DERIVATION

PLAN

I : Dérivée
- 1) Définition
- 2) Opérations
- 3) Dérivées successives
- 4) Théorème de Rolle
- 5) Théorème des accroissements finis

II : Fonctions convexes
- 1) Définition
- 2) Convexité et dérivabilité

I : Dérivée

1– Définition

DEFINITION

f définie sur un intervalle I est dérivable en x_0 élément de I s'il existe un réel noté $f'(x_0)$ tel que :

$$f(x_0 + h) = f(x_0) + hf'(x_0) + o(h)$$

Cette définition est équivalente à :

$$\lim_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} = f\,'(x_0)$$

La droite d'équation $y = f\,'(x_0)(x - x_0) + f(x_0)$ est la droite tangente au graphe représentatif de f, au point d'abscisse x_0. Au lieu de $f\,'$, on trouve également les notations Df ou $\dfrac{\mathrm{d}f}{\mathrm{d}x}$. (La définition précédente s'applique aussi aux fonctions à valeurs complexes. Si f est à valeurs complexes, de parties réelle g et imaginaire h, il résulte du calcul des limites d'une telle fonction en séparant précisément partie réelle et imaginaire que $f\,' = g' + ih'$).

La définition est également équivalente à l'existence d'une fonction φ continue en x_0 telle que :

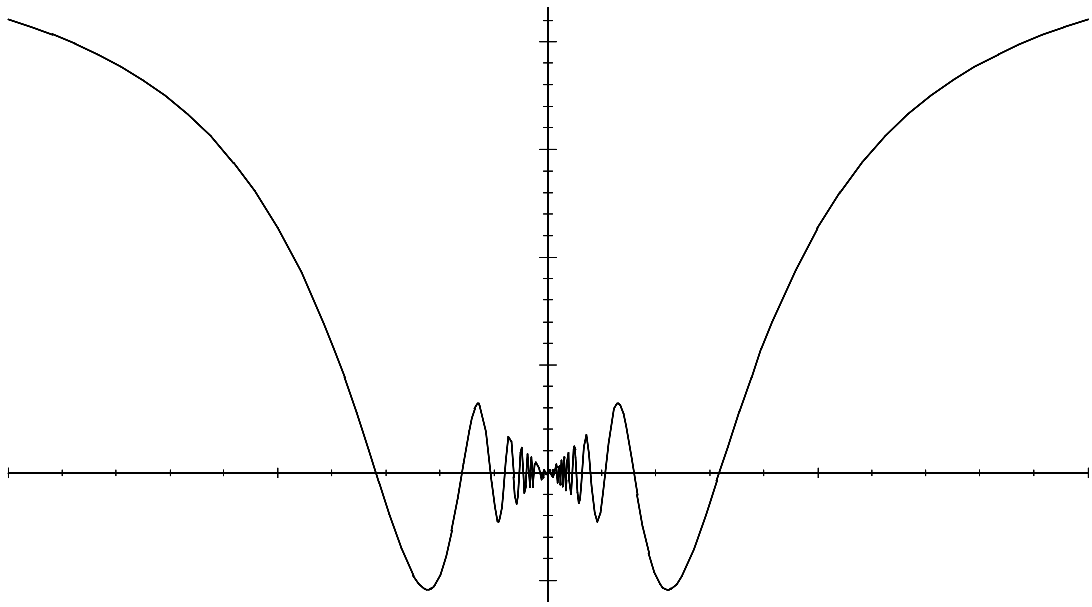
$$f(x) = f(x_0) + (x - x_0)\,\varphi(x)$$

avec $\varphi(x_0) = f\,'(x_0)$. En dehors de x_0, $\varphi(x)$ est égal au taux d'accroissement $\dfrac{f(x) - f(x_0)}{x - x_0}$.

Il résulte immédiatement de la définition qu'une fonction dérivable est continue. La réciproque est fausse comme le prouve l'exemple de $|x|$ en 0.

Si on se limite à $h > 0$, on parle de dérivée à droite. Si l'on se limite à $h < 0$, on parle de dérivée à gauche. Si f est dérivable à droite et à gauche de x_0 et si les deux dérivées sont égales, alors f est dérivable en x_0.

La fonction $x \to |x|$ est dérivable à droite et à gauche de 0. Il existe des fonctions continues n'admettant aucune dérivée à droite et à gauche de 0, par exemple $x \to x \sin(\frac{1}{x})$.



Il existe des fonctions continues dérivables en aucun point, mais la présentation d'un contre-exemple dépasse le niveau de première année.

2– Opérations

Les résultats relatifs à la somme, au produit, au quotient de fonctions dérivables étant censés être bien connus, nous nous limiterons à :

<u>a) Composition</u>

Soit f dérivable en x_0 et g dérivable en f(x_0). Alors $g \circ f$ est dérivable en x_0 et :

$$(g \circ f)'(x_0) = g'(f(x_0)).f\,'(x_0)$$

En effet, il existe deux fonctions φ et ψ, continue en x_0 et $f(x_0) = y_0$ respectivement telles que :

$$f(x) = f(x_0) + (x - x_0)\ \varphi(x) \text{ et } f\,'(x_0) = \varphi(x_0)$$

$$g(y) = g(y_0) + (y - y_0)\ \psi(y) \text{ et } g\,'(y_0) = \psi(y_0)$$

Donc, en prenant $y = f(x)$:

$$\begin{aligned} g(f(x)) &= g(f(x_0)) + (f(x) - f(x_0))\ \psi(f(x)) \\ &= g(f(x_0)) + (x - x_0)\ \varphi(x)\psi(f(x)) \end{aligned}$$

Or la fonction $\varphi(x)\psi(f(x))$ est continue en x_0, donc $g \circ f$ est dérivable en x_0 et :

$$(g \circ f)'(x_0) = \varphi(x_0)\psi(f(x_0)) = f\,'(x_0)\ g'(f(x_0))$$

La règle de dérivation d'une fonction composée se note agréablement en physique, où seules les variables portent un nom et non les fonctions elles-mêmes. Supposons une quantité E dépendant de la position z d'un mobile. On a alors E = E(z). Supposons que z dépendent du temps t de sorte que $z = z(t)$ et que E = E(z(t)) = E(t) pour abréger. On remarquera que cette dernière notation est invalide en mathématique à cause d'une ambiguïté. E(3) désigne-t-il la valeur de E pour $z = 3$ ou pour $t = 3$. Cette ambiguïté n'existe pas en physique où l'on demandera E(3 m) ou E(3 s), l'unité appliquée aux variables levant alors l'ambiguïté. La règle de dérivation des fonctions composées se note alors :

$$\frac{\mathrm{dE}}{\mathrm{d}t} = \frac{\mathrm{dE}}{\mathrm{d}z}\frac{\mathrm{d}z}{\mathrm{d}t}$$

