Convexité

Cornou Jean-Louis

28 décembre 2022

☞ **Remarque**
Illustrations à produire.

1 Parties convexes et barycentres

La notion de barycentre est introduite dans le seul but de caractériser les parties convexes. Aussi ce cours est-il volontairement succinct sur le sujet.

1.1 Barycentres

Dans toute la suite, le symbole E désigne l'espace $\mathbb{R}^2$. On y dispose des opérations habituelles comme pour les vecteurs

$$\forall (x_1, y_1) \in \mathbb{R}^2, \forall (x_2, y_2) \in \mathbb{R}^2, (x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2)$$

$$\forall \lambda \in \mathbb{R}, \forall (x_1, y_1) \in \mathbb{R}^2, \lambda(x_1, y_1) = (\lambda x_1, \lambda y_1)$$

On note également $0_{\mathbb{R}^2} = 0_E = (0, 0)$.

Définition 1 *On appelle système (fini) de points massiques de* E *toute famille finie* $((A_i, \alpha_i))_{i\in I}$ *où* $A_i \in E$ *et* $\alpha_i \in \mathbb{K}$ *pour tout* $i \in I$.

Le concept de barycentre cherche à définir la notion de «point d'équilibre », à l'instar d'un fléau de balance de Roberval au bout duquel deux masses seraient disposées.

Notation
Si $A, B \in E$ *on notera* $\overrightarrow{AB}$ *le vecteur* $B - A$. *Le lecteur remarquera que* $\overrightarrow{AA} = 0_E$ *et* $\overrightarrow{AB} + \overrightarrow{BC} = \overrightarrow{AC}$ *quels que soient* $A, B, C \in E$, *ainsi que l'agréable relation* $A + \overrightarrow{AB} = B$.

Théorème 1 *Soit* $((A_i, \alpha_i))_{i\in I}$ *un système fini de points massiques d'un* $\mathbb{K}$*-espace vectoriel* E. *Les deux assertions suivantes sont équivalentes.*

1. *Il existe un unique point* G *de* E *tel que* $\sum_{i\in I} \alpha_i \overrightarrow{A_iG} = 0_E$.
2. $\sum_{i\in I} \alpha_i \neq 0$

De plus, quand ces assertions sont vraies, $G = \frac{1}{\sum_{i\in I} \alpha_i} \sum_{i\in I} \alpha_i A_i$. *On dit que* G *est le barycentre du système massique* $((A_i, \alpha_i))_{i\in I}$, *ou encore ou que c'en est le centre de gravité.*

Démonstration. Posons, pour tout point $M \in E, \Lambda(M) = \sum_{i\in I} \alpha_i \overrightarrow{A_iM}$ (c'est la fonction vectorielle de Leibniz). Pour tous points M, M' de E,

$$\Lambda(M) = \sum_{i\in I} \alpha_i \left(\overrightarrow{A_iM'} + \overrightarrow{M'M}\right) = \Lambda\left(M'\right) + \left(\sum_{i\in I} \alpha_i\right) \overrightarrow{M'M}.$$

Si $\sum_{i\in I} \alpha_i = 0$, la fonction Λ est donc constante : elle s'annule partout ou jamais, mais pas en un unique point. Si $\sum \alpha_i \neq 0$, prenons $M' = A_{i_0}$ avec $i_0 \in I$ quelconque. La relation ci-dessus permet de dire que l'équation $\Lambda(M) = 0_E$ équivaut à $(\sum_{i\in I} \alpha_i) \overrightarrow{A_{i_0}M} = -\Lambda\left(A_{i_0}\right)$ qui admet l'unique solution

$$\begin{aligned}\mathrm{M} = \mathrm{A}_{i_0} - \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\Lambda\left(\mathrm{A}_{i_0}\right) &= \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i\mathrm{A}_{i_0} - \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i\overrightarrow{\mathrm{A}_i\mathrm{A}_{i_0}} \\ &= \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i\left(\mathrm{A}_{i_0} - \overrightarrow{\mathrm{A}_i\mathrm{A}_{i_0}}\right) = \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i\mathrm{A}_i,\end{aligned}$$

ce qui prouve notre théorème.

Corollaire (coordonnées du barycentre)
On note pour chaque point A_i *ses coordonnées* $\left(x_1^{(i)}, x_2^{(i)}\right)$. *Lorsqu'il est défini, le barycentre* G *de* $((\mathrm{A}_i,\alpha_i))_{i\in\mathrm{I}}$ *a pour coordonnées*

$$x_1 = \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i x_1^{(i)}, x_2 = \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i x_2^{(i)}.$$

Définition 2 *Quand toutes les masses sont identiques, et non nulles, on parle d'isobarycentre. Ainsi, l'isobarycentre des points* $\mathrm{A}_1,\ldots,\mathrm{A}_n$ *est le point* G *défini* par $\mathrm{G} = \frac{1}{n}\sum_{i=1}^n \mathrm{A}_i$

Théorème 2 *Soit* $((\mathrm{A}_i,\alpha_i))_{i\in\mathrm{I}}$ *un système fini de points massiques admettant un barycentre* G. *Alors pour tout* $k\in\mathbb{R}^*$, $((\mathrm{A}_i,k\alpha_i))_{i\in\mathrm{I}}$ *admet* G *comme barycentre.*

Démonstration. Soit k un réel non nul, alors $\sum_{i\in\mathrm{I}}(k\alpha_i) = k\sum_{i\in\mathrm{I}}\alpha_i \neq 0$. Le théorème résulte alors de

$$\frac{1}{\sum_{i\in\mathrm{I}}k\alpha_i}\sum_{i\in\mathrm{I}}k\alpha_i\mathrm{A}_i = \frac{1}{\sum_{i\in\mathrm{I}}\alpha_i}\sum_{i\in\mathrm{I}}\alpha_i\mathrm{A}_i.$$

1.2 Parties convexes

Définition 3 *Soit* A, B *deux points de* E, *le segment d'extrémités* A *et* B *est l'ensemble*

$$\{t\mathrm{A} + (1-t)\mathrm{B} | t\in[0,1]\}$$

Notation
Ce segment est noté $[\mathrm{A},\mathrm{B}]$.

⚠ **Attention**
Cette notation est en conflit avec la notion des intervalles de $\mathbb{R}$. *Dans le cas de* $\mathbb{R}^2$, $[\mathrm{A},\mathrm{B}] = [\mathrm{B},\mathrm{A}]$ *puisque* $[0,1]\to[0,1], t\mapsto(1-t)$ *est une bijection. Dans le cas de* $\mathbb{R}$, $[1,0] = \emptyset$ *tandis que* $[0,1]$ *est non vide. Certains auteurs notent dans le cadre réel* "$[x,y]$" $= [\min(x,y),\max(x,y)]$ *le segment réel d'extrémités* x *et* y.

Définition 4 *Soit* C *une partie de* E. *On dit que* C *est convexe (ou une partie convexe de* E*) lorsque*

$$\forall(\mathrm{A},\mathrm{B})\in\mathrm{C}^2, [\mathrm{A},\mathrm{B}]\subset\mathrm{C}$$

☞ **Remarque**
On a vu dans le chapitre sur la topologie de $\mathbb{R}$ *que les seuls convexes de* $\mathbb{R}$ *sont les intervalles.*

Propriété 1 *Une partie* C *de* E *est convexe si et seulement si*

$$\forall(\mathrm{A},\mathrm{B})\in\mathrm{C}^2, \forall t\in[0,1], (1-t)\mathrm{A} + t\mathrm{B}\in\mathrm{C}$$

Exemple 1 *L'ensemble vide est convexe. Toute droite affine de* E *est convexe. L'espace* E *est lui-même convexe. Un cercle de rayon non nul n'est pas convexe. Une boule ouverte (ou fermée) est convexe.*

Exercice 1 *Montrer qu'une intersection quelconque de parties convexes est convexe. Est-ce le cas pour une union ?*

Correction 1 *Soit* $(\mathrm{C}_i)_{i\in\mathrm{I}}$ *une famille de parties convexes. Notons* $\mathrm{C} = \bigcap_{i\in\mathrm{I}}\in\mathrm{C}_i$. *Soit* $(\mathrm{A},\mathrm{B})\in\mathrm{C}^2$ *et* $t\in[0,1]$. *Soit* $i\in\mathrm{I}$. *Alors* A *et* B *appartiennent à* C_i *puisqu'ils sont tous deux dans l'intersection des* $(\mathrm{C}_i)_{i\in\mathrm{I}}$. *Comme la partie* C_i *est convexe,* $(1-t)\mathrm{A}+t\mathrm{B}$ *appartient à* C_i *et ce pour tout indice* i *dans* I. *Ainsi,* $(1-t)\mathrm{A}+t\mathrm{B}$ *appartient à* C. *Par conséquent, la partie* C *est convexe.*
C'est évidemment faux pour l'union. Soit A *et* B *deux points distincts de* E. *Alors les singletons* $\{\mathrm{A}\}$ *et* $\{\mathrm{B}\}$ *sont tous deux convexes. Toutefois, leur union n'est pas convexe, puisque* $\frac{1}{2}\mathrm{A}+\frac{1}{2}\mathrm{B}$ *est distinct de* A *et de* B, *donc n'appartient pas à* $\{\mathrm{A},\mathrm{B}\}$.

