Analyse convexe

Université d'Eleuthéria-Polites

République de Poldévie

Cours de Licence 2 — 2023/2024

Bruno Deschamps

Version 1.0

Les physiciens étudient les lois auxquelles l'Univers à du obéir quand Dieu l'a créé.
Les mathématiciens étudient les lois auxquelles Dieu a du obéir quand il a créé l'Univers...

Table des matières

1 Convexité

1.1 Généralités.

On considère un $\mathbb{R}$-espace vectoriel E.

Définition 1.— *Si $x, y \in E$, on appelle "segment" d'extrémités x et y, l'ensemble*

$$[x,y] = \{t.x + (1-t).y/ \ t \in [0,1]\}$$

On dit d'une partie $A \subset E$ qu'elle est convexe si pour tout $x, y \in A$, $[x,y] \subset A$.

Théorème 2.— (Généralisation) *Soit $A \subset E$. Les propriétés suivantes sont équivalentes :*

i) A est convexe,

ii) pour tout entier $p \geq 2$, toute suite finie $x_1, \cdots, x_p \in A$ et toute suite finie $\alpha_1, \cdots, \alpha_p \in \mathbb{R}^+$ telle que $\sum_{i=1}^{p} \alpha_i = 1$, on a $\sum_{i=1}^{p} \alpha_i.x_i \in A$.

Preuve : Pour $p = 2$, l'énoncé ii) équivaut à i) : il suffit de considérer $t = \alpha_1$, on a alors $\alpha_2 = 1 - t$. Ainsi, on a $ii) \Rightarrow i)$

$i) \Rightarrow ii)$ On va montrer ii) par récurrence sur $p \geq 2$. Comme on vient de le remarquer, le cas $p = 2$ est exactement i). Supposons la propriété vraie pour $p \geq 2$ et considérons $x_1, \cdots, x_{p+1} \in A$ et $\alpha_1, \cdots, \alpha_{p+1} \in \mathbb{R}^+$ telle que $\sum_{i=1}^{p+1} \alpha_i = 1$.

- Si $\alpha_{p+1} = 1$ alors $\alpha_1 = \cdots = \alpha_p = 0$ et donc $\sum_{i=1}^{p} \alpha_i.x_i = x_{p+1} \in A$.
- Si $\alpha_{p+1} \neq 1$ alors on peut diviser par $1 - \alpha_{p+1}$ et l'on a alors

$$\frac{\alpha_1}{1-\alpha_{p+1}} + \cdots + \frac{\alpha_p}{1-\alpha_{p+1}} = 1$$

En utilisant l'hypothèse de récurrence, on a donc

$$y = \frac{\alpha_1}{1-\alpha_{p+1}}.x_1 + \cdots + \frac{\alpha_p}{1-\alpha_{p+1}}.x_p \in A$$

mais comme $x_{p+1}, y \in A$ et $\alpha_{p+1} + (1 - \alpha_{p+1}) = 1$, en appliquant le i), on trouve finalement

$$(1-\alpha_{p+1}).y + \alpha_{p+1}.x_{p+1} = \alpha_1.x_1 + \cdots + \alpha_p.x_p + \alpha_{p+1}.x_{p+1} \in A$$

□

Exemples : 1/ Cas $E = \mathbb{R}$. Si l'on considère deux réel $x < y$ et le segment (au sens réel du terme) $[x,y] = \{u \in \mathbb{R}/ \ x \leq u \leq y\}$ alors on constate que

$$[x,y] = \{tx + (1-t)y/ \ t \in [0,1]\}$$

En effet, si $t \in [0,1]$ et $u = tx + (1-t)y$ alors $u - x = (1-t)(y-x) \geq 0$ et $y - u = t(y-x) \geq 0$ et donc $x \leq u \leq y$. Réciproquement, si $x \leq u \leq y$ on pose $t = \dfrac{y-u}{y-x}$ et l'on voit immédiatement que $u = tx + (1-t)y$ et que $t \in [0,1]$. On voit donc que, dans $\mathbb{R}$, la notion usuelle de segment correspond exactement à la notion de segment au sens de la définition 1 (ce qui justifie la terminologie). On fera attention toutefois au fait que la notion usuelle de segment $[x,y]$ exige que $x \leq y$, alors qu'au sens de la définition 1 on a $[x,y] = [y,x]$.

Rappel : Intervalles.

Définition.— *Un intervalle est une partie I de $\mathbb{R}$ qui vérifie la propriété suivante :*

$$\forall x,y \in I,\ \forall t \in \mathbb{R},\ x < t < y \Longrightarrow t \in I$$

En considérant, dans la droite numérique achevée $\overline{\mathbb{R}}$, les bornes inférieure et supérieure de I, on montre que I a l'une des 9 formes suivantes :

- $]-\infty,+\infty[= \mathbb{R}$,
- $]-\infty,a[= \{x \in \mathbb{R}/\ x < a\}$ pour un certain $a \in \mathbb{R}$,
- $]-\infty,a] = \{x \in \mathbb{R}/\ x \leq a\}$,
- $]a,+\infty[= \{x \in \mathbb{R}/\ x > a\}$,
- $[a,+\infty[= \{x \in \mathbb{R}/\ x \geq a\}$,
- $]a,b[= \{x \in \mathbb{R}/\ a < x < b\}$, pour certains $a,b \in \mathbb{R}$ vérifiant $a < b$,
- $]a,b] = \{x \in \mathbb{R}/\ a < x \leq b\}$, pour certains $a,b \in \mathbb{R}$ vérifiant $a < b$,
- $[a,b[= \{x \in \mathbb{R}/\ a \leq x < b\}$, pour certains $a,b \in \mathbb{R}$ vérifiant $a < b$,
- $[a,b] = \{x \in \mathbb{R}/\ a \leq x \leq b\}$, pour certains $a,b \in \mathbb{R}$ vérifiant $a < b$,

et réciproquement.