On notera que cette notation "fractionnaire" n'est valide qu'au premier ordre, puisqu'on s'exercera à vérifier que :

$$\frac{\mathrm{d}^2\mathrm{E}}{\mathrm{d}t^2} = \frac{\mathrm{d}^2\mathrm{E}}{\mathrm{d}z^2}\left(\frac{\mathrm{d}z}{\mathrm{d}t}\right)^2 + \frac{\mathrm{dE}}{\mathrm{d}z}\frac{\mathrm{d}^2z}{\mathrm{d}t^2}$$

<u>b) Reciproque</u>

Soit f continue sur un intervalle I, bijective sur I, dérivable en x_0 et telle que $f\,'(x_0) \neq 0$. Alors f^{-1} est dérivable en $y_0 = f(x_0)$ et :

$$(f^{-1})'(y_0) = \frac{1}{f\,'(x_0)} = \frac{1}{f\,'(f^{-1}(y_0))}$$

En effet, il existe une fonction φ, continue en x_0 telle que

$$f(x) = f(x_0) + (x - x_0)\,\varphi(x) \text{ et } f\,'(x_0) = \varphi(x_0)$$

On souhaite écrire $f^{-1}(y) = x$ avec $y = f(x)$, or :

$$x = x_0 + \frac{y - y_0}{\varphi(x)} \Rightarrow f^{-1}(y) = f^{-1}(y_0) + \frac{y - y_0}{\varphi(f^{-1}(y))}$$

La fonction $y \to \varphi(f^{-1}(y))$ est continue en y_0 comme composée de fonctions continues, et sa valeur en y_0 est $\varphi(x_0) = f\,'(x_0)$ qui est non nul. Elle est donc définie au voisinage de y_0. f^{-1} est donc dérivable en y_0 de dérivée $(f^{-1})'(y_0) = \dfrac{1}{\varphi(f^{-1}(y_0))} = \dfrac{1}{f\,'(x_0)} = \dfrac{1}{f\,'(f^{-1}(y_0))}$.

<u>*EXEMPLES*</u> :

Pour $f(x) = \ln(x)$, de dérivée $\dfrac{1}{x}$, on obtient : $(e^x)' = \dfrac{1}{\frac{1}{e^x}} = e^x$.

Pour $f(x) = \sin(x)$ sur $[-\pi/2, \pi/2]$, on obtient : $(\arcsin(x))' = \dfrac{1}{\cos(\arcsin(x))} = \dfrac{1}{\sqrt{1-x^2}}$ pour $x \in\ ]-1,1[$

De même, $(\arccos(x))' = -\dfrac{1}{\sqrt{1-x^2}}$

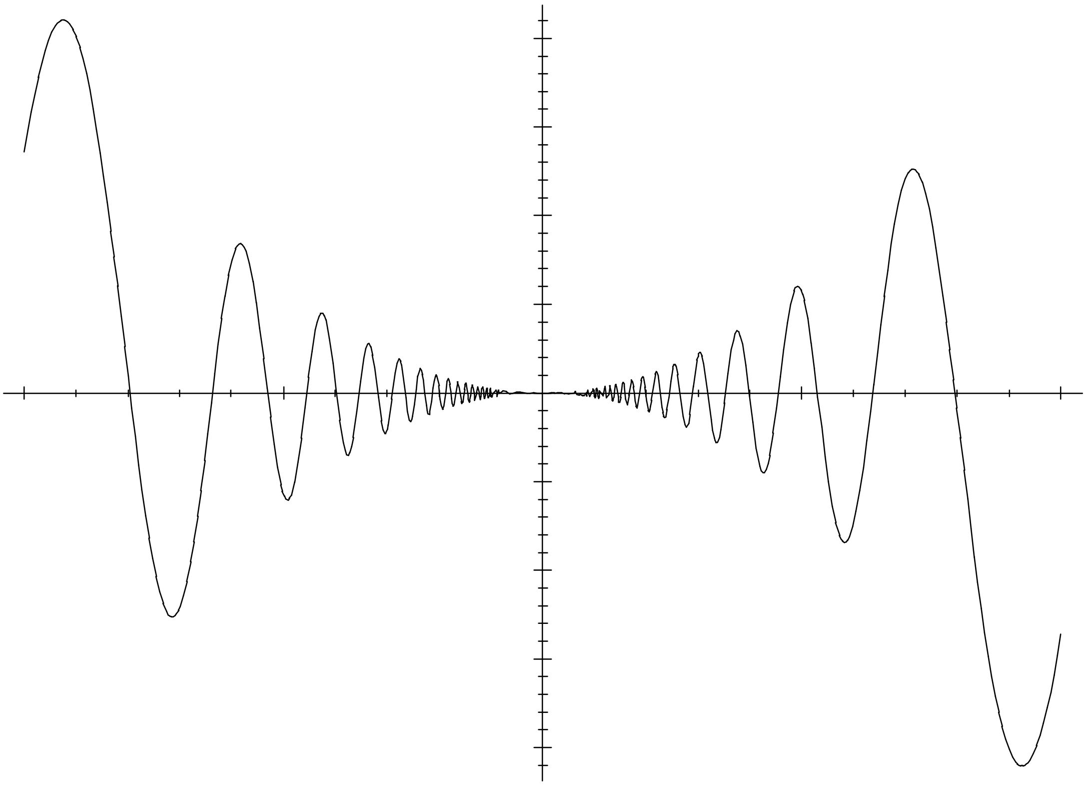
Pour $f(x) = \tan(x)$ sur $]-\pi/2,\pi/2[$, on obtient : $(\arctan(x))' = \dfrac{1}{1 + \tan^2(\arctan(x))} = \dfrac{1}{1 + x^2}$

3– Dérivées successives

Si f est dérivable sur un intervalle I, on peut définir la fonction dérivée $f\,'$, et se poser la question de savoir si elle est elle–même continue ou dérivable. Si c'est le cas, on peut définir sa dérivée f'' , etc ...