1.3 Caractérisation des parties convexes

Exemple 2 *Soit* A, B *deux points de* $\mathbb{R}^n$. *L'ensemble des barycentres que l'on peut faire avec* A *et* B *est la droite* (AB). *En effet, ces barycentres sont de la forme* $\frac{1}{a+b}(aA+bB)$ *avec* $(a,b)\in\mathbb{R}^2$ *tel que* $a+b\neq 0$. *En posant* $\lambda=\frac{b}{a+b}$, *on obtient l'expression* $G_\lambda=(1-\lambda)A+\lambda B=A+\lambda\overrightarrow{AB}$. *Quand* a,b *sont quelconques,* λ *parcourt* $\mathbb{R}$ *et* G_λ *décrit la droite passant* A *dirigée par* $\overrightarrow{AB}$, *c'est-à-dire la droite* (AB).
Si l'on impose à λ *de ne parcourir que* $[0,1]$, G_λ *parcourt le segment* [AB]. *Cette condition est équivalente à dire que a et b sont de même signe (laissé à titre d'exercice).*

Exemple 3 *Soit trois points* A, B, C *non alignés dans* $\mathbb{R}^2$. *L'ensemble des barycentres que l'on peut former avec* A, B, C *est le plan* $(ABC)=\mathbb{R}^2$. *En effet, ceux-ci sont tous de la forme* $\frac{1}{a+b+c}(aA+bB+cC)$ *avec* $(a,b,c)\in\mathbb{R}^3$ *tels que* $a+b+c\neq 0$. *En posant* $\lambda=\frac{b}{a+b+c}$ *et* $\mu=\frac{c}{a+b+c}$, *cette expression devient* $(1-\lambda-\mu)A+\lambda B+\mu C$ *soit encore* $A+\lambda\overrightarrow{AB}+\mu\overrightarrow{AC}$: *quand* a,b,c *décrivent* $\mathbb{R}$, (λ,μ) *décrit* $\mathbb{R}^2$ *si bien que* $A+\lambda\overrightarrow{AB}+\mu\overrightarrow{AC}$ *décrit le plan* (ABC).
Si l'on veut que les barycentres parcourent l'intérieur du triangle ABC, *on doit imposer les conditions* $\lambda\geqslant 0,\mu\geqslant 0$ *et* $\lambda+\mu\leqslant 1$. *Cela revient encore à imposer à* a,b,c *d'être de même signe.*

Dans les deux exemples précédents, le barycentre du système massique considéré peut toujours s'écrire $\sum_{i\in I}\lambda_i A_i$ avec $\sum_{i\in I}\lambda_i=1$: $\lambda_1=\frac{a}{a+b},\lambda_2=\frac{b}{a+b}$ pour l'exemple 1, $\lambda_1=\frac{a}{a+b+c},\lambda_2=\frac{b}{a+b+c},\lambda_3=\frac{c}{a+b+c}$ pour l'exemple 2 .

Définition 5 *Soit* $n\in\mathbb{N}^*$ *et* $A_1,\dots,A_n$ *des points de* $E=\mathbb{R}^2$. *On appelle combinaison linéaire convexe (abrégé en CLC) de* $A_1,\dots,A_n$ *tout élément de* E *de la forme* $\sum_{i=1}^n\lambda_i A_i$ *avec* $\lambda_i\in\mathbb{R}_+$*pour tout* $i\in[\![1,n]\!]$ *et* $\sum_{i=1}^n\lambda_i=1$.

On vient de prouver que l'ensemble des CLC de A et B est le segment [A, B] et que celui de A, B, C est l'intérieur du triangle ABC.

Théorème 3 (caractérisation des convexes par les barycentres) *Soit* C *une partie non vide de* E. *Les propriétés suivantes sont équivalentes.*

1. *La partie* C *est convexe.*
2. *La partie* C *est stable par combinaison linéaire convexe.*
3. *Tout système fini* $((A_i,\alpha_i))_{i\in I}$ *de points massiques dans* C *pondérés par des masses positives et non identiquement nulles a un barycentre dans* C.

Démonstration. 1 ⇒ 2 Si C est convexe, alors toute CLC de deux points A_1 et A_2 est encore dans C, puis l'ensemble de ces CLC n'est autre que [AB] (cf. exemple ci-dessus). Soit $n\geq 2$. Supposons que toute CLC de n points de C soit encore dans C et considérons $A_1,\dots,A_{n+1}$ dans C. Une CLC de ces $n+1$ points s'écrit $M=\sum_{i=1}^{n+1}\lambda_i A_i$ avec et $\lambda_1,\dots,\lambda_{n+1}\in\mathbb{R}_+$de somme 1 . Les réels λ_i ne pouvant manifestement pas être tous égaux à 1 , on dispose d'un entier i_0 tel que $\lambda_{i_0}\neq 1$ si bien que $\sum_{\substack{i=1\\ i\neq i_0}}^{n+1}\lambda_i\neq 0$: notons m cette somme non nulle. Alors, $m>0$ et $M=\sum_{i=1}^{n+1}\lambda_i A_i=m\sum_{\substack{i=1\\ i\neq i_0}}^{n+1}\frac{\lambda_i}{m}A_i+\lambda_{i_0}A_{i_0}=mB+(1-m)A_{i_0}$, où $B=\sum_{\substack{i=1\\ i\neq i_0}}^{n+1}\frac{\lambda_i}{m}A_i\in C$ par HR. Ainsi, M est CLC de B et A_{i_0} qui sont deux points de C, donc $M\in C$ et la récurrence s'achève.

2 ⇒ 3 Tout barycentre à masses positives non identiquement nulles est une CLC par homogénéité

3 ⇒ 1 Si $A,B\in C$ tout point de [AB] est un barycentre à coefficients positifs de A et B, c'est donc un donc un point de C.

2 Fonctions convexes

Le programme se limite aux fonctions convexes d'une variable réelle.

2.1 Définition et caractérisations

Définition 6 *Soit* I *un intervalle de longueur non nulle de* $\mathbb{R}$ *et* $f:I\to\mathbb{R}$ *une fonction à valeurs réelles. On dit que* f *est une fonction convexe sur* I *quand*

$$\forall (x,y) \in \mathrm{I}^2, \forall \lambda \in [0,1], \quad f((1-\lambda)x+\lambda y) \leqslant (1-\lambda)f(x)+\lambda f(y).$$

On dit que f est concave sur I quand $-f$ est convexe sur I.

☞ Remarque

Il n'y a pas de notion de partie concave, pourtant rencontrée dans le langage courant.

Définition 7 *Si $f : \mathrm{I} \to \mathbb{R}$ est une fonction définie sur un intervalle I, on appelle épigraphe de f l'ensemble des couples (x,y) de $\mathrm{I} \times \mathbb{R}$ tels que $y \geqslant f(x)$. On le note $\mathrm{Epi}(f)$.*

Ainsi, Epi (f) est l'ensemble des points sont au-dessus de la courbe de f.

