CHAPITRE 2

Fonctions convexes

I Définitions, pentes, régularité

Dans cette partie, on note I un intervalle de $\mathbb{R}$, non trivial et on considère f une fonction de I dans $\mathbb{R}$. Pour tout $(a,b) \in \mathbb{R}^2$, on note

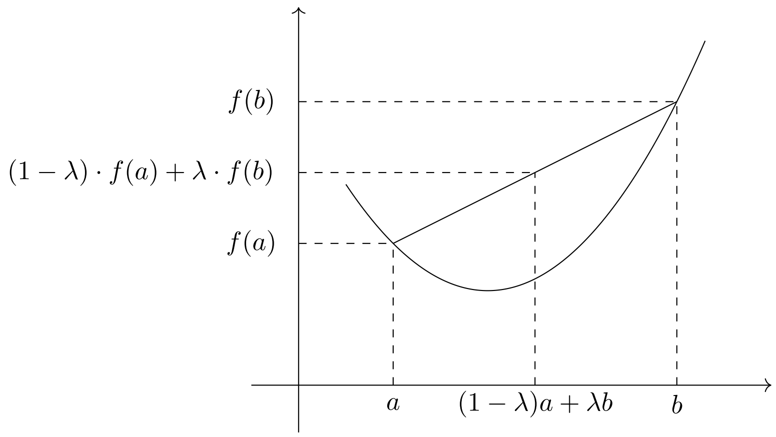
$$[a,b] = \{(1-t)a + tb,\ t \in [0,1]\}.$$

Définition I.1.

On dit que la fonction f est convexe lorsque

$$\forall (a,b) \in I^2\ ,\ \forall \lambda \in [0,1]\ ,\ f\left((1-\lambda)a + \lambda b\right) \leq (1-\lambda)f(a) + \lambda f(b).$$

Remarque. Graphiquement, la définition ci-haut se voit facilement sur un dessin : une fonction convexe sur un intervalle a, b est en dessous de la droite joignant les points $(a, f(a))$ et $(b, f(b))$ sur l'intervalle $[a,b]$. Ceci est aussi valable pour tout intervalle inclus dans I.



La droite qui joint $(a, f(a))$ et $(b, f(b))$ est la courbe de la fonction $g : x \mapsto \dfrac{f(b) - f(a)}{b-a} \cdot (x-a) + f(a)$. On a ainsi $g((1-\lambda)a + \lambda b) = (1-\lambda)\cdot f(a) + \lambda f(b)$.

Avant de faire l'exercice ci-dessous, le lecteur souhaitant revoir la définition et propriétés des ensembles convexes est invité à voir le chapitre 11.9 sur la convexité.

Exercice I.2.

Montrer que f est convexe si et seulement si l'ensemble $\mathrm{Epi}(f) = \{(x,y) \in I \times \mathbb{R}, f(x) \leq y\}$ est convexe.

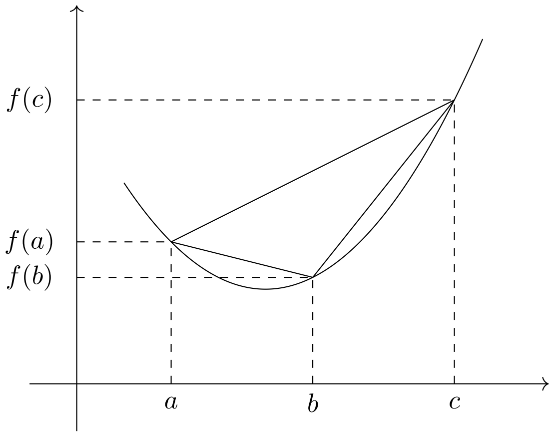
Théorème (Théorème des trois pentes) I.3.

On note pour tous $a, b \in I$ distincts, $p_{a,b}(f) = \frac{f(b)-f(a)}{b-a}$, la pente de la droite joignant les points $(a, f(a))$ et $(b, f(b))$. f est convexe sur I si et seulement si, pour tous $a, b, c \in I$, si $a < b < c$, alors on a

$$p_{a,b}(f) \leq p_{a,c}(f) \leq p_{b,c}(f).$$

Démonstration.

→ (⇒) Observons tout d'abord le dessin suivant, qui nous donne une idée de la preuve.



Lorsque f est convexe, la pente de la droite passant par deux points $(a, f(a)), (b, f(b))$ est croissante lorsque a et/ou b sont translatés vers la droite, chose qu'on observe sur le dessin ci-dessus. Montrons à présent ce résultat. Pour le montrer, il suffit de remarquer que étant donné que $b \in]a, c[$, on peut dire qu'il existe $\lambda \in]0, 1[$ tel que $b = \lambda a + (1 - \lambda)c$, et alors, on écrit

$$\begin{aligned} p_{a,b}(f) &= \frac{f(b) - f(a)}{b - a} = \frac{f(\lambda a + (1-\lambda)c) - f(a)}{(1-\lambda)(c-a)} \\ &\leq \frac{\lambda f(a) + (1-\lambda)f(c) - f(a)}{(1-\lambda)(c-a)} = \frac{f(c) - f(a)}{c - a} = p_{a,c}(f), \end{aligned}$$

$$\begin{aligned} p_{b,c}(f) &= \frac{f(b) - f(c)}{b - c} = \frac{f(\lambda a + (1-\lambda)c) - f(c)}{\lambda \underbrace{(a-c)}_{<0}} \\ &\geq \frac{\lambda f(a) + (1-\lambda)f(c) - f(c)}{\lambda(a-c)} = \frac{f(a) - f(c)}{a - c} = p_{a,c}(f). \end{aligned}$$

→ (⇐) Soient $a, b \in I$ tels que $a < b$, $\lambda \in]0, 1[$, et $m = \lambda a + (1 - \lambda)b$. L'inégalité supposée vraie donne $p_{a,m}(f) \leq p_{a,b}(f)$. On a alors

$$\begin{aligned} p_{a,m}(f) \leq p_{a,b}(f) &\Longrightarrow \frac{f(m) - f(a)}{m - a} \leq \frac{f(b) - f(a)}{b - a} \\ &\Longrightarrow \frac{f(m) - f(a)}{(1-\lambda)(b-a)} \leq \frac{f(b) - f(a)}{b - a} \\ &\Longrightarrow f(m) \leq f(a) + (1-\lambda)(f(b) - f(a)) = \lambda f(a) + (1-\lambda)f(b) \\ &\Longrightarrow f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda)f(b), \end{aligned}$$

ce qui est bien le résultat voulu. Les cas $\lambda \in \{0, 1\}$ et $a = b$ sont évidents, nous ne les traiterons donc pas.

