Convexité

Cornou Jean-Louis

4 janvier 2024

L'étude de la convexité provient tout d'abord de la notion naturelle de barycentre, chère à la mécanique classique. Elle permet d'étudier certains objets géométriques tels les polygônes comme intersection de demi-plans, les polyèdres comme intersections de demi-espaces. Ces notions géométriques ont des prolongements analytiques très riches, amenant à reformuler des problèmes sous forme de recherche d'extrema (on parle d'optimisation). Les conséquences en analyse numérique, théorie des graphes, économie mathématique, problème de frottement, sont légion.

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 $0_{\mathbb{R}^2} = 0_{\mathrm{E}} = (0, 0)$.

Définition 1 *On appelle système (fini) de points massiques de* E *toute famille finie* $((\mathrm{A}_i, \alpha_i))_{i\in \mathrm{I}}$ *de* $\mathrm{E} \times \mathbb{R}$ *indexée par un ensemble* I *non vide.*

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 $\mathrm{A}, \mathrm{B} \in \mathrm{E}$ *on note* $\overrightarrow{\mathrm{AB}}$ *le vecteur* $\mathrm{B} - \mathrm{A}$. *On remarque que* $\overrightarrow{\mathrm{AA}} = 0_{\mathrm{E}}$ *et* $\overrightarrow{\mathrm{AB}} + \overrightarrow{\mathrm{BC}} = \overrightarrow{\mathrm{AC}}$ *quels que soient* $\mathrm{A}, \mathrm{B}, \mathrm{C} \in \mathrm{E}$, *ainsi que* $\mathrm{A} + \overrightarrow{\mathrm{AB}} = \mathrm{B}$.

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

- $\exists! \mathrm{G} \in \mathrm{E}, \sum_{i\in \mathrm{I}} \alpha_i \overrightarrow{\mathrm{A}_i\mathrm{G}} = 0_{\mathrm{E}}$.
- $\sum_{i\in \mathrm{I}} \alpha_i \neq 0$

Le cas échéant, $\mathrm{G} = \frac{1}{\sum_{i\in \mathrm{I}} \alpha_i} \sum_{i\in \mathrm{I}} \alpha_i \mathrm{A}_i$. *On dit que* G *est le barycentre du système massique* $((\mathrm{A}_i, \alpha_i))_{i\in \mathrm{I}}$.

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

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

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 $\sum \alpha_i \neq 0$, on dispose de $i_0 \in I$ car I non vide, prenons $M' = A_{i_0}$. 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

$$\begin{aligned} M &= A_{i_0} - \frac{1}{\sum_{i\in I}\alpha_i}\Lambda\left(A_{i_0}\right) = \frac{1}{\sum_{i\in I}\alpha_i}\sum_{i\in I}\alpha_i A_{i_0} - \frac{1}{\sum_{i\in I}\alpha_i}\sum_{i\in I}\alpha_i\overrightarrow{A_iA_{i_0}} \\ &= \frac{1}{\sum_{i\in I}\alpha_i}\sum_{i\in I}\alpha_i\left(A_{i_0} - \overrightarrow{A_iA_{i_0}}\right) = \frac{1}{\sum_{i\in I}\alpha_i}\sum_{i\in I}\alpha_i A_i, \end{aligned}$$

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* $((A_i,\alpha_i))_{i\in I}$ *a pour coordonnées*

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

Définition 2 *Quand toutes les masses sont identiques, et non nulles, on parle d'isobarycentre. Ainsi, l'isobarycentre des points* $A_1,\ldots,A_n$ *est le point* G *défini* par $G = \frac{1}{n}\sum_{i=1}^n A_i$

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

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

$$\frac{1}{\sum_{i\in I}k\alpha_i}\sum_{i\in I}k\alpha_i A_i = \frac{1}{\sum_{i\in I}\alpha_i}\sum_{i\in I}\alpha_i A_i.$$

1.2 Parties convexes

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

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

Notation
Ce segment est noté $[A,B]$.