Théorème 4 *La fonction f est convexe sur I si et seulement si son épigraphe est une partie convexe de $\mathbb{R}^2$.*

Démonstration. Si f est convexe sur I, considérons deux points $\mathrm{A} = (x_\mathrm{A}, y_\mathrm{A})$ et $\mathrm{B} = (x_\mathrm{B}, y_\mathrm{B})$ de $\mathrm{Epi}(f)$. Un point M quelconque de [AB] s'écrit $(1-\lambda)\mathrm{A}+\lambda\mathrm{B}$ avec $\lambda \in [0,1]$: ses coordonnées sont donc $x_\mathrm{M} = (1-\lambda)x_\mathrm{A}+\lambda x_\mathrm{B}$ et $y_\mathrm{M} = (1-\lambda)y_\mathrm{A}+\lambda y_\mathrm{B}$. Comme f est une fonction convexe, $f(x_\mathrm{M}) \leqslant (1-\lambda)f(x_\mathrm{A})+\lambda f(x_\mathrm{B})$ et comme A et B sont dans $\mathrm{Epi}(f), f(x_\mathrm{A}) \leqslant y_\mathrm{A}$ et $f(x_\mathrm{B}) \leqslant y_\mathrm{B}$ si bien qu'en sommant, $f(x_\mathrm{M}) \leqslant y_\mathrm{M}$ ce qui prouve que C est dans $\mathrm{Epi}(f)$. Ainsi, l'épigraphe de f contient tous les points de [AB], il est donc convexe.
Réciproquement, si $\mathrm{Epi}(f)$ est convexe, considérons $x, y \in \mathrm{I}$ et $\lambda \in [0,1]$. Les points $\mathrm{A} = (x, f(x))$ et $\mathrm{B} = (y, f(y))$ étant dans $\mathrm{Epi}(f)$, le point $\mathrm{C} = (1-\lambda)\mathrm{A}+\lambda\mathrm{B}$ l'est aussi : la traduction de cela traduit exactement la convexité de f.

Corollaire (inégalité de Jensen discrète))
La fonction $f : \mathrm{I} \to \mathbb{R}$ est convexe sur I si et seulement si pour tout $n \in \mathbb{N}^*$, tout $(x_1, \dots, x_n) \in \mathrm{I}^n$ et tous réels positifs $\lambda_1, \dots, \lambda_n$ tels que $\sum_{i=1}^n \lambda_i = 1$:

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leqslant \sum_{i=1}^n \lambda_i f(x_i).$$

Notons que I étant un intervalle, c'est une partie convexe de $\mathbb{R}$, donc contient toute CLC d'éléments de I, c'est-à-dire $f(\sum_{i=1}^n \lambda_i x_i)$ a bien un sens.

Démonstration. Si f vérifie l'inégalité de Jensen pour tout $n \in \mathbb{N}^*$, elle la vérifie pour $n = 2$, et ceci n'est autre que la définition d'une fonction convexe. Réciproquement, si f est convexe sur I, son épigraphe est une partie convexe, donc stable par CLC d'après la caractérisation des parties convexes (§ précédent). Or les points $\mathrm{A}_i = (x_i, f(x_i)) (1 \leqslant i \leqslant n)$ sont dans Epi (f), donc $\mathrm{M} = \sum \lambda_i \mathrm{A}_i$ aussi. L'abscisse de M étant $\sum \lambda_i x_i$ et son ordonnée étant $\sum \lambda_i f(x_i)$, dire que $\mathrm{M} \in \mathrm{Epi}(f)$ c'est exactement annoncer l'inégalité voulue.

☞ Remarque

On peut aussi choisir de procéder par récurrence. Dans ce cas on reproduit presque à l'identique la démonstration de la caractérisation des parties convexes du paragraphe précédent.

Corollaire (position des cordes)
Soit I un intervalle de longueur non nulle et $f : \mathrm{I} \to \mathbb{R}$ une fonction. Alors f est convexe ssi pour tous réels $a < b$ dans I,

$$\forall x \in [a,b], f(x) \leqslant \frac{f(b)-f(a)}{b-a}(x-a)+f(a)$$

autrement dit quand la courbe de f est en dessous de chacune de ses cordes.

On appelle corde (ou sécante dans le programme) de la courbe représentative de f tout segment [AB] où A et B sont des points de cette courbe.

Démonstration. Si f est convexe, les extrémités A et B d'une corde sont dans l'épigraphe de f. Celui-ci étant convexe d'après le théorème précédent, toute la corde [AB] est dans l'épigraphe, ce qui traduit la position relative attendue de $\mathscr{C}_f$ par rapport à la corde [AB].
Inversement, si $\mathscr{C}_f$ est en dessous de chacune de ses cordes, soit $a < b$ dans I. Les points $\mathrm{A} = (a, f(a))$ et $\mathrm{B} = (b, f(b))$ définissent une corde [AB] dont une équation est $y = \frac{f(b)-f(a)}{b-a}(x-a)+f(a)(x \in [a,b])$. Si $\lambda \in [0,1]$, posons $c = (1-\lambda)a+\lambda b$. L'hypothèse faite sur f implique que $f(c) \leqslant \frac{f(b)-f(a)}{b-a}(c-a)+f(a)$. Oui mais $c-a = \lambda(b-a)$ donc cette inégalité devient $f(c) \leqslant \lambda(f(b)-f(a))+f(a)$ soit encore $f(c) \leqslant (1-\lambda)f(a)+\lambda f(b)$. La fonction f est donc convexe.

Exercice 2 *Si $f : \mathrm{I} \to \mathbb{R}$ est convexe et admet un minimum local, montrer que ce minimum est global.*

Correction 2 *Soit a un point de I où f admet un minimum local. Si $f(a)$ n'était pas le minimum global de f sur I, il existerait $b \in \mathrm{I}$ (forcément distincts de a) tel que $f(b) < f(a)$. Tout réel strictement compris entre a et b se met sous la forme $(1-\lambda)a+\lambda b$ avec $\lambda \in]0,1[$. La convexité de f donne $f((1-\lambda)a+\lambda b) \leqslant (1-\lambda)f(a)+\lambda f(b) < (1-\lambda)f(a)+\lambda f(a) = f(a)$: si λ est quelconque au voisinage de 0 , on trouve une contradiction avec le fait que f admet en a un minimum local.*

Théorème 5 (croissance des taux d'accroissement) *Soit I un intervalle de longueur non nulle et $f : \mathrm{I} \to \mathbb{R}$ une fonction réelle définie sur I. La fonction f est convexe sur I si et seulement si pour tout $x_0 \in \mathrm{I}$, la fonction*

$$\begin{aligned} \tau_{x_0} : \mathrm{I}\setminus\{x_0\} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{f(x)-f(x_0)}{x-x_0} \end{aligned}$$

est croissante sur $\mathrm{I}\setminus\{x_0\}$

Démonstration. Supposons f convexe sur I et considérons $x_0 \in \mathrm{I}$. Faisons une première constatation : si $a < b < c$ sont dans I, alors b peut s'écrire $b = (1-\lambda)a+\lambda c$ avec $\lambda \in]0,1[$ si bien que $b-a = \lambda(c-a)$ et alors

$$\tau_a(c)-\tau_a(b) = \frac{\lambda[f(c)-f(a)]}{b-a} - \frac{f(b)-f(a)}{b-a} = \frac{(1-\lambda)f(a)+\lambda f(c)-f(b)}{b-a}.$$

Comme $b-a > 0$ et que $f(b) \leqslant (1-\lambda)f(a)+\lambda f(c)$ (par convexité de f), on en déduit l'inégalité $(*)$: $\tau_a(b) \leqslant \tau_a(c)$. Mais ce n'est pas tout! Puisque la seule chose importante a été que b soit entre a et c, les rôles joués par a et c sont symétriques! Le lecteur suspicieux exigera une preuve : en posant $\mu = 1-\lambda$, on a $b = \mu a+(1-\mu)c$ si bien que $b-c = \mu(a-c)$ et cette fois,

$$\tau_c(a)-\tau_c(b) = \frac{\mu[f(a)-f(c)]}{b-c} - \frac{f(b)-f(c)}{b-c} = \frac{(1-\mu)f(c)+\mu f(a)-f(b)}{b-c}.$$

Comme $b-c < 0$ et que $f(b) \leqslant (1-\mu)f(c)+\mu f(a)$ (toujours par convexité de f!), on en tire que $(**)$: $\tau_c(a) \leqslant \tau_c(b)$ comme attendu.
Maintenant, soit $x < y$ dans $\mathrm{I}\setminus\{x_0\}$. Distinguons trois cas :

- Si $x_0 < x < y$, alors $(*)$ donne $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$.
- Si $x < y < x_0$, alors $(**)$ donne $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$.
- Si $x < x_0 < y$, alors $(*)$ donne $\tau_x(x_0) \leqslant \tau_x(y)$ et $(**)$ donne $\tau_y(x) \leqslant \tau_y(x_0)$. Seulement voilà, $\tau_x(y) = \tau_y(x)$ donc par transitivité, $\tau_x(x_0) \leqslant \tau_y(x_0)$ et encore par le truc astucieux ´$\tau_\star(\square) = \tau_\square(\star)$ˇ, on obtient $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$.