On note $f^{(k)}$ ou D^kf ou $\dfrac{\mathrm{d}^kf}{\mathrm{d}x^k}$ la dérivée d'ordre k.

On remarquera que $f\,'$ peut ne pas être continue. Un exemple est donné par la fonction $f(x) = x^2 \sin \dfrac{1}{x}$ dont la dérivée n'est pas continue en 0. On a en effet, $\dfrac{f(x) - f(0)}{x - 0} = x \sin \dfrac{1}{x}$ qui tend vers $0 = f\,'(0)$, cependant que $f\,'(x) = 2x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}$ n'a pas de limite en 0. Ci-dessous le graphe de f entre -1 et 1, dans un repère non orthonormé. Il y a clairement une tangente horizontale en 0 (le graphe est compris entre les deux paraboles $y = \pm\, x^2$), mais les tangentes au graphe aux points d'intersection avec Ox ont des pentes qui tendent vers $\pm\, 1$.



On note $C^n(I)$ l'espace vectoriel des fonctions n fois continûment dérivables sur I. Montrons que, si f et g sont C^n, il en est de même de $f + g$, de fg, de $\dfrac{f}{g}$ (à condition que g ne s'annule pas), de $g \circ f$, et de f^{-1} (à condition que f soit bijective et que sa dérivée ne s'annule pas).

<u>a) Somme</u>
Il est trivial de vérifier par récurrence que $(f + g)^{(n)} = f^{(n)} + g^{(n)}$

<u>b) Produit</u>
On dispose d'une formule donnant la dérivée $n^{\text{ème}}$ d'un produit.

FORMULE DE LEIBNIZ
Soit u et v deux fonctions n fois dérivables. Alors uv est n fois dérivable et :

$$(uv)^{(n)} = \sum_{p=0}^{n} \binom{n}{p} u^{(p)}.v^{(n-p)}$$

où l'on convient que $u^{(0)} = u$. $\binom{n}{p}$ désigne le coefficient binomial $\dfrac{n!}{p!(n-p)!}$ si $0 \leq p \leq n$ et 0 sinon.

Démonstration :
Elle se fait par récurrence sur n. C'est évidemment vérifié pour $n = 0$ et pour $n = 1$, pour lequel on reconnaît : $(uv)' = u'v + uv'$.
Si la formule est vraie au rang n et que les fonctions sont $n+1$ fois dérivables, on voit que $(uv)^{(n)}$ est dérivable et de dérivée :

$$(uv)^{(n+1)} = \sum_{p=0}^{n} \binom{n}{p} (u^{(p+1)}\, v^{(n-p)} + u^{(p)}\, v^{(n-p+1)}) = \sum_{p=0}^{n} \binom{n}{p} u^{(p+1)}\, v^{(n-p)} + \sum_{p=0}^{n} \binom{n}{p} u^{(p)}\, v^{(n-p+1)}$$

$$= \sum_{p=0}^{n+1} \left[\binom{n}{p-1} + \binom{n}{p}\right] u^{(p)}\, v^{(n-p+1)} \text{ en changeant d'indice } p+1 \to p \text{ dans la première somme}$$

$$= \sum_{p=0}^{n+1} \binom{n+1}{p} u^{(p)}\, v^{(n-p+1)}$$

(On a utilisé le fait que $\binom{n}{p} = 0$ si $p < 0$ ou $p > n$)
La démonstration, ainsi que la nature de la formule, est donc comparable à celle du développement du binôme de Newton. Ce n'est pas un hasard : la formule de Leibniz permet d'en déduire la formule du binôme de Newton :
Il suffit de prendre $u(x) = e^{ax}$ et $v(x) = e^{bx}$ et d'appliquer la formule de Leibniz en $x = 0$.

c) Quotient

Il n'existe pas de formule générale simple donnant la dérivée $n^{\text{ème}}$ d'un quotient, mais on peut montrer par récurrence que cette dérivée existe. C'est le cas pour $n = 1$ et supposons que ce le soit pour $n-1$.
On a alors $\left(\dfrac{f}{g}\right)' = \dfrac{f'g - fg'}{g^2}$. Or :

$f \in C^n(I)$ et $g \in C^n(I) \Rightarrow f \in C^{n-1}(I)$, $g \in C^{n-1}(I)$, $f' \in C^{n-1}(I)$, $g' \in C^{n-1}(I)$
$\Rightarrow$ $f'g \in C^{n-1}(I)$, $fg' \in C^{n-1}(I)$ et $g^2 \in C^{n-1}(I)$ en utilisant le résultat prouvé sur le produit
$\Rightarrow$ $f'g - fg' \in C^{n-1}(I)$ et $g^2 \in C^{n-1}(I)$ en utilisant le résultat sur la somme
$\Rightarrow$ $\dfrac{f'g - fg'}{g^2} \in C^{n-1}(I)$ en appliquant l'hypothèse de récurrence
$\Rightarrow$ $\dfrac{f}{g} \in C^n(I)$

d) Composition

Il n'existe pas non plus de formule générale simple donnant la dérivée $n^{\text{ème}}$ d'une composée de fonctions, mais on peut montrer par récurrence que cette dérivée existe. Le raisonnement est comparable au précédent. On a : $(g \circ f)' = (g' \circ f)f'$. On note $J = f(I)$.

$f \in C^n(I)$ et $g \in C^n(J) \Rightarrow f \in C^{n-1}(I)$, $f' \in C^{n-1}(I)$, $g' \in C^{n-1}(J)$
$\Rightarrow$ $g' \circ f \in C^{n-1}(I)$, $f' \in C^{n-1}(I)$ en appliquant l'hypothèse de récurrence au rang $n-1$
$\Rightarrow$ $(g' \circ f)f' \in C^{n-1}(I)$ en appliquant le résultat sur le produit.
$\Rightarrow$ $g \circ f \in C^n(I)$

e) Réciproque

De même, la formule $(f^{-1})' = \dfrac{1}{f' \circ f^{-1}}$ permet de voir que, si f est de classe C^n, et si f' ne s'annule pas, alors f^{-1} est de classe C^n. Cela se montre également par récurrence :

$f \in C^n(I)$

$\Rightarrow$ $f' \in C^n(I)$, $f^{-1} \in C^{n-1}(I)$ en appliquant l'hypothèse de récurrence

$\Rightarrow$ $f' \circ f^{-1} \in C^{n-1}(I)$ en appliquant le résultat sur la composée de fonction

$\Rightarrow$ $\dfrac{1}{f' \circ f^{-1}} \in C^{n-1}(I)$ en appliquant le résultat sur le quotient

$\Rightarrow$ $f^{-1} \in C^n(I)$

4– Théorème de Rolle

Ce théorème a d'abord été énoncé au XVIIème sous la forme suivante : entre deux racines d'un polynôme P, il y a une racine de sa dérivée P'.

