ouvert, nous pouvons nous donner un élément $b\in I$ pour lequel $a<b$. Par convexité de f, la fonction $x\overset{\varphi}{\longmapsto}\dfrac{f(x)-f(a)}{x-a}$ est croissante sur $I\setminus\{a\}$, donc croissante sur $I\cap\,]-\infty,a[$ et majorée par $\varphi(b)$. Elle possède

ainsi une limite finie à gauche en a d'après d'après le théorème de la limite monotone. Il en découle que : $f(x) = \dfrac{f(x)-f(a)}{x-a} \times (x-a) + f(a) \underset{x \to a^-}{\longrightarrow} f(a)$, i.e. que f est continue à gauche en a. ■

Théorème (Inégalité de Jensen) Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe.
Pour tous $x_1, \ldots, x_n \in I$ et $\lambda_1, \ldots, \lambda_n \in [0,1]$ pour lesquels $\lambda_1 + \ldots + \lambda_n = 1$: $\quad f\left(\sum_{k=1}^n \lambda_k x_k\right) \leqslant \sum_{k=1}^n \lambda_k f(x_k)$.

Démonstration Par récurrence sur n. Pour $n = 2$, le résultat est exactement la définition de la convexité.

- **Initialisation :** Rien à démontrer pour $n = 1$.
- **Hérédité :** Soit $n \in \mathbb{N}^*$. Faisons l'hypothèse que pour tous $x_1, \ldots, x_n \in I$ et $\lambda_1, \ldots, \lambda_n \in [0,1]$ pour lesquels $\lambda_1 + \ldots + \lambda_n = 1$: $\quad f\left(\sum_{k=1}^n \lambda_k x_k\right) \leqslant \sum_{k=1}^n \lambda_k f(x_k)$.

  Fixons $x_1, \ldots, x_{n+1} \in I$ et $\lambda_1, \ldots, \lambda_{n+1} \in [0,1]$ pour lesquels $\lambda_1 + \ldots + \lambda_{n+1} = 1$. Si $\lambda_{n+1} = 0$, l'inégalité à prouver au rang $n+1$ découle trivialement de l'hypothèse de récurrence. Supposons donc $\lambda_{n+1} > 0$ et remplaçons la moyenne coefficientée des $n+1$ réels $x_1, \ldots, x_{n+1}$ par une moyenne coefficientée de n réels auxquels nous pourrons appliquer l'hypothèse de récurrence.

  Posons pour cela $\lambda'_n = \lambda_n + \lambda_{n+1} \in [0,1]$ et $x'_n = \dfrac{\lambda_n}{\lambda'_n}\, x_n + \dfrac{\lambda_{n+1}}{\lambda'_n}\, x_{n+1} \in I$. Sachant que $\lambda_1 + \ldots + \lambda_{n-1} + \lambda'_n = 1$, appliquons l'hypothèse de récurrence aux n réels $x_1, \ldots, x_{n-1}, x'_n$ et aux n coefficients $\lambda_1, \ldots, \lambda_{n-1}, \lambda'_n$:
  $$f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) = f\big(\lambda_1 x_1 + \ldots + \lambda_{n-1} x_{n-1} + \lambda'_n x'_n\big) \overset{\text{HDR}}{\leqslant} \lambda_1 f(x_1) + \ldots + \lambda_{n-1} f(x_{n-1}) + \lambda'_n f(x'_n).$$
  Pour finir, par convexité de f : $\quad f(x'_n) = f\left(\dfrac{\lambda_n}{\lambda'_n}\, x_n + \dfrac{\lambda_{n+1}}{\lambda'_n}\, x_{n+1}\right) \leqslant \dfrac{\lambda_n}{\lambda'_n}\, f(x_n) + \dfrac{\lambda_{n+1}}{\lambda'_n}\, f(x_{n+1})$, $\quad$ donc comme voulu : $\quad f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) \leqslant \sum_{k=1}^{n+1} \lambda_k f(x_k)$. ■

Exemple Pour tous $x_1, \ldots, x_n > 0$:
$$\underbrace{\frac{n}{\displaystyle\sum_{k=1}^n \frac{1}{x_k}}}_{\substack{\text{Moyenne}\\ \text{harmonique}}} \leqslant \underbrace{\sqrt[n]{\prod_{k=1}^n x_k}}_{\substack{\text{Moyenne}\\ \text{géométrique}}} \leqslant \underbrace{\frac{1}{n}\sum_{k=1}^n x_k}_{\substack{\text{Moyenne}\\ \text{arithmétique}}}.$$
L'inégalité de droite s'appelle l'*inégalité arithmético-géométrique*.