Dans tous les cas, $\tau_{x_0}(y)-\tau_{x_0}(x) \geqslant 0$ ce qui établit la croissance de τ_{x_0}.
Réciproquement, supposons que τ_{x_0} soit croissante sur $\mathrm{I}\setminus\{x_0\}$ quel que soit $x_0 \in \mathrm{I}$. Soit alors $x < y$ dans I et $\lambda \in]0,1[$. On pose $x_0 = (1-\lambda)x+\lambda y$ de sorte que $x < x_0 < y$. La croissance de τ_{x_0} donne $\tau_{x_0}(y)-\tau_{x_0}(x) \geqslant 0$ et le calcul fait au début de cette preuve montre que cela équivaut à $(1-\lambda)f(x)+\lambda f(y)-f(x_0) \geqslant 0$. Comme cela est trivialement vrai pour $\lambda \in \{0,1\}$, f est convexe.

Corollaire
La fonction f est convexe sur I, si et seulement si pour tous réels x, y, z dans I tels que $x < y < z$ on a

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

Démonstration. Si f est convexe, l'astuce est (encore) de remarquer que $\tau_a(b) = \tau_b(a)$ quels que soient $a \neq b$. Les inégalités demandées sont alors $\tau_x(y) \leqslant \tau_x(z) = \tau_z(x) \leqslant \tau_z(y)$, qui sont vraies d'après le théorème précédent. Inversement, si les inégalités des pentes sont toujours vraies, la première d'entre elles traduit la croissance de τ_x sur $\mathrm{I}\setminus\{x\}$, et ce quel que soit $x \in \mathrm{I}$. D'après le théorème précédent, f est convexe sur I.

Exercice 3 *Quelles sont les fonctions à la fois convexes et concaves ?*

Correction 3 *Une fonction affine est à la fois convexe et concave puisque sa dérivée seconde est nulle donc à la fois positive et négative. Réciproquement, si f est à la fois convexe et concave, elle n'est a priori pas dérivable donc ce n'est pas aussi facile. Soit I l'intervalle de définition de f, supposé de longueur non nulle, et soit $x_0 \in \mathrm{I}$. L'application τ_{x_0} est croissante (car f est convexe), mais aussi décroissante (car f est concave), donc constante. Si p désigne cette constante, on a donc $\frac{f(x)-f(x_0)}{x-x_0} = p$ pour tout $x \in \mathrm{I}\setminus\{x_0\}$, soit $f(x) = p(x-x_0)+f(x_0)$. Cette égalité est encore vraie pour $x = x_0$, donc f est affine.*

Exercice 4 *Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction convexe sur $\mathbb{R}$ tout entier. Si f est majorée, montrer que f est constante. Et si f est seulement convexe sur $\mathbb{R}_+$?*

Correction 4 *Si f n'était pas constante, il existerait $a < b$ dans $\mathbb{R}$ tels que $f(a) \neq f(b)$. Traitons le cas où $f(a) < f(b)$. Puisque τ_b est croissante sur $\mathbb{R}\backslash\{b\}$, on aurait, pour tout $x > b, \tau_b(x) \geqslant \tau_b(a)$, soit $f(x) \geqslant f(b) + (x-b)\frac{f(b)-f(a)}{b-a}$. Puisque $f(b) - f(a) > 0$ et $b - a > 0$, on aurait $\lim_{+\infty} f = +\infty$, ce qui contredit le caractère majorée de f. Si $f(a) > f(b)$, on obtiendrait la même absurdité avec la croissance de τ_a en regardant ce qui se passe sur $]-\infty, a[$. C'est faux sur $\mathbb{R}_+$: la fonction $x \mapsto \frac{1}{x+1}$ y est convexe et majorée par 1.*

2.2 Cas des fonctions dérivables

Théorème 6 (cas 1 fois dérivable) *Soit I un intervalle de longueur non nulle et $f : \mathrm{I} \to \mathbb{R}$ une fonction dérivable sur I. Les assertions suivantes sont équivalentes.*
(i) La fonction f est convexe sur I.
(ii) La fonction f' est croissante sur I.

Démonstration. Supposons f convexe. Si $x < z$ sont dans I. Pour tout $y \in]x, z[$, l'inégalité des pentes s'écrit $\frac{f(y)-f(x)}{y-x} \leqslant \frac{f(z)-f(x)}{z-x} \leqslant \frac{f(z)-f(y)}{z-y}$.

- En faisant $y \to x^+$on obtient $f'_d(x) \leqslant \frac{f(z)-f(x)}{z-x}$ (dérivée à droite).
- En faisant $y \to z^-$on obtient $\frac{f(z)-f(x)}{z-x} \leqslant f'_g(z)$ (dérivée à gauche).

Puisque f est dérivable, $f'_d(x) = f'(x)$ et $f'_g(z) = f'(z)$ si bien que $f'(x) \leqslant f'(z)$ et la croissance de f' est prouvée.
Inversement, si f' est croissante, considérons $a < b$ dans I et étudions la fonction D mesurant l'écart entre la corde et la fonction : $\mathrm{D}(x) = \frac{f(b)-f(a)}{b-a}(x-a) + f(a) - f(x)$. La fonction D est alors dérivable sur $[a, b]$ et $\mathrm{D}'(x) = \frac{f(b)-f(a)}{b-a} - f'(x)$.
Le théorème des accroissements finis assure l'existence de $c \in]a, b[$ tel que $\frac{f(b)-f(a)}{b-a} = f'(c)$ si bien que $\mathrm{D}'(x) = f'(c) - f'(x)$. Puisque f' est croissante, on en déduit que

- $\forall x \in [a, c], \mathrm{D}'(x) \geqslant 0$, donc D est croissante sur $[a, c]$
- $\forall x \in [c, b], \mathrm{D}'(x) \leqslant 0$, donc D est décroissante sur $[c, b]$. Comme enfin $\mathrm{D}(a) = \mathrm{D}(b) = 0$, on en tire que D est positive sur $[a, b]$. Cela traduit le fait que la corde est au-dessus de la courbe de f, c'est-à-dire que f est convexe.

Corollaire (position des tangentes)
Soit I un intervalle de longueur non nulle et $f : \mathrm{I} \to \mathbb{R}$ une fonction dérivable sur I. Alors f est convexe si et seulement si

$$\forall x \in \mathrm{I}, \quad f(x) \geqslant f'(a) \cdot (x-a) + f(a)$$

autrement dit quand la courbe de f est au-dessus de chacune de ses tangentes

Démonstration. Une très mauvaise idée serait de partir de l'inégalité $f(x) \leqslant \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$ et de faire tendre vers b vers a : ce serait oublier que cette inégalité n'est valable que lorsque $x \in [a, b]$: après la limite on n'obtiendrait seulement $f(a) = f(a)$: super!
Restons sérieux : étudions $\mathrm{D}(x) = f(x) - [f'(a)(x-a) + f(a)]$ dont la dérivée est donnée par $\mathrm{D}'(x) = f'(x) - f'(a)$. Puisque f' est croissante, $\mathrm{D}'(x) \geqslant 0$ quand $x \geqslant a$ et $\mathrm{D}'(x) \leqslant 0$ quand $x \leqslant a$: les variations de D s'en déduisent et montrent que D admet un minimum global en a. Ajouté au fait que $\mathrm{D}(a) = 0$, nous en déduisons que D est toujours positive sur I.

Corollaire (cas 2 fois dérivable)
Soit I un intervalle de longueur non nulle et $f : \mathrm{I} \to \mathbb{R}$ une fonction deux fois dérivable sur I. Les assertions suivantes sont équivalentes.
(i) La fonction f est convexe sur I.
(ii) La fonction f'' est positive sur I.

2.3 Exemples de référence

- $\mathbb{R} \to \mathbb{R}, x \mapsto e^x$ est convexe : $\forall x \in \mathbb{R}, 1+x \leqslant e^x$
- Le sinus est concave sur $\left[0, \frac{\pi}{2}\right]$: $\forall x \in \left[0, \frac{\pi}{2}\right], \frac{2}{\pi} \leqslant \sin(x) \leqslant x$
- $]-1,+\infty[\to \mathbb{R}, x \mapsto \ln(1+x)$ est concave : $\forall x > -1, \ln(1+x) \leqslant x$
- $[-1,+\infty[, x \mapsto \sqrt{1+x}$ est concave : $\forall x \geqslant -1, \sqrt{1+x} \leqslant 1 + \frac{1}{2}x$
- Notons aussi que pour tout réel $\mathbb{R}^{+*} \to \mathbb{R}, \alpha, x \mapsto x^\alpha$ est
  - convexe sur $\mathbb{R}_+^*$ si et seulement si $\alpha \geqslant 1$ ou $\alpha \leqslant 0$
  - concave sur $\mathbb{R}_+^*$ si et seulement si $\alpha \in [0,1]$.