Théorème 3.— *Les parties non vides et convexes de $\mathbb{R}$ sont exactement les intervalles.*

Preuve : Si I est un intervalle, alors pour tous $x,y \in I$ avec $x < y$ et tout $x \leq u \leq y$, on a par définition $u \in I$. On vient de voir que $[x,y] = \{u/\ x \leq u \leq y\}$, ainsi on a bien $[x,y] \subset I$. Réciproquement, si I est convexe et si $x,y \in I$ avec $x < y$ alors on a $[x,y] \subset I$. Mais comme $[x,y] = \{u/\ x \leq u \leq y\}$, on voit donc que $x \leq u \leq y \Longrightarrow u \in I$: I est un intervalle.

$\square$

2/ Tout segment de E est convexe. Si $u,v \in [x,y]$, alors il existe $\alpha,\beta \in [0,1]$ tels que $u = \alpha.x + (1-\alpha).y$ et $v = \beta.x + (1-\beta).y$. Si $t \in [0,1]$, on alors

$$t.u + (1-t).v = (t\alpha + (1-t)\beta).x + (t(1-\alpha) + (1-t)(1-\beta)).y$$

or, $(t\alpha + (1-t)\beta) \leq 0$ et $(t\alpha + (1-t)\beta) + (t(1-\alpha) + (1-t)(1-\beta)) = t + (1-t) = 1$, et donc $t.u + (1-t).v \in [x,y]$. Ceci prouve finalement que $[u,v] \subset [x,y]$ et donc que $[x,y]$ est bien convexe.

3/ Tout sous-espace vectoriel de E est convexe. En effet, si F est un s.e.v. de E alors, pour tout $x,y \in E$ et tout $\lambda,\mu \in \mathbb{R}$, on a $\lambda.x + \mu.y \in F$. En particulier, si l'on choisit $\lambda = t$ et $\mu = 1-t$ pour $t \in [0,1]$ on a $t.x + (1-t).y \in F$.

4/ Si E est en plus normé, alors toute boule de E est une partie convexe de E. Considérons une boule ouverte $B = B(u,\varepsilon) = \{x \in E/\ \|x-u\| < \varepsilon\}$. Si $x,y \in B$ et $t \in [0,1]$, on a $(t.x + (1-t).y) - u = (t.x + (1-t).y) - (t + (1-t)).u = t.(x-u) + (1-t).(y-u)$ et donc

$$\|(t.x + (1-t).y) - u\| = \|t.(x-u) + (1-t).(y-u)\| \leq t\|(x-u)\| + (1-t)\|(y-u)\| < t\varepsilon + (1-t)\varepsilon = \varepsilon$$

et donc $t.x + (1-t).y \in B$. La boule B est convexe. Le même raisonnement marche pour les boules fermées.

1.2 Enveloppe convexe.

Proposition 4.— *Si $(A_i)_{i\in I}$ désigne une famille finie et non vide de parties convexes de E alors l'intersection*

$$A = \bigcap_{i\in I} A_i$$

est une partie convexe de E.

Preuve : Si $x, y \in A$ alors, pour tout $i \in I$, $x, y \in A_i$ et donc $[x, y] \subset A_i$. Ceci étant valable pour tout $i \in I$, on a finalement $[x, y] \subset \bigcap_{i\in I} A_i = A$.

□

Théorème-Définition 5.— *Si $A \subset E$, alors parmi toutes les parties convexe de E qui contiennent A il en existe une plus petite. On l'appelle "l'enveloppe convexe" de la partie A et on la note $\mathfrak{Env}(A)$.*

Preuve : On considère la famille $(A_i)_{i\in I}$ des parties convexe de E qui contiennent A. Cette famille est non vide car la partie E y appartient. Par la proposition 4, la partie $\mathfrak{Env}(A) = \bigcap_{i\in I} A_i$ est une partie convexe et elle contient visiblement la partie A. C'est nécessairement la plus petite car elle appartient *de facto* à la famille $(A_i)_{i\in I}$.

□

Théorème 6.— *Si A est une partie de E alors $\mathfrak{Env}(A)$ est égale à l'ensemble des barycentres de systèmes pondérés d'éléments de A et de somme des poids 1 :*

$$\mathfrak{Env}(A) = \left\{ \alpha_1.x_1 + \cdots + \alpha_n.x_n /\ n \geq 2,\ x_1, \cdots, x_n \in A,\ \alpha_1, \cdots, \alpha_n \in \mathbb{R}^+ \text{ avec } \sum_{i=1}^{n} \alpha_i = 1 \right\}$$

Preuve : Notons Ω l'ensemble décrit dans l'énoncé. Le théorème 2 montre que Ω est incluse dans toute partie convexe de E qui contient A, donc en particulier dans $\mathfrak{Env}(A)$. Puisque $\mathfrak{Env}(A)$ est la plus petite d'entre elle, pour montrer que $\mathfrak{Env}(A) = \Omega$ il suffit de montrer que Ω est convexe. Considérons deux élément $u, v \in \Omega$. Quitte à rajouter des x_i et des poids α_i et β_i nuls, on peut supposer qu'il existe $n \geq 2$, $x_1, \cdots, x_n \in A$, $\alpha_1, \beta_1, \cdots, \alpha_n, \beta_n \in \mathbb{R}^+$ avec $\displaystyle\sum_{i=1}^{n} \alpha_i = \sum_{i=1}^{n} \beta_i = 1$ et tels que

$$\begin{aligned} u &= \alpha_1.x_1 + \cdots + \alpha_n.x_n \\ v &= \beta_1.x_1 + \cdots + \beta_n.x_n \end{aligned}$$

Considérons un réel $t \in [0, 1]$ et posons $w = t.u + (1-t).v \in [u, v]$:

$$w = (t\alpha_1 + (1-t)\beta_1).x_1 + \cdots + (t\alpha_n + (1-t)\beta_n).x_n = \gamma_1.x_1 + \cdots + \gamma_n.x_n$$

où $\gamma_i = t\alpha_i + (1-t)\beta_i$. On a

$$\sum_{i=1}^{n} \gamma_i = t \sum_{i=1}^{n} \alpha_i + (1-t) \sum_{i=1}^{n} \beta_i = t + (1-t) = 1$$

et il s'ensuit que $w \in \Omega$. On a donc $[u,v] \subset \Omega$ et, par suite, Ω est convexe.