THEOREME DE ROLLE

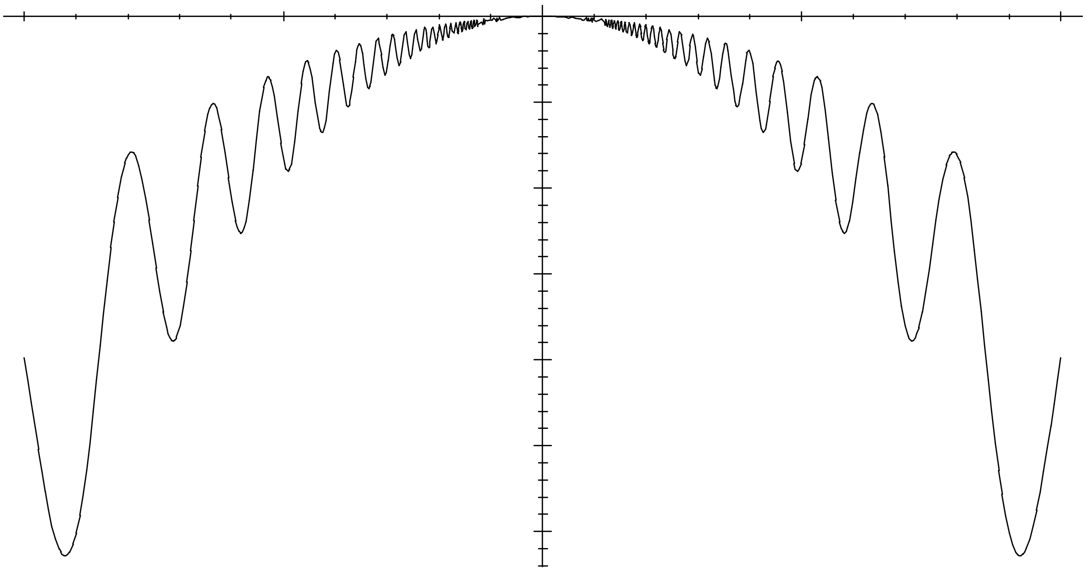
Soit f une fonction continue sur [a,b] à valeur réelles, dérivable sur]a,b[, telle que f(a) = f(b). Alors il existe un élément c de]a,b[, tel que f '(c) = 0.

Démonstration :

f étant continue sur un segment, admet un maximum et un minimum (toute la difficulté du théorème de Rolle repose sur ce résultat, montré dans le chapitre *Fonctions*, fichier FONCTION.PDF). Si ceux–ci se trouvent en a et b, cela signifie que f est constante, et alors c peut être pris de façon quelconque. Sinon, l'un des deux extrema se situe à l'intérieur de l'intervalle, en c. En un tel point, la dérivée s'annule. En effet, si par exemple c est un maximum, on a :

$$\left\{\begin{array}{l} \text{pour } x < c,\ \dfrac{f(x)-f(c)}{x-c} \geq 0 \Rightarrow f'(c) \geq 0 \\ \text{pour } x > c,\ \dfrac{f(x)-f(c)}{x-c} \leq 0 \Rightarrow f'(c) \leq 0 \end{array}\right. \quad \Rightarrow\ f'(c) = 0.$$

Il est faux de croire que, si c est un maximum, alors f est croissante à gauche de c, puis décroissante après. $f(c)$ est certes la valeur maximale, mais f peut ne pas être monotone, ni à gauche, ni à droite. Prendre par exemple $f(x) = -\dfrac{x^2}{2} - x^2 \sin^2 \dfrac{1}{x}$.



Le théorème de Rolle ne s'applique pas aux fonctions à valeurs complexes. Par exemple $f(x) = e^{ix}$, définie sur $[0, 2\pi]$, vérifie $f(0) = f(2\pi) = 1$, mais la dérivée de f, égale à ie^{ix} ne s'annule pas.

5– Théorème des accroissements finis

THEOREME DES ACCROISSEMENTS FINIS

Soit f une fonction continue sur [a,b] à valeurs réelles, dérivable sur]a,b[. Alors il existe un élément c de]a,b[, tel que $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.

<u>Démonstration</u> :

Il suffit de se ramener au théorème de Rolle. Pour cela, il suffit de poser, pour x dans $[a, b]$:

$$g(x) = f(x) - \left[f(a) + \frac{f(b) - f(a)}{b - a}(x - a) \right]$$

g représente l'écart entre f et la fonction affine qui coïncide avec f aux bornes de l'intervalle. g est continue sur $[a, b]$, dérivable sur $]a, b[$, et $g(a) = 0 = g(b)$. Donc il existe c élément de $]a, b[$ tel que $g'(c) = 0$, ce qui donne le résultat cherché.

Pas plus que le théorème de Rolle, ce théorème ne s'applique aux fonctions à valeurs complexes.