□

Remarque. Le résultat démontré ci-dessus permet de montrer que la pente de la droite passant par deux points $(x, f(x))$ et $(a, f(a))$, lorsque f est convexe, est croissante en x. En effet, rigoureusement, cela signifie que la fonction

$$g : \begin{cases} I \setminus \{a\} & \longrightarrow \mathbb{R} \\ x & \longmapsto \dfrac{f(x)-f(a)}{x-a} \end{cases}$$

est croissante lorsque f est convexe. En effet, pour tout $x, y \in I \setminus \{a\}$ tels que $x < y$, on a

$$g(x) = \frac{f(x)-f(a)}{x-a} = p_{a,x}(f) \leq p_{a,y}(f) = g(y).$$

On va maintenant énoncer quelques propriétés de régularité des fonctions convexes.

Proposition I.4.

Si f est convexe dans I, alors les propositions suivantes sont vraies.

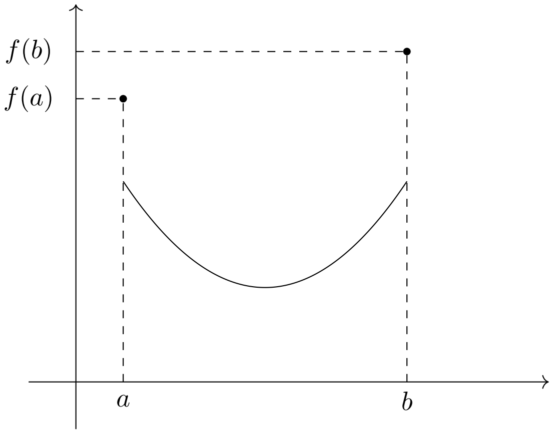
1. En tout point dans l'intérieur de I (différent des bords de I), f admet une dérivée à gauche et à droite. De plus, pour tout a dans l'intérieur de I, en notant f'_g et f'_d respectivement les dérivées à droite et à gauche de f en a, on a $f'_g(a) \leq f'_d(a)$.
2. f est continue en tout point dans l'intérieur de I.

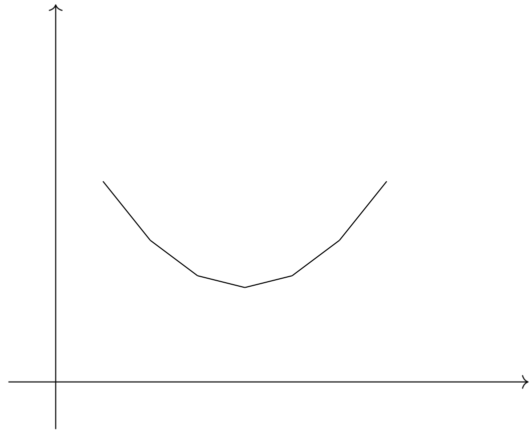
Démonstration.

1. Soit a dans l'intérieur de I. Il s'agit de montrer que le taux d'accroissement $x \longmapsto p_{a,x}(f)$ admet une limite à droite et à gauche de a. Posons encore une fois
$$g : \begin{cases} I \setminus \{a\} & \longrightarrow \mathbb{R} \\ x & \longmapsto \dfrac{f(x)-f(a)}{x-a} \end{cases}.$$
g est croissante sur $I \setminus \{a\}$. g admet donc des limites à droite et à gauche de a. L'implication entre monotonie et existence des limites à droite et à gauche est assez connue, mais nous la montrons quand même. Soit $b \in I$ tel que $b > a$. La restriction de g à $I \cap]-\infty, a[$ est croissante majorée par $g(b)$. On en déduit que g admet alors une limite à gauche de a. L'existence de la limite à droite peut être obtenue de la même manière. De plus, la croissance de g nous donne (par comparaison des limites à droite et à gauche) directement l'inégalité $f'_g(a) \leq f'_d(a)$.
2. D'après le point précédent, on sait que $g(x) \xrightarrow[x \to a^+]{} f'_d(a)$. Il existe alors $\eta > 0$ tel que pour tout $x \in]a, a+\eta[$
$$\left| \frac{f(x)-f(a)}{x-a} - f'_d(a) \right| \leq 1 \text{ et alors } |f(x)-f(a)| \leq |x-a| + |f'_d(a)(x-a)| \xrightarrow[x \to a^+]{} 0.$$
f est donc continue à droite de a. La continuité à gauche peut être démontrée de la même manière.

□

Remarque. Attention, il est possible que f soit convexe sur I, mais discontinue sur les bords de l'intervalle où elle est définie. En effet, on peut remarquer par exemple que la fonction ci-dessous à gauche, bien que discontinue aux bords, est convexe. Le lecteur pourra vérifier que la définition de la convexité s'applique bien dans ce cas. De même, il est possible qu'une fonction soit convexe mais pas forcément dérivable sur tout point de I. On peut voir de phénomène sur la figure de droite ci-dessous, par la présence de plusieurs points anguleux.





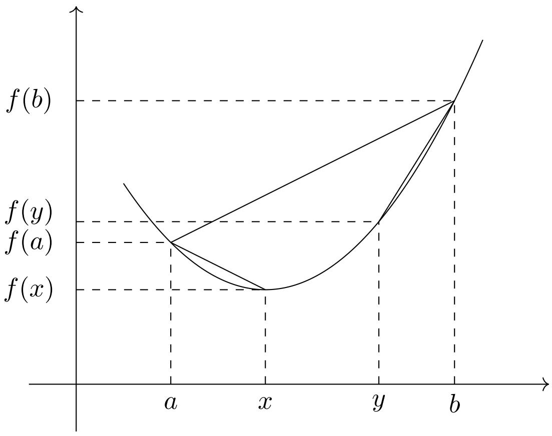
Corollaire I.5.

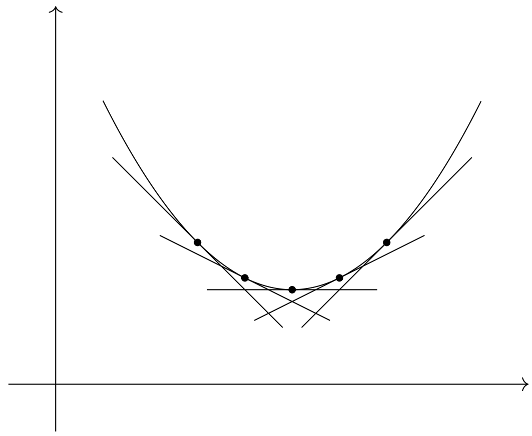
Si f est convexe et dérivable sur I, alors f' est croissante sur I.

Démonstration. Soit $a, b \in I$ tels que $a < b$. f est dérivable, donc pour tout point dans l'intérieur de I, les dérivées à droite et à gauche de f sont égales. Notre intuition pour montrer ce résultat est de comparer les dérivées de f respectivement à droite de a et à gauche de b. Pour ce faire, on va utiliser le théorème des trois pentes (théorème I.3) pour comparer la pente des tangentes en a et en b. En effet, on a pour tout $x, y \in I$ tels que $a < x < y < b$,

$$p_{a,x}(f) \leq p_{a,y}(f) \leq p_{b,y}(f).$$

On peut observer cette inégalité sur le dessin de gauche ci-dessous.