□

2 Fonctions convexes

2.1 Définition et caractérisations.

Dans ce qui suit, I désigne un intervalle de $\mathbb{R}$ d'intérieur non vide et $f : I \longrightarrow \mathbb{R}$ une application.

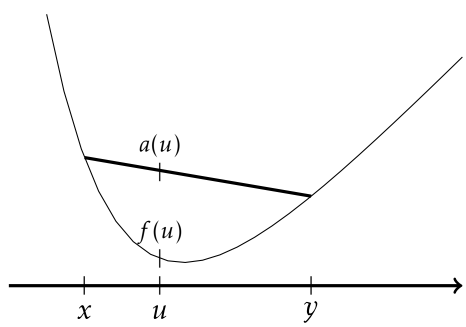
Définition 7.— *On appelle graphe (resp. épigraphe) de f l'ensemble des couples $(x,y) \in I \times \mathbb{R}$ tels que $y = f(x)$ (resp. $y \geq f(x)$). On dit que la fonction f est convexe si son épigraphe est une partie convexe de $\mathbb{R}^2$.*

Théorème 8.— (Caractérisation de la convexité d'une fonction) *Les propriétés suivantes sont équivalentes :*

i) f est convexe,

ii) pour tout $x,y \in I$ et tout $t \in [0,1]$, $f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$,

iii) pour tout $x,y \in I$ $(x<y)$ et tout $u \in [x,y]$, on a $f(u) \leq a(u)$, où a est l'unique fonction affine qui coïncide avec f en x et en y.



(la corde est "au dessus" de la courbe.)

Preuve : $i) \Rightarrow ii)$ Les points $A=(x,f(x))$ et $B=(y,f(y))$ sont dans le graphe de f, donc dans son épigraphe. Par convexité, le segment $[A,B]$ est donc inclus dans l'épigraphe de f, c'est-à-dire que pour tout $t \in [0,1]$, le point $t.A+(1-t).B = (tx+(1-t)y, tf(x)+(1-t)f(y))$ appartient à l'épigraphe : $f(tx+(1-t)y) \leq tf(x)+(1-t)f(y)$.

$ii) \Leftrightarrow iii)$ La fonction affine a a pour équation cartésienne $a(u) = \frac{f(y)-f(x)}{y-x}(x-u)+f(x)$. Puisque $[x,y] = \{tx+(1-t)y/\ t \in [0,1]\}$, on a donc

$$\forall u \in [x,y],\ f(u) \leq a(u) \Longleftrightarrow \forall t \in [0,1],\ f(tx+(1-t)y) \leq a(tx+(1-t)y) = tf(x)+(1-t)f(y)$$

$ii) \Rightarrow i)$ Soient $A=(x,\alpha)$ et $B=(y,\beta)$ deux points de l'épigraphe de f. Pour tout $t \in [0,1]$, on a

$$f(tx+(1-t)y) \leq tf(x)+(1-t)f(y) \leq t\alpha+(1-t)\beta$$

et donc $t.A+(1-t).B$ est aussi dans l'épigraphe. L'épigraphe de f est donc convexe.

□

Corollaire 9.— *Les propriétés suivantes sont équivalentes :*

i) f est convexe,

ii) pour tout entier $p \geq 2$, toute suite finie $x_1, \cdots, x_p \in I$ et toute suite finie $\alpha_1, \cdots, \alpha_p \in \mathbb{R}^+$ telle que $\sum_{i=1}^{p} \alpha_i = 1$, on a $f\left(\sum_{i=1}^{p} \alpha_i x_i\right) \leq \sum_{i=1}^{p} \alpha_i f(x_i)$.

Preuve : La preuve s'effectue exactement comme celle du théorème 2. On constate i) équivaut au ii) avec $p = 2$ et sous i), on montre ii) par récurrence sur $p \geq 2$: la propriété est vraie pour $p = 2$ et, si l'on suppose que la propriété est vraie pour $p \geq 2$, alors au rang $p+1$ on considère $x_1, \cdots, x_{p+1} \in I$ et $\alpha_1, \cdots, \alpha_{p+1} \in \mathbb{R}^+$ telle que $\sum_{i=1}^{p+1} \alpha_i = 1$.

• Si $\alpha_{p+1} = 1$ alors $\alpha_1 = \cdots = \alpha_p = 0$ et donc

$$f\left(\sum_{i=1}^{p} \alpha_i . x_i\right) = f(x_{p+1}) = \sum_{i=1}^{p} \alpha_i f(x_i)$$

• Si $\alpha_{p+1} \neq 1$ alors on peut diviser par $1 - \alpha_{p+1}$ et l'on a alors

$$\frac{\alpha_1}{1-\alpha_{p+1}} + \cdots + \frac{\alpha_p}{1-\alpha_{p+1}} = 1$$

Par convexité de I, on en déduit que

$$y = \frac{\alpha_1}{1-\alpha_{p+1}} . x_1 + \cdots + \frac{\alpha_p}{1-\alpha_{p+1}} . x_p \in I$$

et, en utilisant l'hypothèse de récurrence, on a donc

$$f\left(\sum_{i=1}^{p} \frac{\alpha_i}{1-\alpha_{p+1}} x_i\right) \leq \sum_{i=1}^{p} \frac{\alpha_i}{1-\alpha_{p+1}} f(x_i)$$

Puisque $x_{p+1}, y \in I$ et $\alpha_{p+1} + (1 - \alpha_{p+1}) = 1$, en appliquant le i), on trouve finalement

$$\begin{aligned} f\left(\sum_{i=1}^{p+1} \alpha_i x_i\right) &= f\Big((1-\alpha_{p+1})y + \alpha_{p+1} x_{p+1}\Big) \leq (1-\alpha_{p+1}) f(y) + \alpha_{p+1} f(x_{p+1}) \\ &\leq (1-\alpha_{p+1}) \sum_{i=1}^{p} \frac{\alpha_i}{1-\alpha_{p+1}} f(x_i) + \alpha_{p+1} f(x_{p+1}) = \sum_{i=1}^{p+1} \alpha_i f(x_i) \end{aligned}$$

□

Corollaire 10.— *1/ Si $f, g : I \longrightarrow \mathbb{R}$ sont des applications convexes alors, pour tous $\lambda, \mu \in \mathbb{R}^+$, l'application $\lambda f + \mu g$ est convexe.*

2/ Si $(f_n)_n$ désigne une suite de fonctions convexes sur I qui converge simplement vers une fonction f, alors f est une application convexe. En particulier, la somme d'une série de fonctions convexes convergente est convexe.