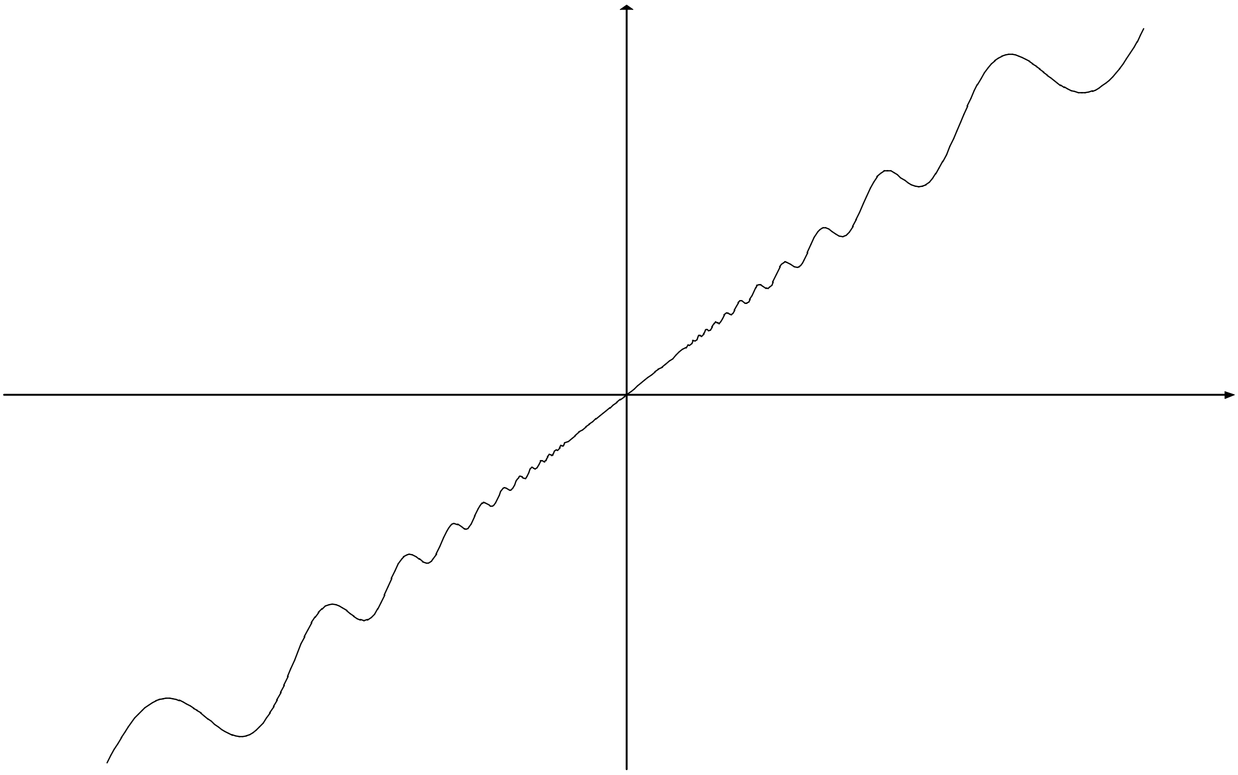
APPLICATIONS :

<u>a) sens de variation d'une fonction à valeurs réelles</u> :

f croissante sur $]a, b[\Leftrightarrow f' \geq 0$.
f décroissante sur $]a, b[\Leftrightarrow f' \leq 0$
f constante sur $]a, b[\Leftrightarrow f' = 0$

Le sens $\Rightarrow$ découle d'un passage à la limite sur des taux d'accroissements de signe constant. Il peut être montré dès la classe de Première. La réciproque est admise en lycée. Elle utilise en effet le théorème des accroissements finis. Si f' est de signe constant (ou nul), il en est de même de tout taux d'accroissement $\dfrac{f(y) - f(x)}{y - x}$ puisque ce dernier est égal à $f'(c)$ avec c entre x et y.

Il est faux de croire que $f'(x_0) > 0 \Rightarrow f$ est strictement croissante sur un intervalle contenant x_0. Il suffit de la stricte positivité sur ***tout un intervalle***, mais la positivité en un point unique ne suffit pas. Considérer par exemple $f(x) = x + 2x^2 \sin \frac{1}{x}$ en 0. On a $f'(0) = 1$, mais f' n'est de signe constant dans aucun voisinage de 0.



Si f est dérivable et si $f' > 0$, f est strictement croissante. La réciproque est fausse. Il se peut que f soit strictement croissante et dérivable, et que f' s'annule. Il suffit de prendre $f(x) = x^3$. L'équivalence est la suivante. Notons Z l'ensemble des x où f' s'annule :

$$f \text{ strictement croissante} \Leftrightarrow f' \geq 0 \text{ et Z ne contient aucun intervalle }]a, b[\text{ avec } a < b$$

En effet, dire que f est croissante sans l'être strictement, c'est dire qu'il existe $x < y$ tel que $f(x) = f(y)$, ou encore que f est constante sur un intervalle, ou encore que f' s'annule sur un intervalle, ou enfin que Z contient un intervalle ouvert.

<u>b) Dérivation à une borne d'un intervalle</u>

f continue sur $[a, b]$, dérivable sur $]a, b[$, et $\lim_{x \to a} f'(x) = l \Rightarrow f$ dérivable à droite de a et $f'(a) = l$. On a un résultat analogue à gauche de a. Si l est infini, on peut conclure qu'il existe une tangente verticale au graphe de f, au point d'abscisse a.

On a en effet, pour $x > a$, $\frac{f(x) - f(a)}{x - a} = f'(c)$ avec $a < c < x$. c est une fonction implicite de x. Lorsque x tend vers a, c tend vers a, de sorte que $f'(c)$ tend vers l. Le taux d'accroissement admet donc une limite.

c) Inégalité des accroissements finis :
L'inégalité des accroissements finis s'énonce, sous les mêmes hypothèses que pour l'égalité, sous la forme :

$$|f\,'| \leq k \Rightarrow \left|\frac{f(b)-f(a)}{b-a}\right| \leq k$$

C'est évident puisqu'il existe c tel que $\left|\dfrac{f(b)-f(a)}{b-a}\right| = |f\,'(c)| \leq k$.

Cette inégalité est également valable, sous les mêmes hypothèses, pour les fonctions f à valeurs complexes. Considérons θ l'argument de $f(b) - f(a)$. Ce θ est tel que $e^{-i\theta}(f(b) - f(a))$ soit réel. Posons $g(t) = \mathrm{Re}(e^{-i\theta}f(t))$, de sorte que g est à valeurs réelles et que :

$$g(b) - g(a) = \mathrm{Re}(e^{-i\theta}(f(b) - f(a)) = e^{-i\theta}(f(b) - f(a))$$

D'où $|g(b) - g(a)| = |f(b) - f(a)|$

g est dérivable et :

$$|g\,'(t)| = |\mathrm{Re}(e^{i\theta} f\,'(t))| \leq |e^{i\theta} f\,'(t)| = |f\,'(t)| \leq k$$

$\Rightarrow$ $|g(b) - g(a)| \leq k\,|b - a|$ en appliquant l'inégalité des accroissements finis à g.