⚠ **Attention**
Cette notation est en conflit avec la notion des intervalles de $\mathbb{R}$*. Dans le cas de* $\mathbb{R}^2$, $[A,B] = [B,A]$ *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*

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

☞ **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*

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

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 ?*

1.3 Caractérisation des parties convexes

Exemple 2 *Soit* A, B *deux points de* $\mathbb{R}^2$*. L'ensemble des barycentres que l'on peut produire à l'aide de* A *et* B *est la droite* (AB)*. En effet, ces barycentres sont de la forme* $\frac{1}{a+b}(a\mathrm{A}+b\mathrm{B})$ *avec* $(a,b)\in\mathbb{R}^2$ *tel que* $a+b\neq 0$*. En posant* $\lambda=\frac{b}{a+b}$*, on obtient l'expression* $\mathrm{G}_\lambda=(1-\lambda)\mathrm{A}+\lambda\mathrm{B}=\mathrm{A}+\lambda\overrightarrow{\mathrm{AB}}$*. Quand* a,b *sont quelconques,* λ *parcourt* $\mathbb{R}$ *et* G_λ *décrit la droite passant* A *dirigée par* $\overrightarrow{\mathrm{AB}}$*, 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. En effet, si* a *et* b *sont tous deux positifs,* $0\leq b\leq a+b$*, donc* $0\leq\lambda\leq 1$ *puisque* $a+b>0$*. Si* a *et* b *sont tous deux négatifs, on a* $a+b\leq b\leq 0$*, ce qui entraîne* $0\leq\lambda\leq 1$*, puisque* $a+b<0$*. Réciproquement, si* $\lambda\in[0,1]$*,* b *et* $a+b$ *sont de même signe. Dans le cas* $b\leq 0$*, alors* $a+b>0$*. Comme* $\lambda\leq 1$*, on obtient* $b\leq a+b$ *soit* $a\geq 0$*. Dans le cas* $b\leq 0$*, alors* $a+b<0$*, comme* $\lambda\leq 1$*, on obtient* $a+b\leq b$ *soit* $a\leq 0$*.*

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* $(\mathrm{ABC})=\mathbb{R}^2$*. En effet, ceux-ci sont tous de la forme* $\frac{1}{a+b+c}(a\mathrm{A}+b\mathrm{B}+c\mathrm{C})$ *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* $(1-\lambda-\mu)\mathrm{A}+\lambda\mathrm{B}+\mu\mathrm{C}$ *soit encore* $\mathrm{A}+\lambda\overrightarrow{\mathrm{AB}}+\mu\overrightarrow{\mathrm{AC}}$ *: quand* a,b,c *décrivent* $\mathbb{R}$*,* (λ,μ) *décrit* $\mathbb{R}^2$ *si bien que* $\mathrm{A}+\lambda\overrightarrow{\mathrm{AB}}+\mu\overrightarrow{\mathrm{AC}}$ *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 $\sum_{i\in\mathrm{I}}\lambda_i\mathrm{A}_i$ avec $\sum_{i\in\mathrm{I}}\lambda_i=1$: $\lambda_1=\frac{a}{a+b},\lambda_2=\frac{b}{a+b}$ 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* $\mathrm{A}_1,\dots,\mathrm{A}_n$ *des points de* $\mathrm{E}=\mathbb{R}^2$*. On appelle combinaison linéaire convexe (abrégé en CLC) de* $\mathrm{A}_1,\dots,\mathrm{A}_n$ *tout élément de* E *de la forme* $\sum_{i=1}^n\lambda_i\mathrm{A}_i$ *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* $((\mathrm{A}_i,\alpha_i))_{i\in\mathrm{I}}$ *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 combinaison linéaire convexe de deux points A_1 et A_2 est encore dans C, puis l'ensemble de ces combinaisons linéaires convexes n'est autre que [AB] (cf. exemple ci-dessus). Soit $n\geq 2$. Supposons que toute combinaison linéaire convexe de n points de C soit encore dans C et considérons $\mathrm{A}_1,\dots,\mathrm{A}_{n+1}$ dans C. Une combinaison linéaire convexe de ces $n+1$ points s'écrit $\mathrm{M}=\sum_{i=1}^{n+1}\lambda_i\mathrm{A}_i$ avec et $\lambda_1,\dots,\lambda_{n+1}\in\mathbb{R}_+$de somme 1 . Comme $n\geq 2$, les réels λ_i ne peuvent 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

$$\mathrm{M}=\sum_{i=1}^{n+1}\lambda_i\mathrm{A}_i=m\sum_{\substack{i=1\\i\neq i_0}}^{n+1}\frac{\lambda_i}{m}\mathrm{A}_i+\lambda_{i_0}\mathrm{A}_{i_0}=m\mathrm{B}+(1-m)\mathrm{A}_{i_0},$$

où $\mathrm{B}=\sum_{\substack{i=1\\i\neq i_0}}^{n+1}\frac{\lambda_i}{m}\mathrm{A}_i\in\mathrm{C}$ par hypothèse de réucurrence. Ainsi, M est combinaison linéaire convexe de B et A_{i_0} qui sont deux points de C, donc $\mathrm{M}\in\mathrm{C}$ et la récurrence s'achève.

