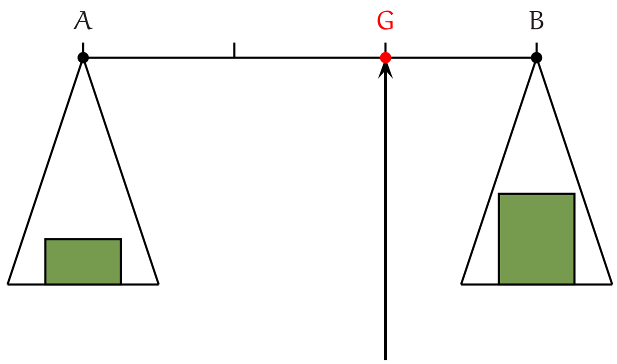
Chapitre 17. Fonctions convexes

Plan du chapitre

1 Parties convexes du plan

1.1 Barycentres

Si on place dans chacun des deux plateaux d'une balance des masses respectives de 1 kg et 2 kg, le point d'équilibre « est aux deux tiers ». On partage le segment $[A, B]$ en trois segments égaux. On laisse un segment du côté de B qui est deux fois plus lourd et deux segments du côté de A et on place le fléau de la balance au point G dessiné ci-dessous.



Le point G vérifie l'égalité $\overrightarrow{AG} = \dfrac{2}{3}\overrightarrow{AB}$ ou aussi $\overrightarrow{GA} + 2\overrightarrow{GB} = \overrightarrow{0}$. L'affixe g du point G vérifie (en notant a et b les affixes respectives des points A et B)

$$\overrightarrow{GA} + 2\overrightarrow{GB} = \overrightarrow{0} \Leftrightarrow (g - a) + 2(g - b) = 0 \Leftrightarrow g = \frac{a + 2b}{1 + 2}.$$

Cette dernière égalité s'écrit aussi $g = \dfrac{1}{3}a + \dfrac{2}{3}b$. On généralise cette situation :

Définition 1. Soient $n \in \mathbb{N}^*$ puis $A_1, \ldots, A_n$ n n points du plan. Soit $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ tel que $\lambda_1 + \ldots + \lambda_n \neq 0$. Le **barycentre du système de points pondérés** $(A_1(\lambda_1), \ldots, A_n(\lambda_n))$ est le point G d'affixe g telle que

$$g = \frac{\lambda_1 a_1 + \ldots + \lambda_n a_n}{\lambda_1 + \ldots + \lambda_n}.$$

Notation. Le barycentre du système $(A_1(\lambda_1), \ldots, A_n(\lambda_n))$ se note $\mathrm{bar}(A_1(\lambda_1), \ldots, A_n(\lambda_n))$.

Ensuite,

$$\begin{aligned} g = \frac{\lambda_1 a_1 + \ldots + \lambda_n a_n}{\lambda_1 + \ldots + \lambda_n} &\Leftrightarrow (\lambda_1 + \ldots + \lambda_n)\, g = \lambda_1 a_1 + \ldots + \lambda_n a_n \\ &\Leftrightarrow \lambda_1 (g - a_1) + \ldots + \lambda_n (g - a_n) = 0 \Leftrightarrow \lambda_1 \overrightarrow{GA_1} + \ldots + \lambda_n \overrightarrow{GA_n} = \overrightarrow{0}. \end{aligned}$$

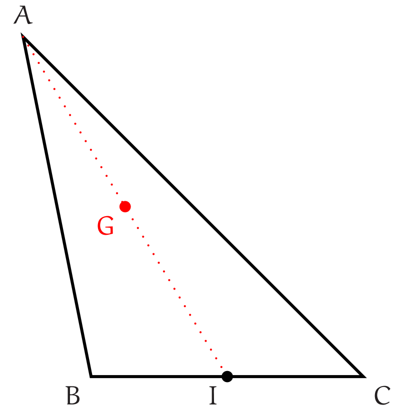
et donc

Théorème 1. G est l'unique point de E vérifiant $\lambda_1 \overrightarrow{GA_1} + \ldots + \lambda_n \overrightarrow{GA_n} = \overrightarrow{0}$.

Exemple. Soit ABC un triangle du plan. Construisons $G = \mathrm{bar}(A(2), B(1), C(1))$. Notons I le milieu de $[BC]$.

$$2\overrightarrow{GA} + \overrightarrow{GB} + \overrightarrow{GC} = \overrightarrow{0} \Rightarrow 2\overrightarrow{GA} + \overrightarrow{GI} + \overrightarrow{IB} + \overrightarrow{GI} + \overrightarrow{GI} = \overrightarrow{0} \Rightarrow 2\overrightarrow{GA} + 2\overrightarrow{GI} = \overrightarrow{0} \Rightarrow \overrightarrow{GA} + \overrightarrow{GI} = \overrightarrow{0}.$$

Le point G est donc le milieu du segment $[AI]$.