Ensuite, il suffit de faire tendre x vers a à droite et y vers b à gauche dans l'inégalité $p_{a,x}(f) \leq p_{b,y}(f)$ pour obtenir $\underbrace{f'_d(a)}_{=f'(a)} \leq \underbrace{f'_g(b)}_{=f'(b)}$, ce qui donne bien que f' est croissante sur I. □

Cette propriété de croissance de la dérivée peut être facilement vue graphiquement, comme on peut la voir sur la figure de droite ci-dessus.

Remarquons que sous les mêmes hypothèses, la dérivée de f est définie dans l'intérieur de I, et qu'uniquement les dérivées de gauche et de droite respectivement sont définies sur les bords de droite et de gauche de I. Le lecteur est donc invité à montrer que la dérivée à gauche du bord de droite de I est supérieure à la dérivée de droite du bord gauche de I, et que si b et c sont respectivement les bords de gauche et de droite de I, alors pour tout $a \in I$ où f est dérivable, $f'_d(b) \leq f'(a) \leq f'_g(c)$. On peut également montrer de la même manière, même en l'absence d'hypothèse de dérivabilité de f, que f'_d et f'_g sont croissantes sur I.

Proposition I.6.

Supposons que f est dérivable sur I. Si f' est croissante sur I, alors f est convexe sur I.

Démonstration. Supposons que f' est croissante sur I et supposons par l'absurde que f n'est pas convexe. D'après le théorème des 3 pentes (théorème I.3), il existe $a, b, c \in I$ tels que $a < b < c$ et $p_{a,b}(f) > p_{b,c}(f)$. En appliquant le théorème des accroissements finis sur $]a, b[$ et $]b, c[$, on peut affirmer qu'il existe $u \in]a, b[$ et $v \in]b, c[$ tels que

$$f'(u) = \frac{f(b) - f(a)}{b - a} = p_{a,b}(f) \text{ et } f'(v) = \frac{f(b) - f(c)}{b - c} = p_{b,c}(f).$$

Le fait que f' est croissante sur I nous donne que $f'(u) \leq f'(v)$, ce qui est absurde car cela implique que $p_{a,b}(f) \leq p_{b,c}(f)$. □

Corollaire I.7.

Si f est deux fois dérivable sur I, alors f est convexe si et seulement si f'' est positive sur I.

Démonstration. Il suffit d'utiliser le fait que lorsque f est deux fois dérivable, qu'il y a équivalence entre f'' positive et f' croissante. □

Proposition (position par rapport à la tangente) I.8.

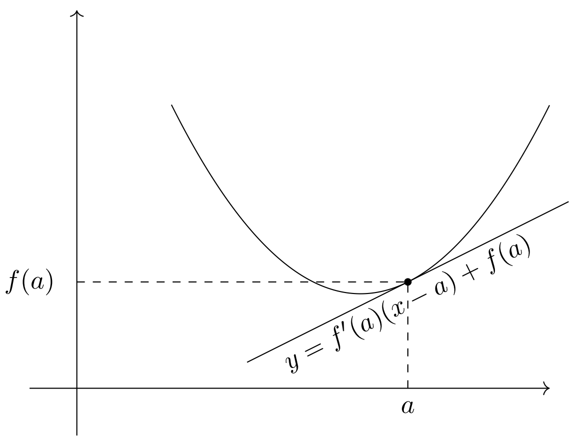
Si f est convexe sur I, alors pour tout $a \in I$, si f est dérivable en a, alors f est au dessus de sa tangente en a, c'est-à-dire que pour tout $x \in I$,

$$f(x) \geq f'(a)(x - a) + f(a).$$

Démonstration. On utilise encore le théorème des trois pentes de la manière suivante. On a pour tout $x, y \in I$ tels que $x > y > a$, $p_{a,x}(f) \geq p_{a,y}(f)$. En passant à la limite lorsque y tend vers a à droite, on obtient que

$$p_{a,x}(f) \geq f'(a) \;\; i.e. \;\; f(x) \geq f'(a)(x - a) + f(a).$$

Le cas $x < a$ se fait de manière identique, et le cas $x = a$ est évident. On a donc bien le résultat voulu. □



Remarques.

→ Attention, ici, on a uniquement supposé que f est dérivable en a, mais pas autre-part.

→ Même lorsque f n'est pas dérivable en a, on peut montrer que la même inégalité est toujours vraie en substituant f' par f'_d ou f'_g.

II Inégalités de convexité

1. Fonctions strictement convexes

On conserve les notations de la partie précédente : on note I un intervalle de $\mathbb{R}$, non trivial et on considère f une fonction de I dans $\mathbb{R}$.

Définition II.1.

On dit que f est strictement convexe lorsque pour tous $a, b \in I$ distincts,

$$\forall \lambda \in]0,1[,\ f((1-\lambda)a+\lambda b) < (1-\lambda)f(a)+\lambda f(b).$$

Remarques.

→ Lorsque f est strictement convexe, f ne peut être plate nulle part, c'est-à-dire qu'aucune portion de sa courbe ne doit être un segment.

→ Lorsque f est strictement convexe, l'inégalité des 3 pentes s'applique et devient stricte.

Proposition II.2.

Les propositions suivantes sont vraies.

1. Si f est dérivable sur I et f' est strictement croissante, alors f est strictement convexe.
2. Si f est deux fois dérivable sur I et f'' est strictement positive sur I, alors f est strictement convexe.

Démonstration. La preuve est quasi-identique aux preuves précédentes pour la convexité et est donc laissée comme exercice au lecteur. □

Remarque. On sait que si f est dérivable, alors f est croissante sur I si et seulement si f' est positive sur I. Par contre, si f est strictement croissante, alors ceci n'est pas équivalent au fait que f' est strictement positive sur I. En effet, la stricte croissance de f est équivalente à une condition qui est plus faible que la stricte positivité de la dérivée de f, qui est la suivante : f' est positive sur I, et $Z(f) = \{x \in I,\ f'(x) = 0\}$ est d'intérieur vide. Montrons cette équivalence.

→ Supposons que f' est positive sur I, et $Z(f) = \{x \in I,\ f'(x) = 0\}$ est d'intérieur vide. Si f n'est pas strictement croissante, alors il existe $\alpha, \beta \in I$ tels que $\alpha < \beta$ et $f(\alpha) = f(\beta)$. On en déduit donc que f est constante sur $[\alpha, \beta]$, et que donc sa dérivée est nulle sur $]\alpha, \beta[$, *i.e.* $]\alpha, \beta[\subset Z(f)$, ce qui est absurde.

→ Supposons que f est strictement croissante sur I. On a alors f' est positive sur I. Supposons que $Z(f)$ n'est pas d'intérieur vide, i.e. il existe $\alpha, \beta \in I$ tels que $\alpha < \beta$ et $]\alpha, \beta[\subset Z(f)$, i.e. f' est nulle sur $]\alpha, \beta[$ et est constante sur cet intervalle, ce qui est absurde car f est strictement croissante.

2. Inégalité fondamentale

Proposition II.3.