2 ⇒ 3 Tout barycentre à masses positives non identiquement nulles est une combinaison linéaire convexe par homogénéité

3 ⇒ 1 Soit $(\mathrm{A},\mathrm{B})\in\mathrm{C}^2$. Tout point de [A, B] est un barycentre à coefficients positifs de A et B, c'est donc un donc un point de C. Ainsi, $[\mathrm{A},\mathrm{B}]\subset\mathrm{C}$.

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 lorsque*

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

On dit que f *est concave quand* $-f$ *est convexe.*

☞ **Remarque**

Notez que l'ensemble de définition de f *doit être une partie convexe pour que le réel* $(1-\lambda)x + \lambda y$ *soit toujours un point de définition de* f*. D'après la caractérisation des convexes de* $\mathbb{R}$*, il faut recourir à des intervalles.*

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

Exemple 4 *Toute fonction affine est convexe et concave. Soit* $(a,b) \in \mathbb{R}^2$ *et* $f : \mathbb{R} \to \mathbb{R}, x \mapsto ax+b$*. Soit* $(x,y) \in \mathbb{R}^2$*,* $\lambda \in [0,1]$*.*

$$f((1-\lambda)x+\lambda y) = a[(1-\lambda)x+\lambda y]+b = a(1-\lambda)+\lambda y+(1-\lambda)b+\lambda b = (1-\lambda)(ax+b)+\lambda(ay+b) = (1-\lambda)f(x)+\lambda f(y)$$

Il y a égalité donc double inégalité.

Exemple 5 *Notons* $g : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$*. Montrons que* g *est convexe. Soit* $(x,y) \in \mathbb{R}^2$*,* $\lambda \in [0,1]$*.*

$$\begin{aligned}(1-\lambda)g(x) + \lambda g(y) - g((1-\lambda)x+\lambda y) &= (1-\lambda)x^2 + \lambda y^2 - ((1-\lambda)x + \lambda y)^2 \\ &= (1-\lambda)x^2 - (1-\lambda)^2x^2 + \lambda y^2 - \lambda^2 y^2 - 2\lambda(1-\lambda)xy \\ &= \lambda(1-\lambda)\left(x^2+y^2-2xy\right) \\ &= \lambda(1-\lambda)(x-y)^2 \\ &\geq 0\end{aligned}$$

Exemple 6 *Notons* $h : \mathbb{R}_+ \to \mathbb{R}, x \mapsto \sqrt{x}$ *la fonction racine carrée. Montrons que* h *est concave. Soit* $(x,y) \in \mathbb{R}_+^2, \lambda \in [0,1]$*. On applique l'inégalité de convexité de la fonction carrée aux réels* $\sqrt{x}$ *et* $\sqrt{y}$ *et au même réel* $\lambda \in [0,1]$*. Cela entraîne*

$$\left((1-\lambda)\sqrt{x} + \lambda\sqrt{y}\right)^2 \leq (1-\lambda)\sqrt{x}^2 + \lambda\sqrt{y}^2 = (1-\lambda)x + \lambda y$$

Mais alors, comme la racine carrée est croissante, on obtient

$$\left|(1-\lambda)\sqrt{x} + \lambda\sqrt{y}\right| \leq \sqrt{(1-\lambda)x + \lambda y}$$