3/ Soit $f : I \longrightarrow J$ une bijection continue (J est alors un intervalle et f est nécessairement strictement monotone). Si f est convexe alors sa réciproque $f^{-1} : J \longrightarrow I$ est une fonction convexe (si f est décroissante) ou concave (si f est croissante).

Preuve : 1/ Pour tous $x, y \in I$ et tout $t \in [0,1]$, on a

$$\begin{array}{rcl}(\lambda f + \mu g)(tx+(1-t)y) & = & \lambda f(tx+(1-t)y) + \mu g(tx+(1-t)y) \\ & \leq & \lambda t f(x) + \lambda(1-t)f(y) + \mu t g(x) + \mu(1-t)g(y) \\ & = & t(\lambda f + \mu g)(x) + (1-t)(\lambda f + \mu g)(y)\end{array}$$

2/ Pour tous $x, y \in I$, tout $t \in [0,1]$ et tout $n \geq 0$, on a

$$f_n(tx+(1-t)y) \leq t f_n(x) + (1-t) f_n(y)$$

Les réels x, y, t étant fixés, par passage à la limite en n on trouve alors

$$f(tx+(1-t)y) \leq t f(x) + (1-t) f(y)$$

Par application du 1/, on voit que, pour tout $n \geq 0$, $S_n(x) = \displaystyle\sum_{k=0}^{n} f_k(x)$ est une fonction convexe. La fonction $\displaystyle\sum_{k\geq 0} f_k(x) = \lim_n S_n(x)$ est donc convexe d'après ce qui précède.

3/ Si $u, v \in J$ et $t \in [0,1]$, puisque $f^{-1}(u), f^{-1}(v) \in I$ et que f est convexe, on a

$$f(tf^{-1}(u) + (1-t)f^{-1}(v)) \leq t f(f^{-1}(u)) + (1-t) f(f^{-1}(v)) = tu + (1-t)v$$

Les fonctions f et f^{-1} ont même sens de monotonie, donc si f est décroissante, on a

$$tf^{-1}(u) + (1-t)f^{-1}(v) = f^{-1}(f(tf^{-1}(u) + (1-t)f^{-1}(v))) \geq f^{-1}(tu + (1-t)v)$$

et donc f^{-1} est convexe, et si f est croissante, on a

$$tf^{-1}(u) + (1-t)f^{-1}(v) = f^{-1}(f(tf^{-1}(u) + (1-t)f^{-1}(v))) \leq f^{-1}(tu + (1-t)v)$$

et donc f^{-1} est concave.

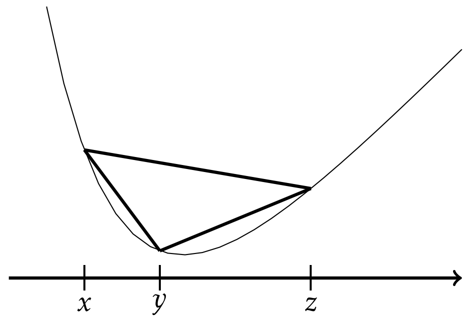
□

Théorème 11.— (Inégalité des pentes) *Les propriétés suivantes sont équivalentes :*

i) f est convexe,

ii) pour tout $x, y, z \in I$ tels que $x < y < z$ on a

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



iii) pour tout $u \in I$, la fonction $p_u : I - \{u\} \longrightarrow \mathbb{R}$ définie par

$$p_u(x) = \frac{f(x) - f(u)}{x - u}$$

est croissante.

Preuve : $i) \Rightarrow ii)$ Considérons la fonction affine a qui coïncide avec f en x et en z. Puisque f est convexe et que y est compris entre x et z, on a donc $f(y) \leq a(y)$. Ainsi on a

$$\frac{f(y) - f(x)}{y - x} \leq \frac{a(y) - f(x)}{y - x} = \frac{a(y) - a(x)}{y - x} = \frac{a(z) - a(x)}{z - x} = \frac{f(z) - f(x)}{z - x}$$

De même, on a

$$\frac{f(z) - f(x)}{z - x} = \frac{a(z) - a(x)}{z - x} = \frac{a(z) - a(y)}{z - y} \leq \frac{f(z) - f(y)}{z - y}$$

$ii) \Rightarrow iii)$ Se déduit du $ii)$ en remplaçant le triplet (x, y, z) par (t_1, t_2, u), (t_1, u, t_2) et (u, t_1, t_2). Dans tous les cas on obtient $t_1 < t_2 \Rightarrow p_u(t_1) \leq p_u(t_2)$.
$iii) \Rightarrow i)$ Considérons x, y, z tels que $x < y < z$. En choisissant $u = x$, on obtient

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

Considérons la fonction affine a qui coïncide avec f en x et z. On a donc

$$\frac{f(z) - f(x)}{z - x} = \frac{a(z) - a(x)}{z - x} = \frac{a(y) - a(x)}{y - x}$$

et donc $f(y) - f(x) \leq a(y) - a(x)$. Comme $a(x) = f(x)$, on en déduit que pour tout $y \in]x, z[$, $f(y) \leq a(y)$. La fonction f est donc convexe.