Si f est convexe, alors pour tout $n \in \mathbb{N}^*$, $x_1, \ldots, x_n \in I$ et $\lambda_1, \ldots, \lambda_n \in [0,1]$, tels que $\lambda_1 + \cdots + \lambda_n = 1$,

$$f\left(\sum_{k=1}^{n} \lambda_k x_k\right) \leq \sum_{k=1}^{n} \lambda_k f(x_k).$$

Si de plus, $x_1, \ldots x_n$ sont non tous égaux, $\lambda_1, \ldots, \lambda_n$ sont tous strictement positifs, et f est strictement convexe, alors l'inégalité devient stricte. En effet, lorsque f est strictement convexe et pour tout $k \in [\![1;n]\!]$, $\lambda_k > 0$, l'inégalité est une égalité si et seulement si $x_1 = \cdots = x_n$.

Démonstration. Nous allons montrer l'inégalité dans le cas où f est convexe. Montrons le résultat par récurrence. Lorsque $n = 1$, le résultat est évident. Soit $n \in \mathbb{N}^*$. Supposons que le résultat est vérifié pour n. Soit $x_1, \ldots, x_{n+1} \in I$, et $\lambda_1, \ldots, \lambda_{n+1} \in [0,1]$ tels que $\lambda_1 + \cdots + \lambda_{n+1} = 1$. On suppose sans perte de généralité que $\lambda_{n+1} < 1$ (le cas $\lambda_{n+1} = 1$ est évident). On a

$$\begin{aligned}
f\left(\sum_{k=1}^{n+1} \lambda_k x_k\right) &= f\left((1-\lambda_{n+1})\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k + \lambda_{n+1} x_{n+1}\right) \\
&\underset{(1)}{\leq} (1-\lambda_{n+1}) f\left(\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} x_k\right) + \lambda_{n+1} f(x_{n+1}) \\
&\underset{(2)}{\leq} (1-\lambda_{n+1})\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} f(x_k) + \lambda_{n+1} f(x_{n+1}) = \sum_{k=1}^{n+1} \lambda_k f(x_k).
\end{aligned}$$

L'inégalité (1) est vraie par convexité de f, et (2) est vraie par hypothèse de récurrence étant donné que f est convexe, $1 - \lambda_{n+1} = \lambda_1 + \cdots + \lambda_n$, et

$$\sum_{k=1}^{n} \frac{\lambda_k}{1-\lambda_{n+1}} = \sum_{k=1}^{n} \frac{\lambda_k}{\underbrace{\lambda_1 + \cdots + \lambda_n}_{\in]0,1[}} = 1.$$

On a donc bien montré l'inégalité voulue.

Montrons à présent par récurrence que le cas d'égalité, lorsque f est strictement convexe et $\lambda_1, \ldots, \lambda_n \in]0,1[$, implique $x_1 = \cdots = x_n$ (l'implication réciproque est évidente). Lorsque $n = 1$, la propriété est vérifiée. Soit $n \in \mathbb{N}^*$ tel que la propriété est vérifiée pour n. Montrons qu'elle est vérifiée pour $n+1$. On veut d'abord essayer d'utiliser l'inégalité de convexité stricte pour le cas $n = 2$, comme on l'a fait dans la preuve pour l'inégalité. On a

$$\sum_{k=1}^{n+1} \lambda_k x_k = \underbrace{(1-\lambda_{n+1})}_{]0,1[} \sum_{k=1}^{n} \underbrace{\frac{\lambda_k}{1-\lambda_{n+1}}}_{\mu_k} x_k + \lambda_{n+1} x_{n+1}.$$

On a donc, si $x_{n+1} \neq \sum_{k=1}^{n} \mu_k x_k$, alors étant donné que $\lambda_{n+1} \in]0,1[$ et $1 - \lambda_{n+1} \in]0,1[$, l'inégalité de stricte convexité donne

$$\begin{aligned}
f\left(\sum_{k=1}^{n+1}\lambda_k x_k\right) &= f\left((1-\lambda_{n+1})\sum_{k=1}^{n}\mu_k x_k + \lambda_{n+1}x_{n+1}\right)\\
&< (1-\lambda_{n+1})f\left(\sum_{k=1}^{n}\frac{\lambda_k}{1-\lambda_{n+1}}x_k\right) + \lambda_{n+1}f(x_{n+1})\\
&\leq (1-\lambda_{n+1})\sum_{k=1}^{n}\frac{\lambda_k}{1-\lambda_{n+1}}f(x_k) + \lambda_{n+1}f(x_{n+1}) = \sum_{k=1}^{n+1}\lambda_k f(x_k),
\end{aligned}$$

ce qui est absurde. On en déduit donc qu'on a bien $x_{n+1} = \sum_{k=1}^{n}\mu_k x_k$, et alors en substituant x_{n+1} par $\sum_{k=1}^{n}\mu_k x_k$, l'égalité devient

$$f\left(\sum_{k=1}^{n}\mu_k x_k\right) = \sum_{k=1}^{n}\mu_k f(x_k),$$

ce qui implique par hypothèse de récurrence, étant donné que pour tout $k \in [\![1;n]\!]$, $\mu_k > 0$, que $x_1 = \cdots = x_n$ et que finalement $x_1 = \cdots = x_n = x_{n+1}$, ce qui est bien le résultat voulu. Remarquons qu'on aurait aussi pu se passer de la récurrence en supposant sans perte de généralité que $x_{n+1} = \max_{k\in[\![1;n+1]\!]} x_k$. $\square$

Exemple. En fait, avec cette inégalité, on peut facilement montrer l'inégalité arithmético-géométrique. En effet, l'énoncé de cet inégalité est le suivant. Pour tout $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n \in \mathbb{R}^+$, on a

$$\frac{x_1 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \ldots x_n},$$

avec égalité si et seulement si tous les x_i sont égaux.

Lorsque l'un des x_i est nul, l'inégalité est facilement vérifiable. Supposons donc que $x_1, \ldots, x_n \in \mathbb{R}_+^*$. L'inégalité est équivalente à

$$\frac{e^{\ln x_1} + \cdots + e^{\ln x_n}}{n} \geq e^{\frac{\ln x_1 + \cdots + \ln x_n}{n}}.$$

Il est donc facile de voir qu'il suffit d'utiliser la stricte convexité de l'exponentielle pour obtenir le résultat voulu.

Exercice II.4.

Soit $a, b, c \in \mathbb{R}_+^*$. Calculer

$$\inf_{x,y\in\mathbb{R}_+^*}\left(ax + by + \frac{c}{xy}\right).$$

Exercice II.5.

Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue.

1. Montrer que g est convexe si et seulement si
$$\forall x, y \in \mathbb{R},\ g\left(\frac{x+y}{2}\right) \leq \frac{g(x)+g(y)}{2}.$$
2. On ne suppose plus que g est continue, mais plutôt que g est bornée sur tout segment de $\mathbb{R}$ et qu'elle vérifie l'inégalité ci-dessus. Montrer que g est convexe.

III Compléments

1. Inégalités de Hölder et de Minkowski

Proposition (Inégalité de Hölder) III.1.