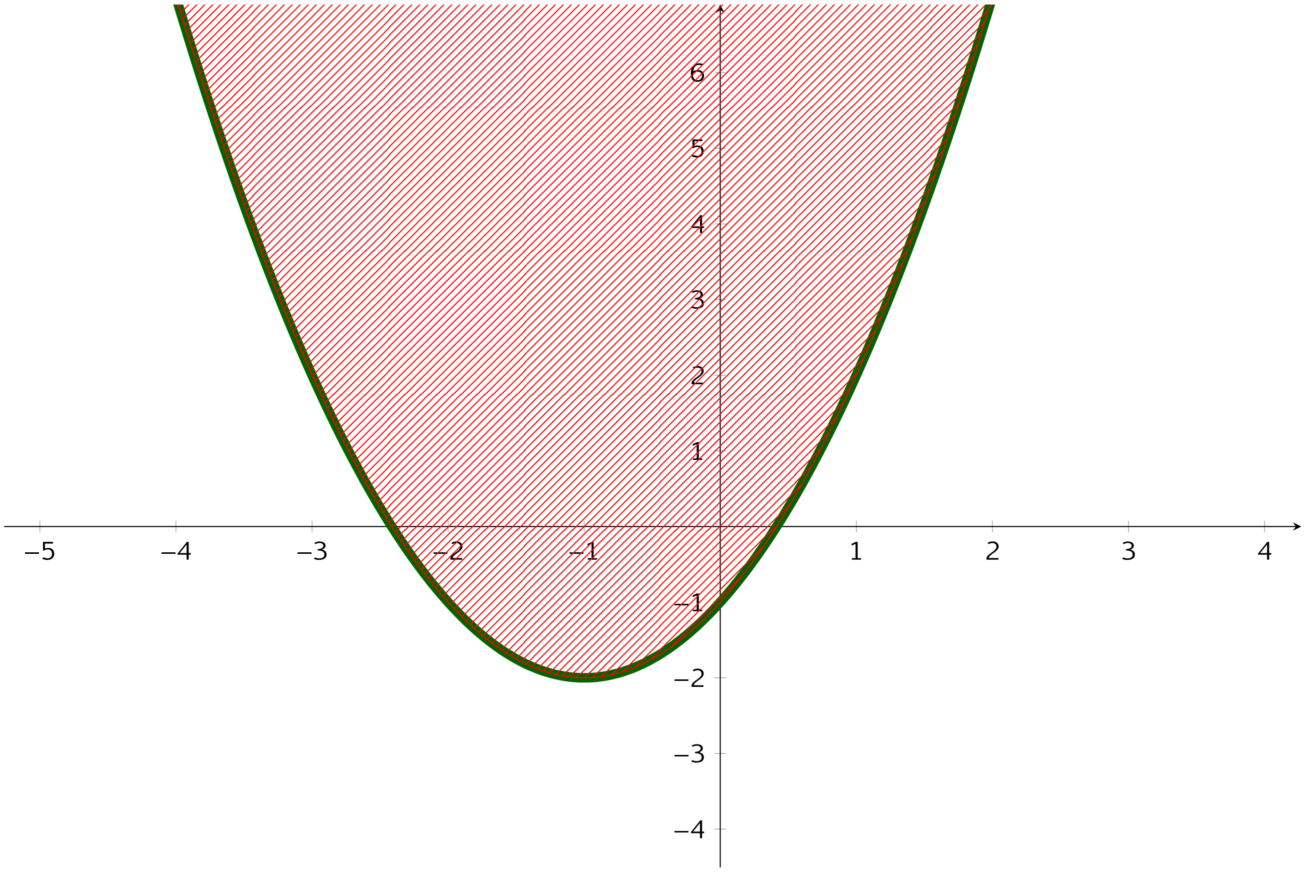
Comme $(1-\lambda)x + \lambda y \geq 0$*, on obtient*

$$(1-\lambda)\sqrt{x} + \lambda\sqrt{y} \leq \sqrt{(1-\lambda)x + \lambda y}$$

Ainsi, h *est bien concave.*

Définition 7 *Si* $f : I \to \mathbb{R}$ *est une fonction définie sur un intervalle* I*, on appelle épigraphe de* f *l'ensemble des couples* (x,y) *de* $I \times \mathbb{R}$ *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. On procède par double implication