□

Les trois inégalités incriminées dans le ii) de ce théorème sont en fait équivalentes :

Proposition 12.— *Pour tout $x, y, z \in I$ tels que $x < y < z$ on a*

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

Preuve : Puisque $y \in]x,z[$, il existe $t \in]0,1[$ tel que $y = tx + (1-t)z$ et l'on a alors

$$\begin{cases} (y-x) = (1-t)(z-x) \\ (z-y) = t(z-x) \end{cases}$$

Il s'ensuit que

$$\begin{aligned}
\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(y)}{z-y} \quad & \iff \quad \frac{f(y)-f(x)}{(1-t)(z-x)} \leq \frac{f(z)-f(y)}{t(z-x)} \iff t(f(y)-f(x)) \leq (1-t)(f(z)-f(y)) \\
& \iff \quad t(f(z)-f(x)) \leq (f(z)-f(y)) \iff \frac{t(f(z)-f(x))}{t(z-x)} \leq \frac{f(z)-f(y)}{t(z-x)} \\
& \iff \quad \frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y} \iff \\
\frac{f(z)-f(x)}{z-x} \leq \frac{f(z)-f(y)}{z-y} \quad & \iff \quad t(f(z)-f(x)) \leq (f(z)-f(y)) \iff f(y)-f(x) \leq (1-t)(f(z)-f(x)) \\
& \iff \quad \frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} \iff \frac{f(y)-f(x)}{(1-t)(z-x)} \leq \frac{f(z)-f(x)}{z-x} \iff \\
\frac{f(y)-f(x)}{y-x} \leq \frac{f(z)-f(x)}{z-x} &
\end{aligned}$$

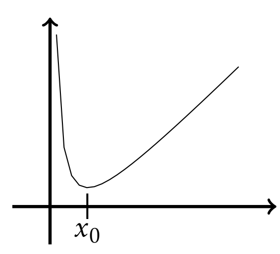
□

La convexité de f est donc équivalente à chacune des trois inégalités du ii) du théorème 11.

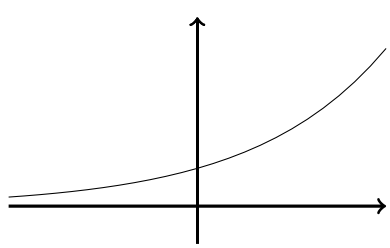
2.2 Convexité et monotonie.

Théorème 13.— *On considère une fonction $f : I \longrightarrow \mathbb{R}$ définie sur un intervalle ouvert I. Si f est convexe alors*

• *soit f possède un minimum local en un point x_0 et, dans ces conditions, x_0 est un minimum global et f est décroissante à gauche de x_0 et croissante à droite de x_0,*



• *soit f ne possède pas de minimum local sur I et, dans ces conditions, f est monotone.*



Preuve : • Si x_0 désigne un minimum local de f alors il existe donc $\varepsilon > 0$ tel que, pour tout $x \in]x_0 - \varepsilon, x_0 + \varepsilon[$, $f(x) \geq f(x_0)$. Supposons que x_0 ne soit pas un minimum global, il existe donc $x_1 \in I$ tel que $f(x_1) < f(x_0)$. Sans perte de généralité, supposons que $x_0 < x_1$. La pente $\dfrac{f(x_1) - f(x_0)}{x_1 - x_0}$ de l'application affine a qui passe par $f(x_0)$ en x_0 et $f(x_1)$ en x_1 est alors strictement négative, si bien que, pour tout $x > x_0$, $a(x) < a(x_0)$. Puisque f est convexe, on a alors $f(x) \leq a(x) < a(x_0) = f(x_0)$, ce qui est impossible pour $x \in]x_0, x_0 + \varepsilon[$. Ainsi, x_0 est un minimum global.

Considérons maintenant $x, y \in I$ tels que $x < y < x_0$. L'inégalité des pentes assure que

$$\frac{f(y) - f(x)}{y - x} \leq \frac{f(x_0) - f(x)}{x_0 - x} \leq 0 \text{ (car } f(x) \geq f(x_0))$$

et donc $f(y) \leq f(x)$: f est décroissante à gauche de x_0. On montre de la même manière que f est croissante à droite de x_0.

• Supposons maintenant que f ne possède aucun minimum local. Si f n'est pas monotone, alors on peut trouver $x < y < z$ dans I tels que $f(x), f(z) < f(y)$ ou $f(x), f(z) > f(y)$. La première hypothèse est exclue, car l'inégalité des pentes impliquerait alors

$$0 < \frac{f(y) - f(x)}{y - x} \leq \frac{f(z) - f(y)}{z - y} < 0$$

ce qui est absurde. Ainsi, on a $f(x), f(z) > f(y)$. D'après le corollaire 15 à venir, f est continue sur $[x, z]$ et elle présente donc un minimum absolu sur $[x, z]$. Ce dernier ne peut être ni en x ni en z puisque $f(x), f(z) > f(y)$. Il est donc dans $]x, z[$ et est donc minimum local de f sur I, ce qui est contraire aux hypothèses.

Remarque : Le sens direct de l'équivalence

$$f \text{ n'est pas monotone} \iff \exists x, y, z \in I,\ x < y < z,\ f(x), f(z) < f(y) \text{ ou } f(x), f(z) > f(y)$$

s'obtient en constant que "f n'est pas monotone $\iff$ f n'est pas croissante et f n'est pas décroissante $\iff \exists a < b$, $f(a) > f(b)$ et $\exists c < d$, $f(d) > f(c)$" et à regarder comme se positionnent a, b, c, d et $f(a), f(b), f(c), f(d)$ les uns par rapport aux autres respectivement. L'étude exhaustive des cas permet de fournir les triplets (x, y, z) recherchés : si $a < b \leq c < d$, alors soit $f(c) \leq f(b)$ et l'on prend $x = a$, $y = c$ et $z = d$ et l'on a $f(x), f(z) > f(y)$, soit $f(c) > f(b)$ et l'on prend $x = a$, $y = b$ et $z = d$ pour avoir $f(x), f(z) > f(y)$. On traite de la même manière les cas $a \leq c < b \leq d$, $a \leq c < d \leq b$, $c < a \leq d < b$ etc.