Soient $p, q \in \mathbb{R}_+^*$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Pour tout $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$, on a
$$\sum_{k=1}^{n} |x_k y_k| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}}.$$

Démonstration. On suppose sans perte de généralité que pour tout $k \in [\![1; n]\!]$, $x_k \neq 0$ et $y_k \neq 0$ (le lecteur pourra s'assurer qu'on peut effectivement faire cette hypothèse). On pose alors pour tout $k \in [\![1; n]\!]$, $e^{s_k} = |x_k|$ et $e^{t_k} = |y_k|$. On a alors pour tout $k \in [\![1; n]\!]$, en utilisant la convexité de l'exponentielle,
$$|x_k y_k| = e^{s_k + t_k} = e^{\frac{1}{p} p s_k + \frac{1}{q} q t_k} \leq \frac{1}{p} e^{p s_k} + \frac{1}{q} e^{q t_k} = \frac{1}{p} |x_k|^p + \frac{1}{q} |y_k|^q$$

Posons $A = \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \neq 0$ et $B = \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}} \neq 0$. En substituant dans l'inégalité ci-dessus pour tout $k \in [\![1; n]\!]$ x_k par $\frac{x_k}{A}$ et y_k par $\frac{y_k}{B}$ et en sommant, on obtient
$$\frac{1}{AB} \sum_{k=1}^{n} |x_k y_k| \leq \frac{1}{p} \underbrace{\frac{\sum_{k=1}^{n} |x_k|^p}{A^p}}_{=1} + \frac{1}{q} \underbrace{\frac{\sum_{k=1}^{n} |x_k|^q}{B^q}}_{=1} = 1,$$

et finalement en multipliant par AB des deux côtés,
$$\sum_{k=1}^{n} |x_k y_k| \leq \left(\sum_{k=1}^{n} |x_k|^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^{n} |y_k|^q\right)^{\frac{1}{q}}.$$

□

Remarque. Cette inégalité est également vérifiée lorsqu'on remplace les sommes par des intégrales.

Autrement dit, pour tout f, g fonctions continues sur un intervalle $[a, b]$, on a

$$\int_a^b |f(x)g(x)|\,dx \leq \left(\int_a^b |f(x)|^p\,dx\right)^{\frac{1}{p}} \left(\int_a^b |g(x)|^q\,dx\right)^{\frac{1}{q}}.$$

la preuve de cette inégalité se fait de manière identique, en substituant somme par intégrale.

Lemme (Quasi-linéarité) III.2.

Soit $n \in \mathbb{N}^*$ et L une fonction de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$, continue et bornée sur les ensembles bornés, telle que pour tout $y \in \mathbb{R}^n$, l'application $x \longmapsto L(x, y)$ est linéaire. Si $B \subset \mathbb{R}^n$ est un borné de $\mathbb{R}^n$, alors $f : x \longmapsto \sup\limits_{y \in B} L(x, y)$ est sous-additive, c'est-à-dire que

$$\forall x_1, x_2 \in \mathbb{R}^n,\ f(x_1 + x_2) \leq f(x_1) + f(x_2).$$

Démonstration. Il suffit d'écrire les définition. En effet, on a pour tout $y \in B$ et $x_1, x_2 \in \mathbb{R}^n$,

$$L(x_1 + x_2, y) = L(x_1, y) + L(x_2, y) \leq \sup_{y \in B} L(x_1, y) + \sup_{y \in B} L(x_2, y) = f(x_1) + f(x_2).$$

En passant alors à la borne supérieure en $y \in B$ dans la partie gauche de l'inégalité, on obtient

$$f(x_1 + x_2) \leq f(x_1) + f(x_2).$$

□

Proposition (Inégalité de Minkowski) III.3.

Soit $p > 1$. On a alors pour tout $n \in \mathbb{N}^*$ et $x_1, \ldots, x_n, y_1, \ldots, y_n \in \mathbb{R}$,

$$\left(\sum_{k=1}^n |x_k + y_k|^p\right)^{\frac{1}{p}} \leq \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} + \left(\sum_{k=1}^n |y_k|^p\right)^{\frac{1}{p}}$$

Démonstration. En fait, cette inégalité est équivalente au fait de dire que la fonction

$$f : \begin{cases} \mathbb{R}^n & \longrightarrow \mathbb{R} \\ (x_1 \ldots, x_n) & \longmapsto \left(\sum\limits_{k=1}^n |x_k|^p\right)^{\frac{1}{p}} \end{cases}$$

est sous-additive. Pour prouver ce résultat, nous allons utiliser le lemme III.2, en montrant qu'il existe $L : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ linéaire en son premier argument, ainsi que bornée sur les bornés, et un ensemble borné $B \subset \mathbb{R}^n$ tels que pour tout $x_1, \ldots, x_n \in R$, $f(x_1, \ldots, x_n) = \sup\limits_{z \in B} L(x_1, \ldots, x_n, z)$. Posons pour tout $x_1, \ldots, x_n, z_1, \ldots, z_n \in \mathbb{R}$,

$$L(x_1, \ldots, x_n, z_1, \ldots, z_n) = \sum_{k=1}^n x_k z_k,$$

puis considérons $q > 1$ tel que $\dfrac{1}{p} + \dfrac{1}{q} = 1$, et le sous-ensemble borné de $\mathbb{R}^n$

$$B = \left\{ (z_1, \ldots, z_n) \in \mathbb{R}^n,\ \sum_{k=1}^{n} |z_k|^q = 1 \right\}.$$

L est clairement linéaire en son premier argument et bornée sur les bornés. On veut alors montrer que pour tout $(x_1, \ldots, x_n) \in \mathbb{R}^n \setminus \{(0, \ldots, 0)\}$,

$$\sup_{(z_1,\ldots,z_n)\in B} L(x_1, \ldots, x_n, z_1, \ldots, z_n) = f(x_1, \ldots, x_n).$$

En utilisant l'inégalité de Hölder (proposition III.1), on a

$$\sum_{k=1}^{n} x_k z_k \leq \sum_{k=1}^{n} |x_k z_k| \leq \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}} \underbrace{\left(\sum_{k=1}^{n} |z_k|^q \right)^{\frac{1}{q}}}_{=1} = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}}.$$

Il suffit maintenant de montrer que cette borne supérieure est atteinte. Trouvons donc $(z_1, \ldots, z_n) \in B$ réalisant cette borne supérieure. Dans le but de faire apparaitre $|x_k|^p$, $k \in [\![1; n]\!]$, on est tentés de poser pour tout $k \in [\![1; n]\!]$, $z_k = \mathrm{signe}(x_k)\, |x_k|^{p-1}$. Cependant, $z_1, \ldots, z_n$ tels que défini risque de ne pas appartenir à B, donc il faut ajouter une constante de normalisation. Posons donc pour tout $k \in [\![1; n]\!]$, $z_k = \lambda \mathrm{signe}(x_k)\, |x_k|^{p-1}$ où λ est une constante positive qui nous permet de garantir que $(z_1, \ldots, z_n) \in B$. On a alors

$$1 = \sum_{k=1}^{n} |z_k|^q = \sum_{k=1}^{n} \lambda^q\, |x_k|^{q(p-1)} = \lambda^q \underbrace{\sum_{k=1}^{n} |x_k|^p}_{\neq 0} \text{ et donc } \lambda = \frac{1}{\left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{q}}}.$$

On en déduit que

$$\sum_{k=1}^{n} x_k z_k = \sum_{k=1}^{n} \lambda \mathrm{signe}(x_k)\, |x_k|^{p-1} x_k = \lambda \sum_{k=1}^{n} |x_k|^p = \frac{\sum_{k=1}^{n} |x_k|^p}{\left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{q}}} = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}}.$$

On a donc bien que

$$\sup_{(z_1,\ldots,z_n)\in B} L(x_1, \ldots, x_n, z_1, \ldots, z_n) = \sup_{(z_1,\ldots,z_n)\in B} \sum_{k=1}^{n} x_k y_k = \left(\sum_{k=1}^{n} |x_k|^p \right)^{\frac{1}{p}} = f(x_1, \ldots, x_n),$$

et alors le lemme III.2 permet de conclure. □

Remarques.