- Supposons f est convexe sur I et considérons deux points $\mathrm{A} = (x_\mathrm{A}, y_\mathrm{A})$ et $\mathrm{B} = (x_\mathrm{B}, y_\mathrm{B})$ de $\mathrm{Epi}(f)$. Soit M un point de $[\mathrm{A}, \mathrm{B}]$. Il exsite $\lambda \in [0,1]$ tel que $\mathrm{M} = (1-\lambda)\mathrm{A} + \lambda\mathrm{B}$: 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 $f(x_\mathrm{B}) \leqslant y_\mathrm{B}$. Comme $\lambda \geq 0$ et $1-\lambda \geq 0$, on a $\lambda f(x_\mathrm{B}) \leq \lambda y_\mathrm{B}$ et $(1-\lambda)f(x_\mathrm{A}) \leq (1-\lambda)y_\mathrm{A}$. En sommant, il vient alors $f(x_\mathrm{M}) \leqslant y_\mathrm{M}$ ce qui prouve que C est dans $\mathrm{Epi}(f)$. Ainsi, l'épigraphe de f contient tous les points de $[\mathrm{AB}]$, il est donc convexe.
- Réciproquement, supposons $\mathrm{Epi}(f)$ 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 est 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 combinaison linéaire convexe d'éléments de I, ainsi $f\left(\sum_{i=1}^n \lambda_i x_i\right)$ 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 combinaison linéaire convexe d'après la caractérisation des parties convexes (paragraphe précédent). Or, pour tout i dans $[\![1, n]\!]$, $\mathrm{A}_i = (x_i, f(x_i))$ est dans Epi (f), donc $\mathrm{M} = \sum_{i=1}^n \lambda_i \mathrm{A}_i$ aussi. L'abscisse de M étant $\sum_{i=1}^n \lambda_i x_i$ et son ordonnée étant $\sum_{i=1}^n \lambda_i f(x_i)$, dire que $\mathrm{M} \in \mathrm{Epi}(f)$ équivaut à l'inégalité souhaitée.

☞ **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.

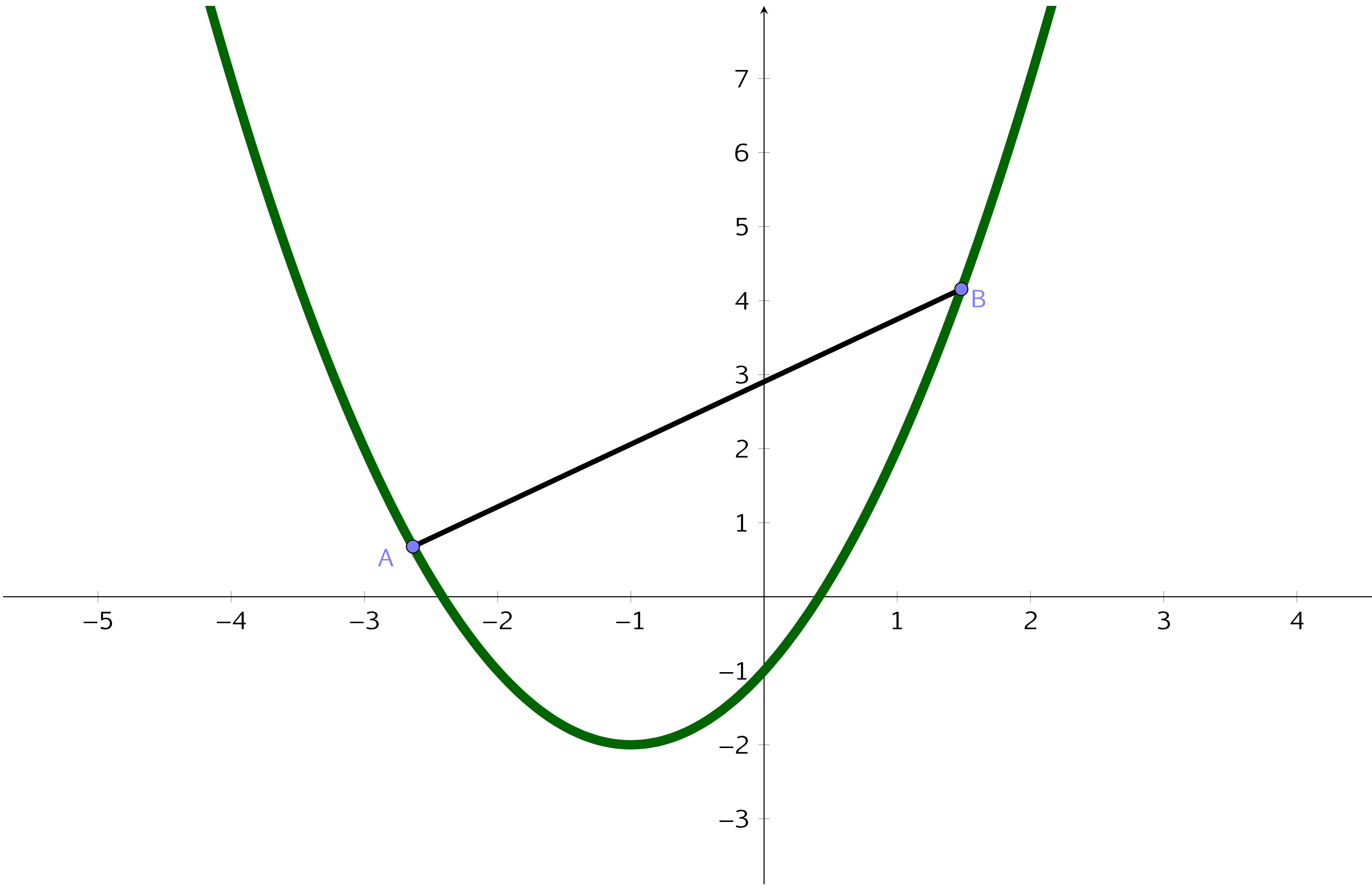
Corollaire (position des cordes)
Soit I *un intervalle de longueur non nulle et* $f : \mathrm{I} \to \mathbb{R}$ *une fonction. Alors* f *est convexe ssi pour tous réels* $a < b$ *dans* I*,*