Démonstration Comme $\displaystyle\sum_{k=1}^n \frac{1}{n} = 1$, l'inégalité de Jensen appliquée à la fonction concave $x \longmapsto \ln x$ s'écrit :
$$\ln\left(\sum_{k=1}^n \frac{1}{n}\, x_k\right) \geqslant \sum_{k=1}^n \frac{1}{n}\, \ln x_k = \ln \sqrt[n]{\prod_{k=1}^n x_k},$$
puis $\displaystyle\frac{1}{n}\sum_{k=1}^n x_k \geqslant \sqrt[n]{\prod_{k=1}^n x_k}$ par croissance de l'exponentielle. Pour l'inégalité de gauche, remplacer x_k par $\dfrac{1}{x_k}$ pour tout $k \in [\![1, n]\!]$ dans l'inégalité de droite.

Rappelons à présent que $0^p = 0$ pour tout $p > 0$, ce qui rend la fonction $x \longmapsto x^p$ continue sur $\mathbb{R}_+$ tout entier alors qu'elle est au départ définie sur $\mathbb{R}_+^*$ seulement par la relation $x^p = \mathrm{e}^{p \ln x}$. Les inégalités de Hölder et Minkowski énoncées ci-dessous dans $\mathbb{R}^n$ n'ont pas beaucoup d'importance en elles-mêmes, mais leur généralisation aux espaces vectoriels de fonctions, où l'on remplace les sommes par des intégrales, est au contraire très fructueuse.

Théorème (Inégalités de Hölder et Minkowski)

(i) **Norme p d'un vecteur de $\mathbb{R}^n$:** Soit $p > 1$.
Pour tout $X \in \mathbb{R}^n$, on appelle *norme p de X* le réel positif $\|X\|_p = \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}$.
Pour tous $X \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$: $\quad \|\lambda X\|_p = |\lambda| . \|X\|_p \quad$ et $\quad \|X\|_p = 0 \iff X = 0_{\mathbb{R}^n}$.

(ii) **Inégalité de Hölder :** Soient $p, q > 1$ deux réels pour lesquels $\dfrac{1}{p} + \dfrac{1}{q} = 1$.
Pour tous $X, Y \in \mathbb{R}^n$: $\quad \left|\sum_{k=1}^n x_k y_k\right| \leqslant \|X\|_p\, \|Y\|_q$.

(iii) **Inégalité de Minkowski :** Soit $p > 1$. Pour tous $X, Y \in \mathbb{R}^n$: $\quad \|X + Y\|_p \leqslant \|X\|_p + \|Y\|_p$.

La norme 2 n'est jamais qu'une généralisation à n réels du concept de norme que vous calculez en géométrie du plan et de l'espace à partir des coordonnées dans une base orthonormale.

L'inégalité de Minkowski n'est finalement qu'une autre manière de nommer l'inégalité triangulaire dans le contexte des normes p. L'inégalité de Hölder est plus mystérieuse au premier abord, mais vous la connaissez dans le cas où $p=q=2$, c'est juste l'inégalité de Cauchy-Schwarz !