→ Attention, l'inégalité de Minkowski n'est pas valable lorsque $p < 1$. En effet, par exemple, pour tout $p < 1$, lorsque $n = 2$, dans le cas où $(x_1, x_2) = (1, 0)$ et $(y_1, y_2) = (0, 1)$, le côté gauche de l'inégalité est égal à $2^{\frac{1}{p}}$ et le côté droit à 2, et on a clairement $2^{\frac{1}{p}} > 2$.

→ Bien entendu, cette inégalité admet également une version continue : pour toutes fonctions f, g continues sur un intervalle $[a, b]$, et $p > 1$,

$$\left(\int_a^b |f(x) + g(x)|^p \, dx \right)^{\frac{1}{p}} \leq \left(\int_a^b |f(x)|^p \, dx \right)^{\frac{1}{p}} + \left(\int_a^b |g(x)|^p \, dx \right)^{\frac{1}{p}}.$$

2. Inégalité de Jensen

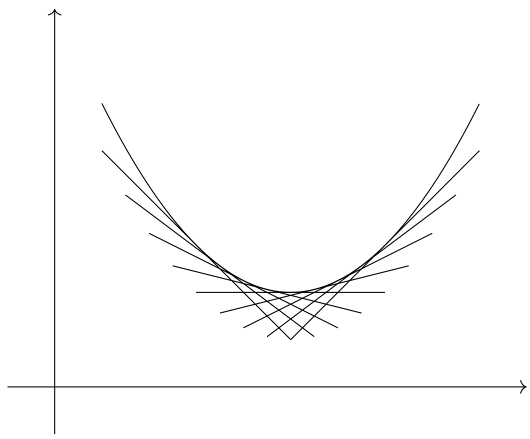
Proposition (Inégalité de Jensen) III.4.

Soit $g : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction de $\mathbb{R}$ dans $\mathbb{R}$, et f une fonction continue de $[a, b]$ dans $\mathbb{R}$ où $[a, b]$ est un intervalle non vide de $\mathbb{R}$ et non réduit à un singleton. Les deux propositions suivantes sont vraies.

1. Si g est convexe, alors il existe une famille de fonctions affines $(\varphi_\lambda)_{\lambda\in\Lambda}$ où Λ est un ensemble, telle que pour tout $x \in \mathbb{R}$, $g(x) = \sup\limits_{\lambda\in\Lambda} \varphi_\lambda(x)$.
2. Si g est convexe et $a < b$, alors l'inégalité
$$g\left(\frac{1}{b-a}\int_a^b f(x)dx\right) \leq \frac{1}{b-a}\int_a^b g\circ f(x)dx$$
est vérifiée.

Démonstration.

1. Pour simplifier, on suppose que g est dérivable en tout point de $\mathbb{R}$. En rejetant un coup d'oeil à la proposition I.8, on est tentés de proposer la famille $(\varphi_\lambda)_{\lambda\in\mathbb{R}}$ des tangentes à la courbe en tout point de $\mathbb{R}$, c'est-à-dire
$$\forall\lambda \in \mathbb{R},\ \forall x \in \mathbb{R},\ \varphi_\lambda(x) = g'(\lambda)(x-\lambda) + g(\lambda).$$
On sait d'après la proposition I.8 que pour tout $x \in \mathbb{R}$ et tout $a \in \mathbb{R}$, $g(x) \geq \varphi_a(x)$. De plus, pour tout $x \in \mathbb{R}$, $g(x) = \varphi_x(x)$, et alors, on a bien que $g(x) = \sup\limits_{\lambda\in\mathbb{R}} \varphi_\lambda(x)$. Maintenant, on peut voir qu'on peut faire exactement la même chose en se passant de l'hypothèse de dérivabilité de g, et en substituant g' par g'_d ou g'_g. En fait, on peut voir cette propriété graphiquement : une fonction convexe (lorsqu'elle est dérivable) est l'enveloppe supérieure de la famille de ses tangentes.



2. Pour tout $\lambda \in \mathbb{R}$, φ_λ étant une fonction affine, on a
$$\varphi_\lambda\left(\frac{1}{b-a}\int_a^b f(x)dx\right) = \frac{1}{b-a}\int_a^b \varphi_\lambda\circ f(x)dx \leq \frac{1}{b-a}\int_a^b g\circ f(x)dx.$$
En passant à la borne supérieure en $\lambda \in \mathbb{R}$ à gauche de l'inégalité, on obtient bien
$$g\left(\frac{1}{b-a}\int_a^b f(x)dx\right) \leq \frac{1}{b-a}\int_a^b g\circ f(x)dx.$$

$\square$

Remarques.

→ En fait, cette inégalité peut être vue comme la version continue de l'inégalité de la proposition II.3 lorsque les poids (la valeur des λ_i) attribués à chaque terme sont égaux. Lorsqu'on attribue des poids quelconques à chaque élément de la somme (qui est ici une somme continue, c'est-à-dire une intégrale), on obtient une inégalité plus générale : pour toute fonction continue $m : [a, b] \longrightarrow \mathbb{R}^+$ telle que

$$\int_a^b m(x)dx = 1,$$

On peut affirmer que

$$g\left(\int_a^b m(x)f(x)dx\right) \leq \int_a^b m(x)g \circ f(x)dx.$$

Cette inégalité se démontre exactement de la même manière que la précédente, à quelques détails mineurs près. Le lecteur est fortement encouragé à la démontrer lui-même. Ici, pour tout $x \in [a, b]$, on attribue à $f(x)$ le poids $m(x)$, contrairement au cas précédent où on attribuait à $f(x)$ le poids $\dfrac{1}{b-a}$.

→ Lorsque g est strictement convexe, alors il y a égalité si et seulement si f est constante sur $[a, b]$.