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

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

☞ **Remarque**
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. L'équation de la droite* (AB) *est* $y = f(a)+(x-a)\frac{f(b)-f(a)}{b-a}$*. Le segment* [A, B] *s'écrit également* $\{(x, f(a)+(x-a)\frac{f(b)-f(a)}{b-a})|x \in [a,b]\}$.

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 A = $(a, f(a))$ et 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.*

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 $I \setminus \{x_0\}$

Démonstration. Supposons f convexe sur I et considérons $x_0 \in I$. Faisons une première constatation : soit $a < b < c$ dans I, alors en posant $\lambda = \frac{b-a}{c-a} \in]0,1[$, $b = (1-\lambda)a + \lambda c$ et $b - a = \lambda(c-a)$, on a

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

Comme $b-a > 0$ et $f(b) = f((1-\lambda)a+\lambda c) \leqslant (1-\lambda)f(a)+\lambda f(c)$ par convexité de f, on en déduit l'inégalité $(*)$: $\tau_a(b) \leqslant \tau_a(c)$. D'autre part, en posant $\mu = 1-\lambda \in]0,1[$, 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 $f(b) = f(\mu a + (1-\mu)c) \leqslant (1-\mu)f(c)+\mu f(a)$ par convexité de f, on obtient $(**)$: $\tau_c(a) \leqslant \tau_c(b)$.
Maintenant, soit $x < y$ dans $I \setminus \{x_0\}$. 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 $\tau_y(x) \leqslant \tau_y(x_0)$. Or, $\tau_x(y) = \tau_y(x)$ donc par transitivité, $\tau_x(x_0) \leqslant \tau_y(x_0)$ et d'après $\tau_\star(\square) = \tau_\square(\star)$, 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 $I \setminus \{x_0\}$ quel que soit $x_0 \in I$. 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 l'inégalité de convexité est trivialement vraie pour $\lambda = 0$ et $\lambda = 1$, f est convexe.

Corollaire (Inégalités des pentes)