3 Petit supplément étoilé

Théorème 7 *Si* I *est un intervalle de longueur non nulle et si $f : \mathrm{I} \to \mathbb{R}$ est convexe sur* I*, elle admet en chaque point x_0 de* I *une dérivée à gauche et à droite et pour tous réels $a < b$ dans* I*,*

$$f_g'(a) \leqslant f_d'(a) \leqslant \frac{f(b)-f(a)}{b-a} \leqslant f_g'(b) \leqslant f_d'(b).$$

En conséquence,

1. *Les fonctions f_d' et f_g' sont croissantes sur* I.
2. *La fonction f est continue sur $\mathring{\mathrm{I}}$ (mais pas forcément sur* I *).*

Démonstration. Soit $x_0 \in \mathrm{I}$. On sait que τ_{x_0} est croissante sur $\mathrm{I}\setminus\{x_0\}$. Si en plus x_0 est intérieur à I, on dispose d'un $\varepsilon > 0$ tel que $]x_0-\varepsilon, x_0+\varepsilon[\subset \mathrm{I}$. On a donc,

$$\forall z \in]x_0-\varepsilon, x_0+\varepsilon\left[\setminus\{x_0\}, \quad \tau_{x_0}(x_0-\varepsilon) \leqslant \tau_{x_0}(z) \leqslant \tau_{x_0}(x_0+\varepsilon).\right.$$

Puisque τ_{x_0} est croissante et bornée sur $]x_0-\varepsilon, x_0+\varepsilon[\setminus\{x_0\}$, le théorème de la limite monotone assure alors l'existence et la finitude des limites à gauche et à droite en x_0 de τ_{x_0}, celles-ci valant respectivement $\sup_{z\in]x_0-\varepsilon,x_0[}\tau_{x_0}(z) = f_g'(x_0)$ et $\inf_{z\in]x_0,x_0+\varepsilon[}\tau_{x_0}(z) = f_d'(x_0)$.
De plus, si $z \in]x_0-\varepsilon, x_0[$ et $z' \in]x_0, x_0+\varepsilon\left[, \tau_{x_0}(z) \leqslant \tau_{x_0}(z')\right.$ par croissance de τ_{x_0}. En faisant $z \to x_0^-$, on obtient $f_g'(x_0) \leqslant \tau_{x_0}(z')$, puis en faisant $z' \to x_0^+$, on trouve finalement $f_g'(x_0) \leqslant f_d'(x_0)$ comme attendu. Si maintenant $a < b$ dans $\dot{o}$, alors, par définition des bornes inf et sup,

$$\inf_{z>a} \tau_a(z) \leqslant \tau_a(b) = \tau_b(a) \leqslant \sup_{z<b} \tau_b(z)$$

c'est-à-dire $f_d'(a) \leqslant \frac{f(b)-f(a)}{b-a} \leqslant f_g'(b)$. À ce stade la quadruple inégalité annoncée est prouvée.
Le fait que f_g' et f_d' soient croissantes en résulte immédiatement $\left(f_d'(a) \leqslant f_d'(b)\right.$ et $f_g'(a) \leqslant f_g'(b)$ quels que soient $\left.a<b\right)$. Enfin, puisque f est dérivable à droite en x_0, elle est a fortiori continue à droite en x_0, c'est-à-dire $\lim_{x\to x_0^+} f(x) = f(x_0)$. De même à gauche, si bien que f admet la limite $f(x_0)$ en x_0 (tout court) : f est continue en x_0.

⚠ **Attention**
Une fonction convexe peut être non continue au bord de I*. Penser par exemple à la fonction constante égale à 0 sur]0,1[et égale à 1 en 0 et en 1 : elle est convexe sur $[0,1]$, mais non continue en 0 et 1.*

Exercice 5 *Soit $f : \mathrm{I} \to \mathbb{R}$ une fonction définie sur un intervalle ouvert. Montrer que f est convexe sur* I *si et seulement si*

$$\forall (x,y) \in \mathrm{I}^2, \quad f\left(\frac{1}{2}x + \frac{1}{2}y\right) \leqslant \frac{1}{2}f(x) + \frac{1}{2}f(y).$$

Correction 5 *Une implication est évidente. Supposons donc que $f\left(\frac{x+y}{2}\right) \leqslant \frac{f(x)+f(y)}{2}$ pour tous $x, y \in \mathrm{I}$. Fixons x et y dans* I *et posons $m = \frac{x+y}{2}$, qui appartient à* I*, car* I *est un intervalle, donc une partie convexe de $\mathbb{R}$. On peut appliquer l'hypothèse à x et m, ainsi qu'à m et y et écrire*

- $f\left(\frac{x+m}{2}\right) \leqslant \frac{f(x)+f(m)}{2}$ *c'est-à-dire* $f\left(\frac{3}{4}x + \frac{1}{4}y\right) \leqslant \frac{3}{4}f(x) + \frac{1}{4}f(y)$.

— $f\left(\frac{m+y}{2}\right) \leqslant \frac{f(m)+f(y)}{2}$ *c'est-à-dire* $f\left(\frac{1}{4}x+\frac{1}{4}y\right) \leqslant \frac{1}{4}f(x)+\frac{3}{4}f(y)$.

On conçoit bien que ce procédé dichotomique peut se reproduire autant de fois que l'on souhaite : on peut montrer par récurrence que pour tout $n \in \mathbb{N}$ *et tout* $p \in [\![0, 2^n]\!]$, *on a* $f(\lambda_n x + (1-\lambda_n)y) \leqslant \lambda_n f(x) + (1-\lambda_n) f(y)$, *où* $\lambda_n = \frac{p}{2^n}$. *Faisons-le ! Le cas* $n = 0$ *est trivial car* λ_0 *ne peut valoir que 0 ou 1 . Soit* $n \in \mathbb{N}$. *Supposons notre hypothèse vraie au rang n. Soit alors* $p \in [\![0, 2^{n+1}]\!]$ *et* $\lambda = \frac{p}{2^{n+1}}$. *Distinguons deux cas.*

— *Si p est pair, alors* $p = 2k$ *avec k dans* $[\![0, 2^n]\!]$ *de sorte que* $\lambda = \frac{k}{2^n}$ *et l'HR assure que* $f(\lambda x+(1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$.

— *Si p est impair, alors* $p = 2k+1$ *avec* $k \in [\![0, 2^n-1]\!]$. *On peut alors écrire, puisque* $\lambda = \frac{1}{2}\left(\frac{k}{2^n}+\frac{k+1}{2^n}\right)$, $\lambda x + (1-\lambda)y$ *vaut*

$$\frac{1}{2}\left[\frac{k}{2^n}x+\frac{k+1}{2^n}x\right]+\frac{1}{2}\left[1+1-\frac{k}{2^n}y-\frac{k+1}{2^n}y\right]$$
$$=\frac{1}{2}\left[\frac{k}{2^n}x+\left(1-\frac{k}{2^n}y\right)\right]$$
$$+\frac{1}{2}\left[\frac{k+1}{2^n}x+\left(1-\frac{k+1}{2^n}y\right)\right].$$

L'hypothèse faite sur f (qui n'a rien à voir avec l'HR !) implique que $f(\lambda x + (1-\lambda)y)$ *est inférieur à*

$$\frac{1}{2}f\left(\frac{k}{2^n}x+\left(1-\frac{k}{2^n}\right)y\right)$$
$$+\frac{1}{2}f\left(\frac{k+1}{2^n}x+\left(1-\frac{k+1}{2^n}\right)y\right).$$

Mais comme k et $k+1$ *sont tous deux dans* $[\![0, 2^n]\!]$, *l'HR s'applique, ce qui nous permet de dire que* $f(\lambda x + (1-\lambda)y)$ *est inférieur à*

$$\frac{1}{2}\left[\frac{k}{2^n}f(x)+\left(1-\frac{k}{2^n}\right)f(y)\right]$$
$$+\frac{1}{2}\left[\frac{k+1}{2^n}f(x)+\left(1-\frac{k+1}{2^n}\right)f(y)\right]$$

qui n'est autre que $\lambda f(x) + (1-\lambda)f(y)$.