□

2.3 Convexité et dérivabilité.

Rappel : Limites et dérivées en un point.

On considère un intervalle ouvert I.

Définition.— *Soient $f : I \longrightarrow \mathbb{R}$ une fonction, $x_0 \in I$ et $\ell \in \mathbb{R}$. On dit que f possède ℓ pour limite (resp. limite épointée, resp. limite gauche, resp. limite droite) au point x_0 si*

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \left\{ \begin{array}{l} \forall x \in I \\ (\textit{resp. } \forall x \in I,\ x \neq x_0) \\ (\textit{resp. } \forall x \in I,\ x < x_0) \\ (\textit{resp. } \forall x \in I,\ x > x_0) \end{array} \right. ,\ |x - x_0| \leq \alpha \Longrightarrow |f(x) - \ell| \leq \varepsilon$$

et l'on note alors $\ell = \lim\limits_{x \to x_0} f(x)$ (resp. $\ell = \lim\limits_{x \to x_0^\bullet} f(x)$ ou encore $\ell = f(x_0^\bullet)$, resp. $\ell = \lim\limits_{x \to x_0^-} f(x)$ ou encore $\ell = f(x_0^-)$, resp. $\ell = \lim\limits_{x \to x_0^+} f(x)$ ou encore $\ell = f(x_0^+)$).

Remarques : 1/ Si f possède ℓ pour limite en x_0 alors nécessairement $\ell = f(x_0)$.

2/ f est continue à gauche (resp. à droite) en $x_0 \Longleftrightarrow f(x_0^-) = f(x_0)$ (resp. $f(x_0^+) = f(x_0)$).

Théorème.— (dit de la limite monotone) *Si $f : I \longrightarrow \mathbb{R}$ est une fonction monotone, alors f possède en tout point une limite droite et une limite gauche. Plus précisément, si f est, par exemple, croissante, alors on a pour tout $x_0 \in I$,*

$$f(x_0^-) \leq f(x_0) \leq f(x_0^+)$$

Preuve : L'ensemble $V^- = \{f(x)/\ x < x_0\}$ est une partie non vide et majorée (par $f(x_0)$), elle possède donc une borne supérieure $\ell \leq f(x_0)$. Par caractérisation de la borne supérieure dans $\mathbb{R}$, on a, $\forall \varepsilon > 0$, $\exists y \in V^-$, $\ell - \varepsilon < y \leq \ell$. Soit $x_1 < x_0$ tel que $y = f(x_1)$ et $\alpha = x_0 - x_1$. Si $x \in I$ est tel que $x < x_0$, alors $|x - x_0| \leq \alpha \Longrightarrow x_1 \leq x < x_0$ et donc, par croissance de f, $\ell - \varepsilon < y = f(x_1) \leq f(x) \leq \ell$. On a donc $|f(x) - \ell| \leq \varepsilon$ et donc $f(x_0^-) = \ell$. L'existence de $f(x_0^+)$ se montre de la même manière en considérant la borne inférieure de l'ensemble $V^+ = \{f(x)/\ x > x_0\}$.

□

Définition.— *Soient $f : I \longrightarrow \mathbb{R}$ une fonction, et $x_0 \in I$. On dit que f est dérivable à gauche (resp. à droite) en x_0 si la fonction "taux de variation"*

$$T(x) = \left\{ \begin{array}{ll} \dfrac{f(x) - f(x_0)}{x - x_0} & \text{si } x \neq x_0 \\ 0 & \text{si } x \neq x_0 \end{array} \right.$$

possède une limite gauche (resp. droite) en x_0. Si elle existe, on note $f_g'(x_0)$ (resp. $f_d'(x_0)$) cette limite.

Remarques : 1/ Si f est dérivable à gauche (resp. à droite) en x_0 alors f est continue à gauche (resp. à droite) en x_0. En effet, sur $I \cap]-\infty, x_0[$, on a $f(x) - f(x_0) = f_g'(x_0)(x - x_0) + o(x - x_0)$ et l'on a donc, $\lim\limits_{x \to x_0^-} f(x) - f(x_0) = 0$.

2/ f est dérivable en x_0 si et seulement si $f_g'(x_0)$ et $f_d'(x_0)$ existent et sont égales et, dans ces conditions $f'(x_0) = f_g'(x_0) = f_d'(x_0)$.

Dans ce qui suit, I désigne un intervalle ouvert de $\mathbb{R}$ et $f : I \longrightarrow \mathbb{R}$ une application.

Théorème 14.— *Si f est convexe alors f est dérivable à droite et à gauche sur I (en particulier f est continue). De plus, pour tout $s, t, x \in I$ tels que $s < x < t$ on a*

$$\frac{f(x) - f(s)}{x - s} \leq f_g'(x) \leq f_d'(x) \leq \frac{f(t) - f(x)}{t - x}$$

Les applications f_d' et f_g' sont croissantes.

Preuve : L'inégalité des pentes assure que $\dfrac{f(x) - f(s)}{x - s} \leq \dfrac{f(t) - f(x)}{t - x}$ et de plus que l'application $p_x : t \longmapsto \dfrac{f(t) - f(x)}{t - x}$ est croissante sur $I \cap]x, +\infty[$ et que l'application $q_x : s \longmapsto \dfrac{f(x) - f(s)}{x - s}$ est croissante sur $I \cap]-\infty, x[$.