Démonstration

(i) Pour tous $X \in \mathbb{R}^n$ et $\lambda \in \mathbb{R}$: $\quad \|\lambda X\|_p = \left(\sum_{k=1}^n |\lambda x_k|^p\right)^{\frac{1}{p}} = \left(|\lambda|^p \sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} = |\lambda| \,.\, \|X\|_p.$

(ii) Montrons d'abord que $xy \leqslant \dfrac{x^p}{p} + \dfrac{y^q}{q}$ pour tous $x, y \geqslant 0$. Il nous suffit de le montrer pour $x, y > 0$. Or par concavité de la fonction logarithme appliquée aux réels x^p et y^q, sachant que $\dfrac{1}{p} + \dfrac{1}{q} = 1$:

$$\ln\left(\frac{x^p}{p} + \frac{y^q}{q}\right) \geqslant \frac{1}{p}\ln(x^p) + \frac{1}{q}\ln(y^q) = \ln(xy), \qquad \text{et on conclut par croissance de l'exponentielle.}$$

Montrons maintenant l'inégalité de Hölder. Fixons $X, Y \in \mathbb{R}^n$.

— Supposons dans un premier temps X et Y *unitaires* au sens où $\|X\|_p = \|Y\|_q = 1$. Pour tout $k \in [\![1,n]\!]$:
$$|x_k y_k| \leqslant \frac{|x_k|^p}{p} + \frac{|y_k|^q}{q}, \quad \text{donc} \left|\sum_{k=1}^n x_k y_k\right| \leqslant \sum_{k=1}^n |x_k y_k| \leqslant \frac{\|X\|_p^p}{p} + \frac{\|Y\|_q^q}{q} = \frac{1}{p} + \frac{1}{q} = 1.$$

— Si $\|X\| = 0$, alors $\sum_{k=1}^n |x_k|^p = 0$, donc $x_k = 0$ pour tout $k \in [\![1,n]\!]$ car on somme des réels positifs et l'inégalité de Hölder est triviale dans ce cas. Même chose si $\|Y\| = 0$.

— Enfin, si $\|X\| > 0$ et $\|Y\| > 0$, on se ramène au cas unitaire grâce à (i), car $\left\|\dfrac{X}{\|X\|_p}\right\|_p = \dfrac{\|X\|_p}{\|X\|_p} = 1$ et $\left\|\dfrac{Y}{\|Y\|_q}\right\|_q = 1$, donc : $\quad \left|\sum_{k=1}^n \dfrac{x_k}{\|X\|_p} \times \dfrac{y_k}{\|Y\|_q}\right| \leqslant \left\|\dfrac{X}{\|X\|_p}\right\|_p \left\|\dfrac{Y}{\|Y\|_q}\right\|_q = 1 \quad$ et c'est fini.

(iii) Pour commencer, $\dfrac{1}{p} + \dfrac{1}{q} = 1$ donc $\dfrac{1}{q} = \dfrac{p-1}{p}$, donc $(p-1)q = p$. Ensuite, pour tous $X, Y \in \mathbb{R}^n$:

$$\begin{aligned}
\|X+Y\|_p^p &= \sum_{k=1}^n |x_k + y_k|^p = \sum_{k=1}^n |x_k + y_k| \,.\, |x_k + y_k|^{p-1} \leqslant \sum_{k=1}^n |x_k| \,.\, |x_k + y_k|^{p-1} + \sum_{k=1}^n |y_k| \,.\, |x_k + y_k|^{p-1} \\
&\overset{\text{Hölder}}{\leqslant} \|X\|_p \left(\sum_{k=1}^n |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}} + \|Y\|_p \left(\sum_{k=1}^n |x_k + y_k|^{(p-1)q}\right)^{\frac{1}{q}} \\
&= \|X\|_p \left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{q}} + \|Y\|_p \left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{q}} = \|X\|_p \|X+Y\|_p^{\frac{p}{q}} + \|Y\|_p \|X+Y\|_p^{\frac{p}{q}}.
\end{aligned}$$

Pour finir, l'inégalité de Minkowski est triviale si $\|X+Y\| = 0$. Dans le cas contraire, on divise par $\|X+Y\|_p^{\frac{p}{q}}$:
$\|X\|_p + \|Y\|_p \geqslant \|X+Y\|_p^{p-\frac{p}{q}} = \|X+Y\|_p^{p\left(1-\frac{1}{q}\right)} = \|X+Y\|_p.$ ■