Concluons : il se trouve que l'ensemble $\mathrm{B} = \left\{\frac{p}{2^n} \mid (p,n) \in \mathbb{N}^2, p \leqslant 2^n\right\}$ *est dense dans* $[0,1]$ *: c'est le principe de l'écriture décimale illimitée des réels en base 2. Comme f est continue sur l'intérieur de* I *qui est ouvert, elle est continue sur* I. *Ainsi, pour tout* $\lambda \in [0,1]$, $f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$. *Comme x et y sont quelconques dans* I, *f est bien convexe sur* I

Théorème 8 *Soit* I *est un intervalle de longueur non nulle. Si* $f : \mathrm{I} \to \mathbb{R}$ *est convexe sur* I, *alors f est dérivable sur* I *sauf peut-être sur une partie finie ou dénombrable.*

La notion de partie dénombrable sera abordée dans le chapitre sur le dénombrement. On y établit par exemple que $\mathbb{Q}$ est dénombrable.

Démonstration. Il suffit de montrer que l'ensemble des points de I où f n'est pas dérivable est fini ou dénombrable, car $\mathrm{I} \setminus \mathring{\mathrm{I}}$ ne comporte qu'au plus deux éléments.
Puisque f admet en chaque point $x \in \mathrm{I}$ une dérivée à gauche $f'_g(x)$ et une dérivée à droite $f'_d(x)$ telles que $f'_g(x) \leqslant f'_d(x)$, l'ensemble des points de non dérivabilité de f sur I est $\mathrm{D} = \left\{x \in \mathring{\mathrm{I}} \mid f'_g(x) < f'_d(x)\right\}$.
Soit $x \in \mathrm{D}$ un tel point. Puisque $\mathbb{Q}$ est dense dans $\mathbb{R}$, on peut trouver un rationnel $r(x)$ dans $]\, f'_g(x), f'_d(x)\, [$. Si $x < y$, le théorème 1 précédent a établi que $f'_d(x) \leqslant f'_g(y)$, si bien que $r(x) < f'_d(x) \leqslant f'_g(y) < r(y)$. L'application $x \mapsto r(x)$ est donc une injection de D dans $\mathbb{Q}$, ce qui prouve, puisque $\mathbb{Q}$ est dénombrable, que D est fini ou dénombrable.

4 Exercices

4.1 Énoncés

Exercice 6 *Soit a un réel et $f : [a, +\infty[\to \mathbb{R}$ une fonction convexe.*

1. *Justifier que $\ell = \lim_{x \to +\infty} \frac{f(x)}{x}$ existe dans $\overline{\mathbb{R}}$.*
2. *Si $\ell \in \mathbb{R}$, montrer que $m = \lim_{x \to +\infty}(f(x) - \ell x)$ existe dans $\overline{\mathbb{R}}$.*
3. *Donner des exemples où $\ell \notin \mathbb{R}$, où ($\ell \in \mathbb{R}$ et $m \notin \mathbb{R}$) et un exemple non affine où ($\ell \in \mathbb{R}$ et $m \in \mathbb{R}$).*

Exercice 7 *Soit a, b des réels positifs tels que $a+b=1$. Montrer que $\forall x, y \in \mathbb{R}_+, 1+x^a y^b \leqslant (1+x)^a(1+y)^b$.*

Exercice 8 *(inégalité arithmético-géométrique and Co). Soit $n \in \mathbb{N}^*$ et $a_1, \ldots, a_n$ des réels positifs. On appelle moyenne géométrique[3], moyenne arithmétique et moyenne quadratique de $a_1, \ldots, a_n$ les quantités suivantes*

$$\mathrm{G} = \sqrt[n]{a_1 \times \ldots \times a_n}, \quad \mathrm{A} = \frac{a_1 + \ldots + a_n}{n}, \quad \mathrm{Q} = \sqrt{\frac{a_1^2 + \ldots + a_n^2}{n}}.$$

De plus, si les $a_1, \ldots, a_n$ sont tous strictement positifs, on pose

$$\mathrm{H} = \frac{n}{\frac{1}{a_1} + \cdots + \frac{1}{a_n}}$$

Les quantités $\mathrm{H}, \mathrm{G}, \mathrm{A}, \mathrm{Q}$ s'appellent respectivement moyennes harmonique, géométrique, arithmétique et quadratique des réels $a_1, \ldots, a_n$.

1. *Montrer que $\mathrm{H} \leqslant \mathrm{G} \leqslant \mathrm{A} \leqslant \mathrm{Q}$.*
2. *On souhaite étudier les cas d'égalité. Pour ce faire, on introduit un raffinement de la notion de convexité. Nous dirons qu'une fonction $f : \mathrm{I} \to \mathbb{R}$(avec I un intervalle de longueur non nulle de $\mathbb{R}$) est strictement convexe sur I quand*

$$\forall x, y \in \mathrm{I}, \forall \lambda \in]0,1[, [x \neq y \Longrightarrow f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y)$$

   c.-à-d. quand, dans la définition d'une fonction convexe, le cas d'égalité se produit uniquement dans les cas triviaux $x = y$ ou $\lambda \in \{0,1\}$. On constate donc qu'une fonction strictement convexe est a fortiori convexe.

   (a) *(Inégalité de Jensen stricte). Si f est strictement convexe sur I, montrer que pour tout $n \geqslant 2$, tous $\lambda_1, \ldots, \lambda_n \in]0,1[$ de somme 1 et tous $x_1, \ldots, x_n \in \mathrm{I}$, si les x_i ne sont pas tous égaux, alors*

$$f\left(\sum_{i=1}^{n} \lambda_i x_i\right) < \sum_{i=1}^{n} \lambda_i f(x_i)$$

   (b) *Si f est dérivable, montrer que f est strictement convexe sur I si et seulement si f' est strictement croissante sur I.*

   (c) *Étudier les cas d'égalité des inégalités vues en 1.*

Exercice 9 1. *A l'aide de la convexité de l'exponentielle, établir*

$$\forall n \in \mathbb{N}^*, \forall (x_1, \ldots, x_n) \in \mathbb{R}_+^n, \forall (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}_{+*}^n, \sum_{i=1}^{n} \alpha_i = 1, \prod_{i=1}^{n} x_i^{\alpha_i} \leq \sum_{i=1}^{n} \alpha_i x_i$$

2. *Comment obtenir un parallélépipède rectangle d'aire minimale à volume donné (emballage le plus économique) ?*

Exercice 10 *(inégalités de Hölder et de Minkowski). Si $p \in]1, +\infty[$, on appelle réel conjugué de p l'unique réel $q > 0$ à vérifier $\frac{1}{p} + \frac{1}{q} = 1$. Ainsi, le réel 2 est son propre conjugué.*

1. *Montrer que* $\forall(x,y)\in(\mathbb{R}_+)^2, xy \leqslant \frac{x^p}{p}+\frac{y^q}{q}$ *(inégalité de Young). 2. En déduire l'inégalité de Hölder : pour tout* $n\in\mathbb{N}^*$ *et tous n-uplets* $(x_1,\ldots,x_n)$ *et* $(y_1,\ldots,y_n)$ *de réels,*

$$\sum_{k=1}^{n}|x_k y_k| \leqslant \left(\sum_{k=1}^{n}|x_k|^p\right)^{\frac{1}{p}}\left(\sum_{k=1}^{n}|y_k|^q\right)^{\frac{1}{q}}.$$

3. *Si* $x=(x_1,\ldots,x_n)\in\mathbb{R}^n$*, on pose* $\|x\|_p=\sqrt[p]{|x_1|^p+\ldots+|x_n|^p}$*. Déduire de l'inégalité de Hölder l'inégalité de Minkowski : pour tous* $x,y\in\mathbb{R}^n$*,*

$$\|x+y\|_p \leqslant \|x\|_p+\|y\|_p$$

Indication : $|a+b|^p=|a+b|\cdot|a+b|^{p-1}$ *et le conjugué de p est* $q=\frac{p}{p-1}$.

Exercice 11 *On dit qu'une fonction* $f:\mathrm{I}\to\mathbb{R}$ *définie sur un intervalle* I *est logarithmiquement convexe (ou log-convexe) sur* I *quand f prend ses valeurs dans* $\mathbb{R}_+^*$ *et quand* $\ln\circ f$ *est convexe sur* I.