$\Rightarrow$ $|f(b) - f(a)| \leq k\,|b - a|$

Interprétation physique :
Si la variable est le temps t, $f(t)$ la position d'un mobile sur un axe, et $f\,'(t)$ est la vitesse instantanée du mobile. On a donc :

Si la vitesse instantanée est majorée par k, alors la vitesse moyenne aussi.

Ce résultat physique a l'air d'aller de soi, et on se demande pourquoi il en faudrait une justification mathématique reposant sur le théorème de accroissements finis, lui-même reposant sur le théorème de Rolle, reposant à son tour sur l'existence d'extrema, qui elle-même repose sur l'existence de borne supérieure et de borne inférieure, ces derniers étant des axiomes, tout cela pour arriver à un résultat physique évident.
En fait, l'inégalité des accroissements finis ne montre pas un résultat physique évident. Elle montre l'adéquation entre la réalité physique et les modèles qui ont été choisis pour décrire cette réalité. Il faut bien avoir conscience que ces modèles sont arbitraires et que rien ne permet d'assurer à coup sûr la correspondance entre une déduction mathématique et une expérience physique. Tant que les deux sont en accord, il n'y a pas de remise en cause à faire, mais rien n'assure qu'il en sera toujours ainsi. Qu'y a-t-il d'arbitraire dans les modèles physiques ? En voici des exemples :

❑ Le choix d'utiliser les éléments de $\mathbb{R}$ pour définir une position x, un instant t, une intensité de courant I et plus généralement d'utiliser $\mathbb{R}^3$ pour modéliser notre espace est difficilement justifiable. En effet, le développement décimal d'un réel s'écrit avec une infinité de chiffres. Cette exigence n'est d'aucune utilité en physique qui se contente de nombres décimaux (ayant un nombre finis de décimales), voire de nombres entiers si l'unité de mesure a été choisie suffisamment petite. De plus, une mesure physique n'est pas en soi un réel, ni un décimal, mais plutôt un intervalle dans lequel se trouve la mesure, qui possède toujours une incertitude. Nous devrions donc définir des opérations (somme, produit...) entre intervalles de décimaux plutôt qu'entre réels pour avoir une représentation plus fidèle de la réalité. La difficulté de mettre en œuvre de telles opérations, ainsi que la commodité consistant à disposer d'un nombre arbitraire et non limité de décimales, adaptable au progrès des précisions des mesures, fait qu'on utilise les réels.

❑ Le choix de la définition de la vitesse instantanée est également arbitraire. On a choisi comme définition de cette vitesse instantanée $V(t_0) = \lim\limits_{h \to 0} \dfrac{f(t_0 + h) - f(t_0)}{h}$, limite de la vitesse moyenne entre l'instant considéré et un autre instant, et donc égale à la dérivée usuelle. On aurait très bien pu choisir d'autres définitions plus raisonnables, par exemple $\lim\limits_{h \to 0} \dfrac{f(t_0 + h) - f(t_0 - h)}{2h}$, limite de la vitesse moyenne sur un intervalle centré en l'instant considéré. On notera qu'un mobile ayant une loi horaire $f(t) = |t|$ (modélisant par exemple un choc) n'a pas de vitesse définie en $t = 0$ dans le premier cas, mais a une vitesse nulle dans le second. On dispose donc de deux modèles radicalement différents et il faut choisir entre l'un et l'autre.
Il serait d'ailleurs plus juste encore de noter qu'étant dans l'impossibilité physique de définir un instant de manière infiniment précise, on devrait plutôt définir la vitesse comme $\lim\limits_{(h,k) \to (0,0)} \dfrac{f(t_0 + h) - f(t_0 - k)}{(h+k)}$. On montre en fait que cette définition signifie qu'une fonction f pour laquelle cette quantité est définie en tout point est dérivable au sens usuel, mais qu'en plus, sa dérivée est continue (autrement dit, f est C^1).

❑ Le fait de raisonner sur des fonctions C^1 et même C^∞ est très courant en physique, lorsque l'évolution des processus est jugée suffisamment régulière. Notons cependant depuis quelques années le développement, aussi bien en mathématiques qu'en physique, de notions rejetant cette régularité et faisant largement usage de fonctions à la rigueur continues, mais dérivables nulle part (chaos, fractales...).

<u>d) Etude de suite récurrente $u_{n+1} = f(u_n)$</u> :
On suppose que f est dérivable, et qu'il existe k élément de $[0, 1[$ tel que $|f'| \leq k$. On suppose qu'il existe un nombre α tel que $f(\alpha) = \alpha$ (point fixe de f). Alors la suite (u_n) converge vers α. En effet :

$|u_{n+1} - \alpha| = |f(u_n) - f(\alpha)| \leq k\,|u_n - \alpha|$ en appliquant l'inégalité des accroissements finis.

On en déduit alors par récurrence que :

$$\forall\, n,\ |u_n - \alpha| \leq k^n\,|u_0 - \alpha|$$

et comme $\lim\limits_{n \to \infty} k^n = 0$, on a bien $\lim\limits_{n \to \infty} u_n = \alpha$.

Cette propriété peut servir à obtenir des valeurs approchées de α en utilisant une suite récurrente.

EXEMPLE :
Soit $f(x) = \dfrac{1}{1 + x}$ dont le point fixe positif vérifie $\alpha = \dfrac{1}{1 + \alpha} \Rightarrow \alpha^2 + \alpha - 1 = 0 \Rightarrow \alpha = \dfrac{\sqrt{5} - 1}{2}$.
On a $f'(x) = -\dfrac{1}{(1 + x)^2}$.
Soit $0 < d < \alpha$. Pour $\alpha - d \leq x \leq \alpha + d$, on a :

$$|f'(x)| = \frac{1}{(1 + x)^2} \leq \frac{1}{(1 + \alpha - d)^2}$$

et $$|f(x) - \alpha| \leq \frac{1}{(1 + \alpha - d)^2}\,|x - \alpha| \leq \frac{d}{(1 + \alpha - d)^2} \leq d$$

de sorte que l'intervalle $[\alpha - d, \alpha + d]$ a son image par f incluse dans lui-même. Il est donc stable par f et non pouvons itérer la fonction f. L'inégalité $|f'(x)| \leq \dfrac{1}{(1 + \alpha - d)^2}$ montre que la fonction f est lipschitzienne de rapport $\dfrac{1}{(1 + \alpha - d)^2}$ strictement inférieur à 1. La convergence de la suite de terme général $u_{n+1} = f(u_n)$ vers α est donc assuré dès lors que u_0 appartient à $[\alpha - d, \alpha + d]$. Comme $\sqrt{5}$ est compris entre 2 et 3, α est compris strictement entre $\dfrac{1}{2}$ et 1. Prenons par exemple $u_0 = 1$ et prenons d tel que $\alpha + d = 1$, soit $d = 1 - \alpha = \dfrac{3 - \sqrt{5}}{2} \Rightarrow \alpha - d = \sqrt{5} - 2 > 0$. La suite (u_n) va converger vers α et permet d'obtenir des valeurs approchées de α. Voici les premiers termes de la suite sous forme rationnelle, puis sous forme décimale :

$$1, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{5}{8}, \frac{8}{13}, \frac{13}{21}, \frac{21}{34}, \frac{34}{55}, \frac{55}{89}, \frac{89}{144}$$