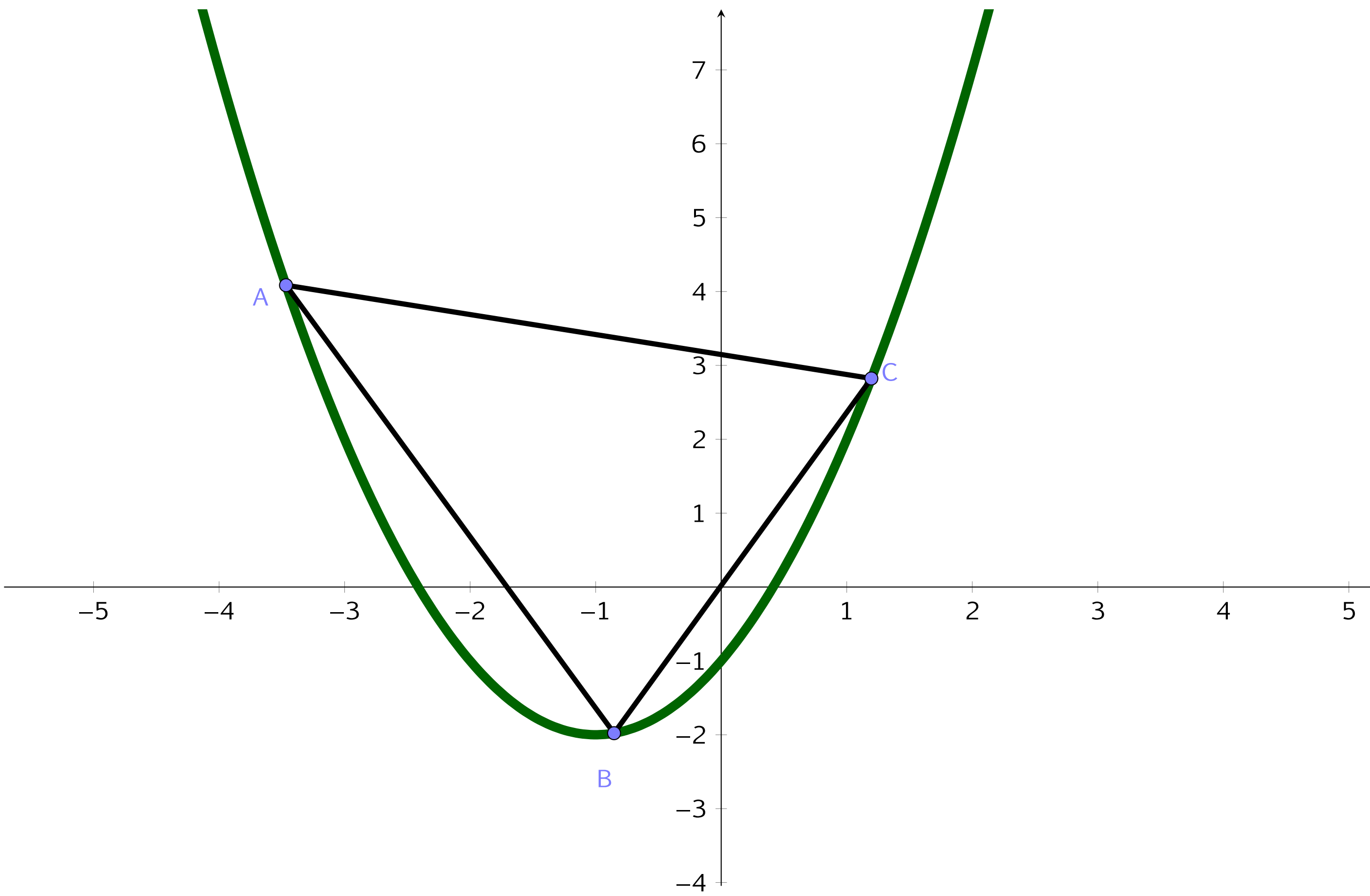
La fonction f est convexe sur I*, si et seulement si pour tous réels x, y, z dans* I *tels que $x < y < z$ on a*

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

Démonstration. Si f est convexe,on remarque toujours 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. Soit $x_0 \in I$. Soit $x < y$ dans $I \setminus \{x_0\}$. Distinguons trois cas :

- Si $x_0 < x < y$, alors la première inégalité des pentes donne $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$.
- Si $x < y < x_0$, alors la deuxième inégalité des pentes donne $\tau_{x_0}(x) \leqslant \tau_{x_0}(y)$.
- Si $x < x_0 < y$, alors les deux inégalités des pentes entraînent $\tau_x(x_0) \leqslant \tau_x(y)$ et $\tau_y(x) \leqslant \tau_y(x_0)$. Or, $\tau_x(y) = \tau_y(x)$ donc par transitivité, $\tau_x(x_0) \leqslant \tau_y(x_0)$ et d'après $\tau_\star(\square) = \tau_\square(\star)$, 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}. D'après le théorème précédente, f est alors convexe.



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

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 $\mathbb{R}_+$?*

2.2 Cas des fonctions dérivables

Exercice 5 *Soit $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction continue croissante. Montrer que toute primitive de φ est convexe. En déduire que l'exponentielle est convexe et le logarithme népérien est concave.*

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.*

- *La fonction f est convexe sur I.*
- *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 prenant la limite $y \to x^+$on obtient $f_d'(x) \leqslant \frac{f(z)-f(x)}{z-x}$ (dérivée à droite).