1. *Montrer que si f est log-convexe sur* I*, alors elle est convexe sur* I.
2. *Démontrer que f est log-convexe si et seulement pour tout* $c>0$*, la fonction* $x\mapsto f(x)c^x$ *est convexe.*
3. *En déduire que la somme de deux fonctions log-convexes est log-convexe.*

4.2 Corrigés

Correction 6 1. *Puisque f est convexe sur* $\left[a,+\infty\right[$, $\tau_a : x\mapsto\frac{f(x)-f(a)}{x-a}$ *est croissante sur* $]a,+\infty[$ *: le théorème de la limite monotone garantit l'existence de sa limite* ℓ *en* $+\infty$*. Comme* $\tau_a(x)\underset{x\to+\infty}{\sim}\frac{f(x)-f(a)}{x}$ *et que* $\lim_{x\to+\infty}\frac{f(a)}{x}=0$*, l'existence de* $\lim_{x\to+\infty}\frac{f(x)}{x}$ *est prouvée.*

2. *Posons* $g(x)=f(x)-\ell x$*. Pour tous* $x_0>a$ *et* $x\neq x_0$, $\frac{g(x)-g(x_0)}{x-x_0}=\tau_{x_0}(x)-\ell$ *si bien que* $x\mapsto\frac{g(x)-g(x_0)}{x-x_0}$ *est croissante sur* $]a,+\infty[\setminus\{x_0\}$*. Or* $\frac{g(x)-g(x_0)}{x-x_0}$ *est équivalent, quand* $x\to+\infty$*, à* $\frac{g(x)-g(x_0)}{x}$ *et* $\lim_{x\to+\infty}\frac{g(x)-g(x_0)}{x}=\lim_{x\to+\infty}\frac{g(x)}{x}=0$ *par définition de* ℓ*. Ainsi,* $x\mapsto\frac{g(x)-g(x_0)}{x-x_0}\leqslant 0$ *sur* $]a,+\infty[$*. Évaluée en* y_0 *avec* $y_0>x_0$*, cette inégalité devient* $g(y)\leqslant g(x_0)$*, si bien que g est décroissante sur* $[a,+\infty[$*. Le TLM s'applique et garantit l'existence de la limite m de g en* $+\infty$*.*
3. *Les fonctions* $x\mapsto x^2$, $x\mapsto-\sqrt{x}$ *et* $x\mapsto\frac{1}{x}+x$ *répondent à la question.*

Correction 7 *C'est un extrait de planche de l'X, qui a souvent dérouté les candidats, car après avoir pensé à prendre le logarithme, il faut encore avoir une petite idée. L'inégalité* $1+x^ay^b\leqslant(1+x)^a(1+y)^b$ *est équivalente à* $\ln\left(1+x^ay^b\right)\leqslant a\ln(1+x)+b\ln(1+y)$*, et on est content car le second membre fleure bon la convexité (car* $a,b\geqslant 0$ *et* $a+b=1$ *). Pour le premier membre, on pense à écrire* $x^ay^b=\mathrm{e}^{a\ln(x)}\mathrm{e}^{b\ln(y)}=\mathrm{e}^{a\ln(x)+b\ln(y)}$ *si bien que ce que l'on cherche à montrer, en fait, c'est*

$$\begin{aligned}&\ln\left(1+\mathrm{e}^{a\ln(x)+b\ln(y)}\right)\\&\leqslant a\ln(1+x)+b\ln(1+y)\\&=a\ln\left(1+\mathrm{e}^{\ln(x)}\right)+b\ln\left(1+\mathrm{e}^{\ln(y)}\right).\end{aligned}$$

Maintenant c'est fini : on pense évidemment à introduire $f:x\mapsto\ln(1+\mathrm{e}^x)$ *qui dérivable sur* $\mathbb{R}$ *avec* $f'(x)=\frac{\mathrm{e}^x}{1+\mathrm{e}^x}=\frac{\mathrm{e}^x+1-1}{1+\mathrm{e}^x}=1-\frac{1}{1+\mathrm{e}^x}$ *ce qui montre que* f' *est croissante (vous ferez meilleure impression en remarquant que* f' *est croissante ainsi, sans la dériver). En conclusion f est convexe.*

Correction 8 1. *Montrons que* $\mathrm{G}\leqslant\mathrm{A}$*. Si l'un des réels* a_k *est nul, c'est évident puisque* $\mathrm{G}=0$*. Sinon il sont tous dans* $\mathbb{R}_+^*$ *et la concavité de ln assure (via l'inégalité de Jensen) que*

$$\begin{aligned}\ln\frac{a_1+\ldots+a_n}{n}&\geqslant\frac{\ln(a_1)+\ldots+\ln(a_n)}{n}\\&=\ln(a_1\ldots a_n)^{\frac{1}{n}}.\end{aligned}$$

Puisque exp est croissante, on obtient $A \geqslant G$.

La convexité de $x \mapsto x^2$ *donne quant à elle* $\left(\frac{1}{n}a_1 + \ldots + \frac{1}{n}\right)^2 \leqslant \frac{1}{n}a_1^2 + \ldots + \frac{1}{n}a_n^2$. *La croissance de* $\sqrt{\ }$ *donne alors* $A \leqslant Q$.

Enfin, si les a_i *sont tous non nuls,* $\frac{1}{H}$ *est la moyenne arithmétique de* $\frac{1}{a_1}, \ldots, \frac{1}{a_n}$ *On sait depuis peu que* $\frac{1}{H} \geqslant G'$ *où* G' *est la moyenne géométrique de ces mêmes réels. Or voilà,* $G' = \frac{1}{G}$, *donc* $H \leqslant G$ *par décroissance de* $x \mapsto \frac{1}{x}$ *sur* $\mathbb{R}_+^*$. *Finalement,* $H \leqslant G \leqslant A \leqslant Q$ *comme annoncé.*

2. (a) *On procède par récurrence, un peu dans l'idée de la preuve de la caractérisation des parties convexes par les barycentres. Si* $n = 2$, *c'est la définition même de la stricte convexité. Soit* $n \geqslant 2$. *Supposons acquis notre inégalité stricte pour n'importe quels choix de* λ_k *et de* x_k *comme dans l'énoncé. Soit maintenant* $\lambda_1, \ldots, \lambda_{n+1}$ *dans* $]0,1[$ *de somme 1 et* $x_1, \ldots, x_{n+1}$ *non tous égaux. Pour se ramener à une CLC de n termes, on pense par exemple à décrire la convexité de f par*

   $f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) = f((1-\lambda_{n+1})x + \lambda_{n+1}x_{n+1}) \leqslant (1-\lambda_{n+1})f(x) + \lambda_{n+1}f(x_{n+1})$

   où $x = \sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i$. *Posons donc, pour tout* $i \in [\![1,n]\!]$, $\mu_i = \frac{\lambda_i}{1-\lambda_{n+1}}$: *nous sommes devant une CLC puisque* $\sum_{i=1}^{n} \mu_i = \frac{1}{1-\lambda_{n+1}} \sum_{i=1}^{n} \lambda_i = \frac{1-\lambda_{n+1}}{1-\lambda_{n+1}} = 1$ *et que pour tout* $i \in [\![1,n+1]\!]$, $\mu_i > 0$. *Notons que* $\mu_i \neq 1$ *car sinon les autres* μ_k *seraient nuls, ce qui est exclu. Distinguons deux cas.*

   — *Cas 1 : les* $x_1, \ldots, x_n$ *ne sont pas tous égaux. L'HR nous permet alors de dire que* $f(x) < \sum_{i=1}^{n} \mu_i x_i$. *On a donc*
   $f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leqslant (1-\lambda_{n+1})f(x) + \lambda_{n+1}f(x_{n+1})$
   $< (1-\lambda_{n+1})\sum_{i=1}^{n} \mu_i x_i + \lambda_{n+1}f(x_{n+1})$
   qui n'est autre que l'inégalité stricte attendue puisque $(1-\lambda_{n+1})\sum_{i=1}^{n} \mu_i x_i = \sum_{i=1}^{n} \lambda_i x_i$.

   — *Cas 2 : les* $x_1, \ldots, x_n$ *sont tous égaux. Cela implique donc, vu ce qu'on a supposé sur* $x_1, \ldots, x_{n+1}$, *que* $x_1 \neq x_{n+1}$. *Cette fois on a* $\sum_{i=1}^{n} \lambda_i x_i = (\sum_{i=1}^{n} \lambda_i) x_1$ *donc* $f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) = f((1-\lambda_{n+1})x_1 + \lambda_{n+1}x_{n+1}) < (1-\lambda_{n+1})f(x_1) + \lambda_{n+1}f(x_{n+1})$

   Puisque $(1-\lambda_{n+1})f(x_1) = \sum_{i=1}^{n} \lambda_i f(x_i)$, *cette dernière inégalité peut encore s'écrire*

   $$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) < \sum_{i=1}^{n+1} \lambda_i f(x_i).$$

   Dans tous les cas on a prouvé l'hérédité, ce qui achève la récurrence.