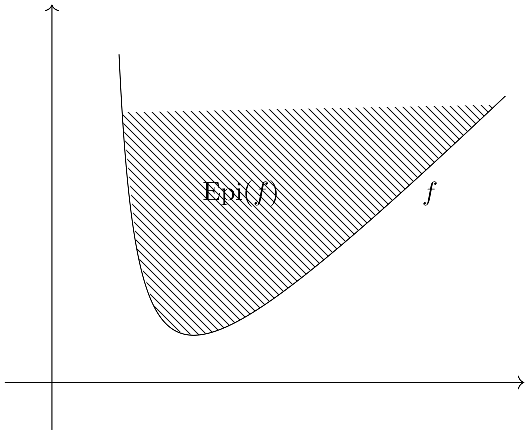
→ La réciproque du point (1) de la proposition III.4 est également vraie. En effet, soit $(\varphi_\lambda)_{\lambda\in\Lambda}$ une famille de fonctions affines telles que pour tout $x \in \mathbb{R}$, $g(x) = \sup_{\lambda\in\Lambda} \varphi_\lambda(x)$. On a alors, $\mathrm{Epi}(g) = \bigcap_{\lambda\in\Lambda} \mathrm{Epi}(\varphi_\lambda)$. De plus, les fonctions de la famille $(\varphi_\lambda)_{\lambda\in\Lambda}$ étant toutes convexes, on peut affirmer en utilisant le résultat de l'exercice I.2 que $\mathrm{Epi}(\varphi_\lambda)$ est convexe pour tout $\lambda \in \Lambda$. $\mathrm{Epi}(g)$ est donc l'intersection d'ensembles convexes, et est alors convexe. En utilisant encore une fois le résultat de l'exercice I.2, on peut affirmer que g est convexe. Remarquons qu'on a uniquement utilisé dans cette preuve le fait que les fonctions de la famille $(\varphi_\lambda)_{\lambda\in\Lambda}$ sont convexes. En particulier, cette preuve nous permet de dire que pour toute famille de fonctions convexes $(\varphi_\lambda)_{\lambda\in\Lambda}$, la fonction $x \longmapsto \sup_{\lambda\in\Lambda} \varphi_\lambda(x)$, lorsqu'elle est bien définie en tout point de $\mathbb{R}$ (*i.e.* la famille $(\varphi_\lambda(x))_{\lambda\in\Lambda}$ est majorée pour tout $x \in \mathbb{R}$), est convexe.

→ Les inégalités de Hölder, Minkowski et Jensen sont également vérifiées pour les variables aléatoires. En particulier, pour tout espace probabilisé Ω, variables aléatoires $X : \Omega \longrightarrow E$ et $Y : \Omega \longrightarrow E$ où E est un sous-ensemble fini de $\mathbb{R}$, fonction convexe $g : \mathbb{R} \longrightarrow \mathbb{R}$, et $p, q \in \mathbb{R}_+^*$ tels que $\dfrac{1}{p} + \dfrac{1}{q} = 1$, on a les inégalités suivantes.

$$\mathbb{E}(|XY|) \leq \mathbb{E}\left(|X|^p\right)^{\frac{1}{p}} \mathbb{E}\left(|Y|^q\right)^{\frac{1}{q}}$$
$$\mathbb{E}(|X+Y|^p)^{\frac{1}{p}} \leq \mathbb{E}\left(|X|^p\right)^{\frac{1}{p}} + \mathbb{E}\left(|Y|^p\right)^{\frac{1}{p}}$$
$$g(\mathbb{E}(X)) \leq \mathbb{E}(g(X)).$$

Correction de l'exercice I.2 :

On rappelle qu'un ensemble A est convexe si et seulement si pour tout $a, b \in A$ et $\lambda \in [0,1]$, $\lambda a+(1-\lambda)b \in A$. L'énoncé de cet exercice est assez intuitif, et peut se reformuler de la manière suivante : f est une fonction convexe si et seulement si la partie au-dessus de sa courbe est convexe. Pour en avoir une intuition plus forte, observons le dessin suivant.



Montrons à présent ce résultat.

→ (⇒) Supposons que f est convexe. Soit $(x_1, y_1), (x_2, y_2) \in \mathrm{Epi}(f)$. On a alors pour tout $\lambda \in [0,1]$,

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \leq \lambda y_1 + (1-\lambda)y_2.$$

On en déduit donc qu'on a bien $\lambda(x_1, y_1) + (1-\lambda)(x_2, y_2) \in \mathrm{Epi}(f)$.

→ (⇐) Supposons que $\mathrm{Epi}(f)$ est convexe. Soit $x_1, x_2 \in I$. On sait que $(x_1, f(x_1)) \in \mathrm{Epi}(f)$ et $(x_2, f(x_2)) \in \mathrm{Epi}(f)$. On a donc pour tout $\lambda \in [0,1]$,

$$\lambda(x_1, f(x_1)) + (1-\lambda)(x_2, f(x_2)) \in \mathrm{Epi}(f) \text{ } i.e. \text{ } f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2),$$

ce qui est bien le résultat voulu.

Correction de l'exercice II.4 :

Il suffit d'utiiser'inégalité arithmético-géométrique. En effet, a on a pour tout $x, y \in \mathbb{R}_+^*$,

$$ax + by + \frac{c}{xy} \geq 3\sqrt[3]{abc},$$

avec égalité si et seulement si $ax = by = \dfrac{c}{xy}$, c'est à dire que

$$x = \sqrt[3]{\frac{bc}{a^2}} \text{ et } y = \sqrt[3]{\frac{ac}{b^2}}.$$

On en déduit donc que

$$\inf_{x,y \in \mathbb{R}_+^*} \left(ax + by + \frac{c}{xy} \right) = 3\sqrt[3]{abc},$$

et que cette borne inférieure est atteinte lorsque $x = \sqrt[3]{\dfrac{bc}{a^2}}$ et $y = \sqrt[3]{\dfrac{ac}{b^2}}$.

Correction de l'exercice II.5 :

1. L'implication directe est évidente : il s'agit de la définition de la convexité avec $\lambda = \frac{1}{2}$. Montrons l'implication réciproque. Supposons que

$$\forall x, y \in \mathbb{R},\ g\left(\frac{x+y}{2}\right) \leq \frac{g(x)+g(y)}{2}.$$

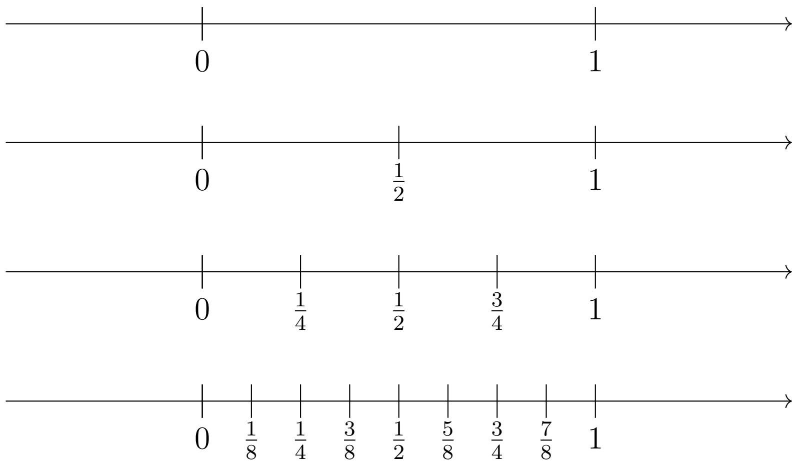
On remarque qu'on peut itérer cette inégalité de la manière suivante : pour tous $x, y \in \mathbb{R}$,

$$g\left(\frac{\frac{x+y}{2}+y}{2}\right) \leq \frac{g\left(\frac{x+y}{2}\right)+g(y)}{2} \leq \frac{g(x)+3g(y)}{4}\ \ i.e.\ g\left(\frac{x+3y}{4}\right) \leq \frac{g(x)+3g(y)}{4}.$$