— En prenant la limite $y \to z^-$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 : $\mathrm{D} : x \mapsto \frac{f(b)-f(a)}{b-a}(x-a) + f(a) - f(x)$. 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 $[a, c]$

— $\forall x \in [c,b], D'(x) \leqslant 0$, donc D est décroissante sur $[c,b]$. Comme enfin $D(a) = D(b) = 0$, 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* $f : I \to \mathbb{R}$ *une fonction dérivable sur* I*. Alors* f *est convexe si et seulement si*

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

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

Démonstration. Supposons f convexe. Soit $a \in I$. Étudions $D : I \to \mathbb{R}, x \mapsto f(x) - [f'(a)(x-a) + f(a)]$ dont la dérivée est donnée par $D' : x \mapsto f'(x) - f'(a)$. Puisque f' est croissante, $D'(x) \geqslant 0$ pour tout $x \geqslant a$ et $D'(x) \leqslant 0$ pour tout $x \leqslant a$: les variations de D s'en déduisent et montrent que D admet un minimum global en a. Ajouté au fait que $D(a) = 0$, nous en déduisons que D est toujours positive sur I, ce qui donne l'inégalité souhaitée.
Réciproquement, supposons $\forall x \in I, \forall a \in I \quad f(x) \geqslant f'(a)(x-a) + f(a)$. Soit $a < b$ dans I. Alors $f(a) \geq f(b) + f'(b)(a-b)$ et $f(b) \geq f(a) + f'(a)(b-a)$. On en déduit par sommation $(f'(b) - f'(a))(b-a) \geq 0$, ce qui équivaut à la croissance de f'

Corollaire (cas 2 fois dérivable)
Soit I *un intervalle de longueur non nulle et* $f : I \to \mathbb{R}$ *une fonction deux fois dérivable sur* I*. Les assertions suivantes sont équivalentes.*

- *La fonction* f *est convexe sur* I*.*
- *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 $\left[0, \frac{\pi}{2}\right]$: $\forall x \in \left[0, \frac{\pi}{2}\right], \frac{2}{\pi}x \leqslant \sin(x) \leqslant x$
- $]-1, +\infty[\to \mathbb{R}, x \mapsto \ln(1+x)$ est concave : $\forall x > -1, \ln(1+x) \leqslant x$
- $[-1, +\infty[, x \mapsto \sqrt{1+x}$ est concave : $\forall x \geqslant -1, \sqrt{1+x} \leqslant 1 + \frac{1}{2}x$
- Pour tout réel α, la fonction puissance $\mathbb{R}^{+*} \to \mathbb{R}, x \mapsto x^\alpha$ 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 Compléments

Théorème 7 *Si* I *est un intervalle de longueur non nulle et si* $f : I \to \mathbb{R}$ *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,

- *Les fonctions* f'_d *et* f'_g *sont croissantes sur* I*.*
- *La fonction* f *est continue sur l'intérieur de* I *(mais pas nécessairement sur* I *).*

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

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

Ainsi, τ_{x_0} bornée sur $]x_0 - \varepsilon, x_0 + \varepsilon[\setminus \{x_0\}$. Comme elle est croissante, le théorème de la limite monotone assure alors l'existence de limites finies 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, soit $z \in]x_0 - \varepsilon, x_0[$ et $z' \in]x_0, x_0 + \varepsilon[$, On a $\tau_{x_0}(z) \leqslant \tau_{x_0}(z')$ par croissance de τ_{x_0}. En prenant la limite $z \to x_0^-$, on obtient $f_g'(x_0) \leqslant \tau_{x_0}(z')$, puis en passant à la limite $z' \to x_0^+$, on trouve finalement $f_g'(x_0) \leqslant f_d'(x_0)$ comme attendu. Soit maintenant $a < b$ dans l'intérieur de I, 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, considérons le cas x_0 dans l'intérieur de I. 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 6 *Soit* $f : \mathrm{I} \to \mathbb{R}$ *une fonction continue sur un intervalle ouvert. Montrer que f est convexe sur* I *si et seulement si*

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

(on dit qu'une telle fonction est mid-convexe).

4 Quelques exercices

4.1 Énoncés

Exercice 7 *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 8 *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 9 *(inégalité arithmético-géométrique and Co). Soit* $n \in \mathbb{N}^*$ *et* $a_1, \ldots, a_n$ *des réels positifs. On appelle moyenne géométrique, 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*

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

Cette dernière quantité est appelée moyenne harmonique.

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 $x_1, \ldots, x_n \in I$, 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 10 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 11 *(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 12 *On dit qu'une fonction $f : I \to \mathbb{R}$ 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

A vous de jouer !