1, 0.5, 0.666..., 0.6, 0.625, 0.615..., 0.619, 0.617..., 0.61818..., 0.617977..., 0.61805...

La calculatrice donne $\alpha = 0.6180339890$, mais il a bien fallu que quelqu'un programme l'algorithme de calcul de la racine carrée. Comment a-t-il fait ?

II : Fonctions convexes

La fin du chapitre est réservée aux MPSI, PCSI, et PTSI suivant l'option mathématiques

De nombreuses inégalités se montrent en analyse en utilisant certaines propriétés de convexité des fonctions. Par exemple :

$$\sqrt[n]{y_1 y_2 \dots y_n} \leq \frac{y_1 + \dots + y_n}{n}$$

C'est cette notion que nous allons développer ici.

1– Définition

On dit qu'une partie A du plan (ou plus généralement d'un espace affine) est convexe si :

$$\forall x \in \mathrm{A}, \forall y \in \mathrm{A}, [x, y] \subset \mathrm{A}$$

$[x, y]$ est l'ensemble des points de la forme $\lambda x + (1 - \lambda)y$, $\lambda \in [0, 1]$, barycentres de x et y à coefficients positifs ou nuls.

PROPOSITION

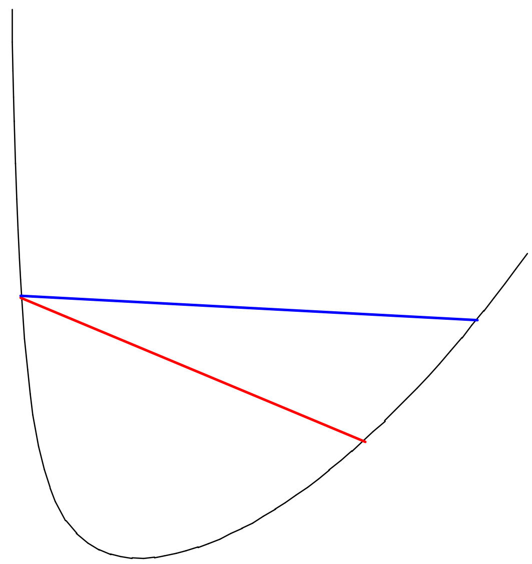
Soit f une fonction définie sur un intervalle I. Il y a équivalence entre :

i) $A = \{(x, y) \mid y \geq f(x)\}$ est une partie convexe du plan.

ii) $\forall x \in I, \forall y \in I, \forall \lambda \in [0, 1], f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y)$

iii) Pour tout u de I, la fonction qui à x associe $\dfrac{f(x) - f(u)}{x - u}$ est croissante.

Une fonction f vérifiant ces propriétés est dite convexe.



Lorsque λ décrit [0,1], le point $(\lambda x + (1-\lambda)y\ ,\ f(\lambda x + (1-\lambda)y))$ décrit l'arc de la courbe représentative de f, entre les points d'abscisse x et y. Le point $(\lambda x + (1-\lambda)y\ ,\ \lambda f(x) + (1-\lambda)f(y))$, lui, est barycentre de $(x, f(x))$ et $(y, f(y))$ avec les coefficients λ et $1-\lambda$. Il décrit donc le segment de droite joignant $(x, f(x))$ à $(y, f(y))$. La propriété ii) s'interprète en disant que l'arc limité par les points d'abscisse x et y se trouve sous sa corde correspondante.

La propriété iii) exprime que les pentes des cordes augmentent lorsque l'abscisse de l'une des extrémités de la corde croît. La pente bleue est ainsi plus grande que la pente rouge.

Démonstration :
i) $\Rightarrow$ ii)
$(x, f(x))$ et $(y, f(y))$ sont éléments de A. Donc le segment joignant ces deux points est élément de A puisque A est supposé convexe. Donc, $(\lambda x + (1-\lambda)y\ ,\ \lambda f(x) + (1-\lambda)f(y))$ qui appartient à ce segment est dans A, ce qui signifie que $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$.

ii) $\Rightarrow$ iii)
Afin de considérer toutes les positions possibles par rapport à u, nous montrerons que :

$$\text{si } x < y < u < z < t, \text{ alors } \frac{f(x) - f(u)}{x - u} \leq \frac{f(y) - f(u)}{y - u} \leq \frac{f(z) - f(u)}{z - u} \leq \frac{f(t) - f(u)}{t - u}.$$

Ces trois inégalités équivalent respectivement à :

$$\text{avec } \lambda = \frac{u-y}{u-x}, \qquad y = \lambda x + (1-\lambda)u, \qquad f(y) \leq \lambda f(x) + (1-\lambda)f(u)$$

$$\text{avec } \lambda = \frac{z-u}{z-y}, \qquad u = \lambda y + (1-\lambda)z, \qquad f(u) \leq \lambda f(y) + (1-\lambda)f(z)$$

$$\text{avec } \lambda = \frac{t-z}{t-u}, \qquad z = \lambda u + (1-\lambda)t, \qquad f(z) \leq \lambda f(u) + (1-\lambda)f(t)$$

et ces trois inégalités sont bien vérifiées, d'après ii).

iii) $\Rightarrow$ i)
Soient deux points (x, y) et (t, z) de A. $y \geq f(x)$, $z \geq f(t)$. Soit λ élément de [0, 1]. Nous voulons montrer que $\lambda y + (1-\lambda)z \geq f(\lambda x + (1-\lambda)t)$. Il suffit pour cela de montrer :

$$\lambda f(x) + (1-\lambda)f(t) \geq f(\lambda x + (1-\lambda)t)$$

Cette inégalité est évidemment vraie pour $\lambda = 0$ ou $\lambda = 1$. Dans le cas contraire, posons $u = \lambda x + (1-\lambda)t$. L'inégalité est alors équivalente à :

$$\frac{f(x) - f(u)}{x - u} \leq \frac{f(t) - f(u)}{t - u}$$

qui est vérifiée d'après la croissance du taux d'accroissement relatif à u.