Le théorème de la limite monotone assure que p_x a une limite à droite en x et donc que $f_d^{'}(x)$ existe et que l'on a de plus

$$\frac{f(x)-f(s)}{x-s} \leq f_d^{'}(x) \leq \frac{f(t)-f(x)}{t-x}$$

En appliquant de nouveau le théorème de la limite monotone, on voit que q_x a une limite à gauche en x et donc que $f_g^{'}(x)$ existe et que l'on a de plus

$$\frac{f(x)-f(s)}{x-s} \leq f_g^{'}(x) \leq f_d^{'}(x) \leq \frac{f(t)-f(x)}{t-x}$$

Soient $x, y \in I$ tels que $x < y$. Pour tout $u \in]x, y[$ on a

$$f_g^{'}(x) \leq f_d^{'}(x) \leq \frac{f(u)-f(x)}{u-x} \leq \frac{f(y)-f(u)}{y-u} \leq f_g^{'}(y) \leq f_d^{'}(y)$$

ce qui assure la croissance des applications $f_g^{'}$ et $f_d^{'}$.

□

Corollaire 15.— *Une fonction convexe sur un intervalle ouvert est continue.*

Preuve : D'après ce qui précède f est dérivable à droite et à gauche en tout point de I, elle est donc continue à droite et à gauche en tout point, c'est-à-dire tout simplement continue.

□

Attention : Ce dernier corollaire est faux si l'on ne suppose plus I ouvert. Par exemple, la fonction $x \longmapsto [x]$ est convexe sur $I = [0, 1]$, mais discontinue en 1.

Corollaire 16.— *Si f est dérivable alors les propriétés suivantes sont équivalentes*

i) f est convexe,

ii) $f^{'}$ est croissante.

Preuve : $i) \Rightarrow ii)$ D'après le théorème 14 pour tout $x < y$ dans I, on a pour tout $u \in]x, y[$

$$f^{'}(x) = f_g^{'}(x) = f_d^{'}(x) \leq \frac{f(u)-f(x)}{u-x} \leq \frac{f(y)-f(u)}{y-u} \leq f_g^{'}(y) = f_d^{'}(y) = f^{'}(y)$$

La fonction dérivée est donc croissante.

$ii) \Rightarrow i)$

Rappel : Théorème des accroissements finis.

C'est le théorème suivant :

Théorème.— *Si $f : [a, b] \longrightarrow \mathbb{R}$ désigne une fonction continue sur $[a, b]$ et dérivable sur l'ouvert $]a, b[$ alors, il existe $c \in]a, b[$ tel que $f(b) - f(a) = f^{'}(c)(b - a)$.*

Considérons $x < y < z$ trois éléments de I. D'après le théorème des accroissements finis, appliqués deux fois, il existe $c \in]x, y[$ et $d \in]y, z[$ tels que

$$\begin{cases} \dfrac{f(y)-f(x)}{y-x} = f^{'}(c) \\ \dfrac{f(z)-f(y)}{z-x} = f^{'}(d) \end{cases}$$

et, puisque f' est croissante et que $x < c < y < d < z$, on a

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

ce qui assure par le théorème 11 et la proposition 12 que f est convexe.

□

Corollaire 17.— *Si f est deux fois dérivable alors les propriétés suivantes sont équivalentes*

i) f est convexe,

ii) f'' est positive sur I.

Preuve : Immédiat.

□

Corollaire 18.— *Si f est dérivable alors les propriétés suivantes sont équivalentes*

i) f est convexe,

ii) le graphe de f est situé au dessus de chacune de ses tangentes.

Preuve : $i) \Rightarrow ii)$ Soit $a \in I$. Si $t < a$ alors $\dfrac{f(t)-f(a)}{t-a} \leq f'(a)$ et si $t > a$ alors $\dfrac{f(t)-f(a)}{t-a} \geq f'(a)$. Donc pour tout $t \in I$ on a $f(t) \geq f(a)+(t-a)f'(a)$.

$ii) \Rightarrow i)$ Supposons que pour tout $a, t \in I$ on ait $f(t) \geq f(a)+(t-a)f'(a)$. Soient $a < b$ dans I, on a donc

$$\begin{cases} f(b) \geq f(a)+(b-a)f'(a) \\ f(a) \geq f(b)+(a-b)f'(b) \end{cases}$$

ce qui donne, après addition et simplification,

$$0 \geq (b-a)(f'(a)-f'(b))$$

Puisque $b-a > 0$, on a alors $f'(a)-f'(b) \leq 0$, ce qui prouve que f' croissante et, grâce au corollaire 16, que f est convexe.

□

2.4 Sur l'ensemble des points de non dérivabilité d'une fonction convexe.

On s'intéresse ici aux points de non dérivabilité des fonctions convexes. On sait qu'une fonction convexe est dérivable à droite et à gauche en tout point. Ses points de non dérivabilité correspondent donc aux points où les dérivées droite et gauche ne sont pas égales.

Lemme 19.— *Soit $f : I \longrightarrow \mathbb{R}$ une fonction convexe et $x_0 \in I$. Pour que f soit dérivable en x_0 il suffit que la fonction f'_g (resp. f'_d) soit continue à droite (resp. à gauche) en x_0.*

Preuve : Le théorème 14 assure que si $x_0 < x$ alors $f'_g(x_0) \leq f'_d(x_0) \leq f'_g(x) \leq f'_d(x)$. En faisant tendre x vers x_0, On voit donc que, si $\lim\limits_{x \to x_0^+} f'_g(x) = f'_g(x_0)$ (c'est-à-dire si f'_g est continue à droite en x_0), on a, par passage à la limite de l'encadrement, $f'_g(x_0) = f'_d(x_0)$ et donc f est dérivable en x_0.

□

Théorème 20.— *Une fonction convexe est dérivable sauf éventuellement sur un ensemble dénombrable de points.*

Preuve :

Rappel : fonctions réglées.

Définition.— *Une fonction $f : [a, b] \longrightarrow \mathbb{R}$ définie sur un segment $[a, b]$ est dite "réglée", si elle est la limite uniforme d'une suite de fonctions en escalier de $[a, b]$.*

Théorème.— *Toute fonction monotone sur un segment est réglée.*

Discontinuité : Soient $f : [a, b] \longrightarrow \mathbb{R}$ une fonction réglée et $(e_n)_n$ et une suite de fonctions en escalier convergeant uniformément vers f. Pour tout $n \geq 0$, l'ensemble D_n des points de discontinuité de e_n est fini, si bien que l'ensemble $D = \bigcup_n D_n$ est dénombrable. Si $x_0 \notin D$, alors chaque e_n est continue en x_0 et, par passage à la limite uniforme, on en déduit que f est continue en x_0. Ainsi, l'ensemble des points de discontinuité de f est inclus dans D. On vient ainsi de montrer qu'une fonction réglée ne possède qu'un nombre dénombrable de points de discontinuité.