   (b) *C'est la même démonstration que dans le cours en remplaçant les* $\leqslant$ *par des* $<$ *(revenir aux taux d'accroissement, prouver qu'ils sont tous strictement croissants, etc.)*

   (c) *D'après (b), ln et* $x \mapsto x^2$ *sont strictement convexes. D'après (a), vu que* $\frac{1}{n} \in \left]0,1\right[$, *les cas d'égalité se produisent si et seulement si* $a_1 = \ldots = a_n$.

Correction 9 1. *Si l'un des réels* x_i *est nul, alors le produit de gauche est nul et l'inégalité est trivialement vérifiée. On suppose à présent que tous les réels* x_i *sont strictement positifs. Alors, comme l'exponentielle est convexe, l'inégalité de Jensen discrète appliquée aux réels* $(\ln(x_i))_i$ *entraîne*

$$\exp\left(\sum_{i=1}^{n} \alpha_i \ln(x_i)\right) \leq \sum_{i=1}^{n} \alpha_i \exp(\ln(x_i))$$

On reconnaît alors

$$\prod_{i=1}^{n} x_i^{\alpha_i} \leq \sum_{i=1}^{n} \alpha_i x_i$$

2. *Notons* x, y, z *les dimensions strictement positives d'un parallélipède rectangle. Alors son volume* V *vaut* xyz *et sa surface vaut* $2(xy+yz+zx)$. *On applique alors l'inégalité précédente aux inverses des dimensions avec* $\alpha_1 = \alpha_2 = \alpha_3 = 1/3$, *ce qui entraîne*

$$\left(\frac{1}{xyz}\right)^{1/3} \leq \frac{1}{3}\left(\frac{1}{x} + \frac{1}{y} + \frac{1}{z}\right) = \frac{zy + zx + xy}{3xyz}$$

soit encore

$$V^{-1/3} \leq \frac{S}{6V}$$

On en déduit que $S \geq V^{2/3}$. *De plus, en utilisant les mêmes notions de convexité stricte que dans l'exercice précédent (l'exponentielle est strictement convexe et le logarithme est injectif), on constate qu'il y a égalité si et seulement si $x = y = z$. Ainsi, l'emballage le plus économique est cubique.*

Correction 10 1. *Soit $x, y > 0$ quelconques. Puisque $\frac{x^p}{p} + \frac{y^q}{q} = \frac{1}{p}e^{p\ln(x)} + \frac{1}{q}e^{q\ln(x)}$, que $p, q > 0$ avec $\frac{1}{p} + \frac{1}{q} = 1$ et que* exp *est convexe, on en déduit que $\frac{x^p}{p} + \frac{y^q}{q} \geqslant \exp\left(\frac{p\ln(x)}{p} + \frac{q\ln(y)}{q}\right) = xy$. Si x ou y est nul, l'inégalité reste vraie : elle est donc vraie pour tout $x, y \in \mathbb{R}_+$.*

Notons que si $p = 2$, cette inégalité est célèbre et simple à montrer : elle s'obtient en écrivant $(x-y)^2 \geqslant 0$ et en développant.

2. *Encore une fois, si l'un des réels mis en jeu est nul, l'inégalité est triviale. On suppose donc que tout le monde est dans* $\mathbb{R}^*$. Si *$i \in [\![1, n]\!]$, on applique l'inégalité précédente avec $x = \frac{|x_i|}{(\sum |x_k|^p)^{\frac{1}{p}}}$ et $y = \frac{|y_i|}{(\sum |y_k|^q)^{\frac{1}{q}}}$ pour obtenir*

$$\frac{|x_i y_i|}{(\sum |x_k|^p)^{\frac{1}{p}} (\sum |y_k|^q)^{\frac{1}{q}}} \leqslant \frac{1}{p}\frac{|x_i|^p}{\sum |x_k|^p} + \frac{1}{q}\frac{|y_i|^q}{\sum |y_k|^q}.$$

En sommant sur tous les i, on obtient l'inégalité attendue car $\frac{1}{p} + \frac{1}{q} = 1$. Si $p = 2$, on retrouve l'inégalité de Cauchy-Schwarz provenant du produit scalaire canonique de $\mathbb{R}^n$.

3. *Notons que l'inégalité demandée reste vraie si $p = 1$, mais pour pouvoir utiliser ce qui précède, nous supposons $p > 1$. Appliquons l'indication donnée :*

$$\begin{aligned}\left(\|x+y\|_p\right)^p &= \sum_{k=1}^{n} |x_k + y_k| \cdot |x_k + y_k|^{p-1} \\ &\leqslant \sum_{k=1}^{n} |x_k| \cdot |x_k + y_k|^{p-1} \\ &\quad + \sum_{k=1}^{n} |y_k| \cdot |x_k + y_k|^{p-1}.\end{aligned}$$

L'inégalité de Hölder fraîchement démontrée en question 2 appliquée aux deux dernières sommes ci-dessus montre que $\left(\|x+y\|_p\right)^p$ est inférieure à

$$\left[\sum_{k=1}^{n} |x_k|^p\right]^{\frac{1}{p}} \left[\sum_{k=1}^{n} |x_k + y_k|^{q(p-1)}\right]^{\frac{1}{q}} + \left[\sum_{k=1}^{n} |y_k|^p\right]^{\frac{1}{p}} \left[\sum_{k=1}^{n} |x_k + y_k|^{q(p-1)}\right]^{\frac{1}{q}}.$$

c'est gagné!

Correction 11 1. *Supposons f log-convexe. Soit $x, y \in$* I *et $t \in [0, 1]$. Par hypothèse,*
$\ln f(tx + (1-t)y) \leqslant t \ln f(x) + (1-t) \ln f(y)$

Par concavité de ln *et croissance de* exp*, f est convexe sur* I

2. *Supposons que f soit log-convexe. Si $c > 0$, l'application $x \mapsto \ln(f(x)) + x\ln(c)$ est convexe comme somme de deux fonctions convexes (la deuxième est même affine), autrement dit $x \mapsto \ln(f(x)c^x)$ est convexe. D'après 1, $x \mapsto f(x)c^x$ est convexe.*

Récip., supposons que $x \mapsto f(x)c^x$ soit convexe pour tout $c > 0$. Si $x \neq y$ dans I *et $t \in [0, 1]$, alors $f(tx+(1-t)y)c^{tx+(1-t)y} \leqslant tf(x)c^x + (1-t)f(y)c^y$, et ce pour tout $c > 0$. Après division par $c^{tx+(1-t)y} > 0$:*

$$\forall c > 0, \quad f(tx + (1-t)y) \leqslant tf(x)c^{(1-t)(x-y)} + (1-t)f(y)c^{t(y-x)}.$$

L'astuce est ici de choisir « le meilleur c possible ≫ : le membre de droite de l'inégalité précédente définit une fonction dérivable φ de la variable c avec

$$\varphi'(c) = t(1-t)(x-y)c^{(1-t)(x-y)-1} + (1-t)t(y-x)f(y)c^{t(y-x)-1} = t(1-t)(x-y)c^{(y-x)t-1}\left[f(x)c^{x-y} - f(y)\right].$$

Puisque φ est continue et à valeurs dans $\mathbb{R}_+$, elle admet un minimum global, et puisque φ' ne s'annule que pour $c_0 = \mathrm{e}^{-\frac{f(y)-f(x)}{y-x}}$ (c'està-dire quand ${c_0}^{x-y} = \frac{f(y)}{f(x)}$), ce minimum vaut $\varphi(c_0) = f(y)^{1-t}f(x)^t$. On peut ainsi dire que $f(tx + (1-t)y) \leqslant f(y)^{1-t}f(x)^t$ et en composant par ln qui est croissante, on obtient la convexité de $\ln \circ f$.

3. *Si f et g sont log-convexes, alors d'après 2) pour tout $c > 0, x \mapsto f(x)c^x$ et $x \mapsto g(x)c^x$ sont convexes, donc leur somme aussi, c'est-à-dire $x \mapsto (f(x)+g(x))c^x$ est convexe, et d'après 2) encore, $f+g$ est logconvexe.*