Notre intuition nous pousse à chercher les $\lambda \in [0,1]$ tels que

$$\forall x, y \in \mathbb{R},\ g(\lambda x + (1-\lambda)y) \leq \lambda g(x) + (1-\lambda)g(y).$$

On pose Λ l'ensemble des $\lambda \in [0,1]$ qui vérifient l'inégalité ci-dessus. On peut facilement montrer que pour tout $\lambda_1, \lambda_2 \in \Lambda$, $\frac{\lambda_1+\lambda_2}{2} \in \Lambda$. On a de plus, $1, 0, \frac{1}{2} \in \Lambda$. On veut maintenant trouver tous les éléments de Λ. Regardons à quoi rassemble Λ sur un segment. On sait que pour chaque paire de points de Λ sur le segment, le milieu du segment délimité par ces deux points est aussi dans Λ. On peut donc retrouver les éléments de Λ en ajoutant tous milieux de segments qu'on peut trouver à chaque étape, comme sur le dessin ci-dessous.



En regardant le dessin ci-dessus, on peut intuitivement en déduire qu'on peut montrer que pour tout $n \in \mathbb{N}$ et $k \in [\![0; 2^n]\!]$

$$\forall x, y \in \mathbb{R},\ g\left(\frac{k}{2^n}x + \left(1-\frac{k}{2^n}\right)y\right) \leq \frac{k}{2^n}g(x) + \left(1-\frac{k}{2^n}\right)g(y),$$

c'est-à-dire que $\frac{k}{2^n} \in \Lambda$. Pour le montrer, il suffit de faire une récurrence, dont l'idée est exactement décrite pas les dessins ci-dessus. Pour $n = 0$, la propriété est vraie. Soit $n \in \mathbb{N}$ tel que la propriété est vérifiée pour n. Soit $k \in [\![0; 2^{n+1}]\!]$. Si k est pair, alors en posant $k = 2l, l \in \mathbb{N}$, on a $\frac{k}{2^{n+1}} = \frac{l}{2^n} \in \Lambda$. Sinon, on pose $k = 2l+1, l \in \mathbb{N}$. On a alors (on cherche parmi les milieux des segments obtenus par la subdivisons précédente)

$$\frac{\overbrace{\frac{l}{2^n}}^{\in\Lambda} + \overbrace{\frac{l+1}{2^n}}^{\in\Lambda}}{2} \in \Lambda \text{ et donc } \frac{k}{2^{n+1}} \in \Lambda,$$

ce qui est bien la propriété désirée. On en déduit donc que Λ contient tous les nombres rationnels entre 0 et 1 dont le dénominateur est une puissance de 2. On va donc approcher tout nombre $\lambda \in [0,1]$ par une suite à valeurs dans Λ. Soit $\lambda \in [0,1]$. On pose pour tout $n \in \mathbb{N}$, $\lambda_n = \dfrac{\lfloor 2^n \lambda \rfloor}{2^n}$. La suite (λ_n) est à valeurs dans Λ, et de plus, on a

$$\lambda_n = \frac{\lfloor 2^n \lambda \rfloor}{2^n} = \lambda - \frac{\overbrace{2^n\lambda - \lfloor 2^n \lambda \rfloor}^{\in[0,1]}}{2^n} \xrightarrow[n\to+\infty]{} \lambda.$$

On a alors pour tous $x, y \in \mathbb{R}$ et $n \in \mathbb{N}$,

$$g\left(\lambda_n x + (1-\lambda_n) y\right) \leq \lambda_n g(x) + (1-\lambda_n) g(y).$$

En passant à la limite à gauche et a droite lorsque n tends vers l'infini, et par continuité de g, on obtient finalement

$$g\left(\lambda x + (1-\lambda) y\right) \leq \lambda g(x) + (1-\lambda) g(y).$$

2. Pour montrer que g est convexe, il suffit de montrer qu'elle est continue et ensuite d'utiliser la question précédente pour obtenir la convexité. Soit $x \in \mathbb{R}$. On veut montrer que g est continue en x. Quitte à remplacer g par $t \longmapsto g(t+x) - g(x)$, on suppose sans perte de généralité que $x = 0$ et $g(x) = 0$ (le lecteur pourra vérifier que cette nouvelle fonction vérifie toujours les propriétés vérifiées par g). On a alors pour tout $y \in \mathbb{R}$

$$g\left(\frac{y+0}{2}\right) \leq \frac{g(y)+g(0)}{2} = \frac{g(y)}{2},$$

et donc par une récurrence immédiate, pour tout $n \in \mathbb{N}$, $g\left(\dfrac{y}{2^n}\right) \leq \dfrac{g(y)}{2^n}$. On a de plus pour tout $y \in \mathbb{R}$,

$$g\left(\frac{y-y}{2}\right) \leq \frac{1}{2}\left(g(y) + g(-y)\right) \ \ i.e. \ g(-y) \geq -g(y).$$

Pour montrer la continuité de g en 0, on va montrer que si $|y|$ est petit, alors $|g(y)|$ est petit. Soit M un majorant de $|g|$ sur $[-1,1]$. Soit $n \in \mathbb{N}^*$ et $y \in \left[-\frac{1}{2^n}, \frac{1}{2^n}\right]$.

→ Si $g(y) \geq 0$, alors on peut affirmer que

$$|g(y)| \leq \frac{|g(2^n y)|}{2^n} \leq \frac{M}{2^n},$$

→ Sinon, si $g(y) < 0$, alors

$$0 \leq -g(y) \leq g(-y) \leq \frac{g(-2^n y)}{2^n} \text{ et donc } |g(y)| \leq \frac{|g(-2^n y)|}{2^n} \leq \frac{M}{2^n}.$$

Ceci nous permet alors de dire que

$$\forall n \in \mathbb{N}^*,\ \forall y \in \mathbb{R},\ y \in \left[-\frac{1}{2^n}, \frac{1}{2^n}\right] \Longrightarrow |g(y)| \leq \frac{M}{2^n},$$

ce qui nous donne directement la continuité, car implique la définition de la continuité de g en 0

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ y \in]-\eta, \eta[\Longrightarrow |g(y)| < \varepsilon.$$

* *
*

Document compilé par Omar Bennouna et révisé par Issam Tauil le 02/10/2023 pour cpge-paradise.com.
Si vous repérez une erreur, ou avez des remarques, prière de me contacter via l'adresse contact@cpge-paradise.com.