EXEMPLE : e^x, x^2, ... sont des fonctions convexes.

2– Convexité et dérivabilité

PROPOSITION

Soit f une fonction dérivable sur un intervalle I. Il y a équivalence entre :

i) f est convexe

ii) f ' est croissante

iii) La courbe représentative de f est au–dessus de chaque tangente

Démonstration :

i) $\Rightarrow$ ii)

Supposons f convexe. Soit $x < y$. On a, pour $x < u < t < y$:

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

La première inégalité provient de la croissance du taux d'accroissement relatif à u, la deuxième de la croissance du taux d'accroissement relatif à t. Si l'on fait tendre u vers x et t vers y, on obtient :

$$f'(x) \leq \frac{f(y) - f(x)}{y - x} \leq f'(y)$$

Donc f' est croissante.

ii) $\Rightarrow$ iii)

Soit f' croissante. Dire que la courbe représentative de f est au–dessus de chaque tangente, par exemple, celle qui passe par le point d'abscisse u, signifie que :

$$\forall x, f(x) \geq (x-u)f'(u) + f(u)$$

ou encore :

$$x < u \Rightarrow \frac{f(x) - f(u)}{x - u} \leq f'(u)$$

$$x > u \Rightarrow \frac{f(x) - f(u)}{x - u} \geq f'(\mathrm{u})$$

Montrons la première inégalité, l'autre étant analogue. D'après le théorème des accroissement fini, le taux d'accroissement entre x et u est égal à $f'(c)$ avec $x < c < u$. f' étant croissante, l'inégalité est vérifiée.

iii) $\Rightarrow$ i)

La deuxième propriété de convexité de f à démontrer peut s'écrire :

$$\frac{f(x) - f(u)}{x - u} \leq \frac{f(y) - f(u)}{y - u}$$

où l'on a posé $u = \lambda x + (1-\lambda)y$, avec $x < u < y$

Or, comme ci-dessus, iii) signifie que :

$$x < u \Rightarrow \frac{f(x) - f(u)}{x - u} \leq f'(u)$$

$$y > u \Rightarrow f\,'(u) \leq \frac{f(y) - f(u)}{y - u}$$

de sorte que $\dfrac{f(x) - f(u)}{x - u} \leq f\,'(u) \leq \dfrac{f(y) - f(u)}{y - u}$

Si f est deux fois dérivable, la condition donnée est équivalente à $f\,'' \geq 0$.

Voici une dernière propiété des fonctions convexes.

PROPRIETE

Soit f convexe sur I, soient $(x_i)_{i\in\{1..n\}}$ des points de I soient $(a_i)_{i\in\{1..n\}}$ des réels positifs de somme 1. Alors $f(\sum_{i=1}^{n} a_i x_i) \leq \sum_{i=1}^{n} a_i f(x_i)$

Démonstration :

Par récurrence sur n.

Cette propriété est vraie pour $n = 2$. Il s'agit en effet de la deuxième propriété de convexité de f. Supposons–la vraie pour $n-1$. La propriété au rang n exprime que l'image du barycentre g des points d'abscisse x_i, affectés des coefficients a_i est inférieure au barycentre des images $f(x_i)$, affectés des mêmes coefficients. Or, si g' est le barycentre de x_1, ..., x_{n-1}, g peut être considéré comme barycentre de g, affecté des coefficients $a_1 + \ldots + a_{n-1} = s_{n-1}$ et de x_n affecté du coefficient $a_n = 1 - s_{n-1}$. Ces propriétés se traduisent par :

$$g' = \frac{a_1 x_1}{s_{n-1}} + \ldots + \frac{a_{n-1} x_{n-1}}{s_{n-1}}$$

$$g = s_{n-1} g' + a_n x_n$$

Selon l'hypothèse de récurrence, on a :

$$f(g') \leq \frac{a_1 f(x_1)}{s_{n-1}} + \ldots + \frac{a_{n-1} f(x_{n-1})}{s_{n-1}}$$

or $\quad f(g) \leq s_{n-1} f(g') + a_n f(x_n)$

donc $\quad f(g) \leq \sum_{i=1}^{n} a_i f(x_i)$

APPLICATION :

Considérons $f(x) = e^x$ et $a_i = \dfrac{1}{n}$. Alors on a $\exp(\sum_{i=1}^{n} \frac{x_i}{n}) \leq \frac{1}{n} \sum_{i=1}^{n} \exp(x_i)$.

ce qui donne, en posant $y_i = \exp(x_i)$:

$$\sqrt[n]{y_1 y_2 \ldots y_n} \leq \frac{y_1 + \ldots + y_n}{n}$$

$\sqrt[n]{y_1 y_2 \ldots y_n}$ s'appelle moyenne géométrique des nombres y_1, ..., y_n alors que $\dfrac{y_1 + \ldots + y_n}{n}$ en est la moyenne arithmétique. La moyenne géométrique est donc inférieure à la moyenne arithmétique.

Une démonstration directe de cette dernière égalité peut se faire de la façon suivante :

Notons $m = \dfrac{y_1 + \ldots + y_n}{n}$, et $z_i = y_i - m$. On veut montrer que :

$$y_1 y_2 \ldots y_n \leq m^n$$

$\Leftrightarrow \quad \ln(y_1) + \ldots + \ln(y_n) \leq n \ln m$

$\Leftrightarrow \quad \ln(\frac{y_i}{m}) + \ldots + \ln(\frac{y_n}{m}) \leq 0$

$\Leftrightarrow \quad \ln(1 + \frac{z_i}{m}) + \ldots + \ln(1 + \frac{z_n}{m}) \leq 0$

inégalité qui résulte des deux relations suivantes :

$\forall\ x$, $\ln(1+x) \leq x$ (qu'on peut montrer par une étude de fonction, mais qui est elle-même une inégalité de convexité appliquée sur la fonction – ln [position de la fonction ln par rapport à sa tangente])

$z_1 + \ldots + z_n = 0$