Puisque tout intervalle de $\mathbb{R}$ est une réunion dénombrable de segments et qu'une réunion dénombrable d'ensembles dénombrables est un ensemble dénombrable, on en déduit que si f est une fonction définie sur un intervalle I qui est réglée sur tout segment inclus dans I, alors l'ensemble des points de discontinuité de f est dénombrable. En combinant ce résultat au théorème ci-dessus, on trouve alors

Corollaire.— *Toute fonction monotone sur un intervalle ne possède qu'un nombre dénombrable de points de discontinuité.*

La fonction f_g' est monotone, elle est donc réglée sur tout segment et son ensemble de points de discontinuité est donc dénombrable. Le lemme précédent permet alors de conclure.

□

Théorème 21.— *Si $I =]\alpha, \beta[$ désigne un intervalle ouvert borné, alors pour tout partie dénombrable $C \subset I$, il existe une fonction convexe $f : I \longrightarrow \mathbb{R}$ dont l'ensemble des points de non dérivabilité vaut exactement C.*

Preuve : On énumère les éléments de C en une suite $(a_n)_n$ et l'on considère la fonction

$$f(x) = \sum_{n \geq 0} \frac{|x - a_n|}{2^n}$$

Puisque I est borné, $\sup(|x - a_n| / \, x \in I) \leq (\beta - \alpha)$ et donc la série incriminée converge normalement. La fonction f est donc bien définie. Pour $n \geq 0$, introduisons la fonction $\varepsilon_n : \mathbb{R} \longrightarrow \mathbb{R}$ définie par

$$\varepsilon_n(x) = \begin{cases} 1 & \text{si } x > a_n \\ -1 & \text{si } x < a_n \\ 0 & \text{si } x = a_n \end{cases}$$

On a donc $f(x) = \displaystyle\sum_{n \geq 0} \frac{(x - a_n)\varepsilon_n(x)}{2^n}$. Fixons, $x_0 \in I$ et, pour tout $x \neq x_0$ évaluons le taux de variation

$$\begin{aligned} T(x) &= \frac{f(x)-f(x_0)}{x-x_0} = \sum_{n\geq 0} \frac{(x-a_n)\varepsilon_n(x)-(x_0-a_n)\varepsilon_n(x_0)}{2^n(x-x_0)} \\ &= \sum_{n\geq 0} \frac{(x-x_0)\varepsilon_n(x_0)+(x-a_n)(\varepsilon_n(x)-\varepsilon_n(x_0))}{2^n(x-x_0)} \\ &= \sum_{n\geq 0} \frac{\varepsilon_n(x_0)}{2^n} + \sum_{n\geq 0} \frac{g_n(x)}{2^n} \end{aligned}$$

où $g_n(x) = (x-a_n)\dfrac{\varepsilon_n(x)-\varepsilon_n(x_0)}{(x-x_0)}$ (la première série convergeant absolument, ceci assure que la deuxième série converge aussi).

• Si $x_0 > a_n$ alors $g_n(x) = 0$ sur $[a_n, \beta[$ et, en particulier g_n est prolongeable par continuité en x_0 en posant $g_n(x_0) = 0$. Sur $]\alpha, a_n]$, on a $g_n(x) = \dfrac{-2(x-a_n)}{(x-x_0)}$ qui est une fonction monotone. On en déduit que $\|g_n\|_\infty = \dfrac{2(a_n-\alpha)}{(x_0-\alpha)} \leq \dfrac{2(\beta-\alpha)}{(x_0-\alpha)}$.

• Si $x_0 < a_n$, on montre de la même manière que $\|g_n\|_\infty = \dfrac{2(\beta-a_n)}{(\beta-x_0)} \leq \dfrac{2(\beta-\alpha)}{(\beta-x_0)}$.

En conclusion, si pour tout $n \geq 0$ on a $x_0 \neq a_n$ (i.e. si $x_0 \notin C$), alors la suite $(\|g_n\|_\infty)_n$ est bornée et donc la série $\sum \dfrac{g_n(x)}{2^n}$ converge normalement sur I (après avoir pris soin de prolonger par continuité chaque g_n en x_0). On peut donc écrire

$$\lim_{x\to x_0}\left(T(x) - \sum_{n\geq 0}\frac{\varepsilon_n(x_0)}{2^n}\right) = \lim_{x\to x_0}\left(\sum_{n\geq 0}\frac{g_n(x)}{2^n}\right) = \sum_{n\geq 0}\lim_{x\to x_0}\frac{g_n(x)}{2^n} = \sum_{n\geq 0} 0 = 0$$

ce qui montre que f est dérivable en tout $x_0 \notin C$ et vérifie $f^{'}(x_0) = \displaystyle\sum_{n\geq 0}\frac{\varepsilon_n(x_0)}{2^n}$.

Considérons maintenant $x_0 \in C$. Par définition, il existe un indice $n_0 \geq 0$ tel que $x_0 = a_{n_0}$. On a alors $\dfrac{|x-x_0|}{2^{n_0}} = f(x) - \displaystyle\sum_{n\neq n_0}\frac{|x-a_n|}{2^n}$. En appliquant ce qui vient d'être montré à la suite $(a_n)_{n\neq n_0}$, on voit que la fonction $x \longmapsto \sum_{n\neq n_0}\frac{|x-a_n|}{2^n}$ est dérivable en x_0. Si f était dérivable en x_0 il en serait de même de $x \longmapsto |x-x_0|$, ce qui est bien sur absurde. Finalement f n'est dérivable en aucun point de C.

□