Théorème 2. Soient $n \in \mathbb{N}^*$ puis $A_1, \ldots, A_n$ n points du plan. Soit $(\lambda_1, \ldots, \lambda_n) \in \mathbb{R}^n$ tel que $\lambda_1 + \ldots + \lambda_n \neq 0$. Soit $k \in \mathbb{R} \setminus \{0\}$. Alors

$$\text{bar}\left(A_1\left(k\lambda_1\right), \ldots, A_n\left(k\lambda_n\right)\right) = \text{bar}\left(A_1\left(\lambda_1\right), \ldots, A_n\left(\lambda_n\right)\right).$$

Démonstration. $k\lambda_1 + \ldots + k\lambda_n \neq 0$ et de plus, en notant a_1, ... , a_n, les affixes respectives des points A_1, ... , A_n et g et g' les affixes respectives des points $G = \text{bar}\left(A_1\left(\lambda_1\right), \ldots, a_n\left(\lambda_n\right)\right)$ et $G' = \text{bar}\left(A_1\left(k\lambda_1\right), \ldots, A_n\left(k\lambda_n\right)\right)$,

$$g' = \frac{k\lambda_1 a_1 + \ldots + k\lambda_n a_n}{k\lambda_1 + \ldots + k\lambda_n} = \frac{\lambda_1 a_1 + \ldots + \lambda_n a_n}{\lambda_1 + \ldots + \lambda_n} = g.$$

❏

Par exemple, le milieu du segment $[AB]$ est $\text{bar}\left(A(1), B(1)\right)$ et est aussi $\text{bar}\left(A\left(\frac{1}{2}\right), B\left(\frac{1}{2}\right)\right)$. De manière plus générale, si $\lambda_1 + \ldots + \lambda_n \neq 0$,

$$\text{bar}\left(A_1\left(\lambda_1\right), \ldots, A_n\left(\lambda_n\right)\right) = \text{bar}\left(A_1\left(\lambda_1'\right), \ldots, A_n\left(\lambda_n'\right)\right),$$

où $\forall i \in [\![1, n]\!]$, $\lambda_i' = \dfrac{\lambda_i}{\lambda_1 + \ldots + \lambda_n}$. Les coefficients λ_i' vérifient

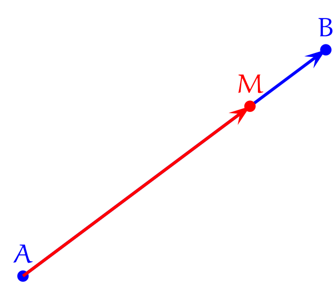
$$\sum_{i=1}^{n} \lambda_i' = \frac{1}{\lambda_1 + \ldots + \lambda_n} \sum_{i=1}^{n} \lambda_i = 1.$$

On peut donc toujours se ramener au cas où la somme des coefficients est égale à 1.

1.2 Convexes

1.2.1 Segments

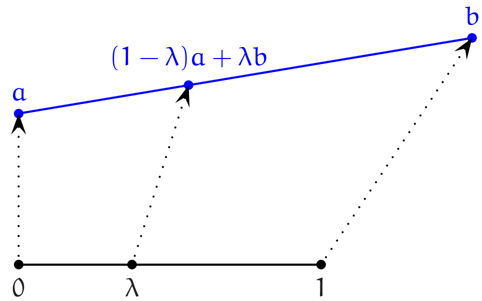
Soient A et B deux points d'un $\mathbb{K}$-espace vectoriel E. Le segment $[A, B]$ est l'ensemble des points M de E tel qu'il existe un réel $\lambda \in [0, 1]$ tel que $\overrightarrow{AM} = \lambda\overrightarrow{AB}$.



En notant m, a et b les affixes respectives des points M, A et B, l'égalité $\overrightarrow{AM} = \lambda\overrightarrow{AB}$ s'écrit encore $m - a = \lambda(b - a)$ ou enfin $m = (1 - \lambda)a + \lambda b$. Donc

Définition 2. Soient A et B deux points du plan d'affixes respectives a et b.
Le segment $[A, B]$ est l'ensemble des points du plan d'affixes $(1 - \lambda)a + \lambda b$ où $\lambda \in [0, 1]\}$ ou encore l'ensemble des barycentres à coefficients positifs des points A et B.

Commentaire. La présentation d'un segment comme ensemble des points d'affixes $(1 - \lambda)a + \lambda b$, $\lambda \in [0, 1]$, est meilleure que la présentation de ce segment comme ensemble des points d'affixe $\lambda a + (1 - \lambda)b$. Quand λ croît de 0 à 1, le point d'affixe $(1-\lambda)a+\lambda b$ parcourt le segment $[A,B]$ de A à B alors que le point d'affixe $\lambda a+(1-\lambda)b$ parcourt le segment $[A,B]$ de B à A. Notons que, lorsque $a \neq b$, l'application $\lambda \mapsto (1-\lambda)a+\lambda b$ « est » une bijection du segment $[0,1]$ de $\mathbb{R}$ sur le segment $[A,B]$ du plan.



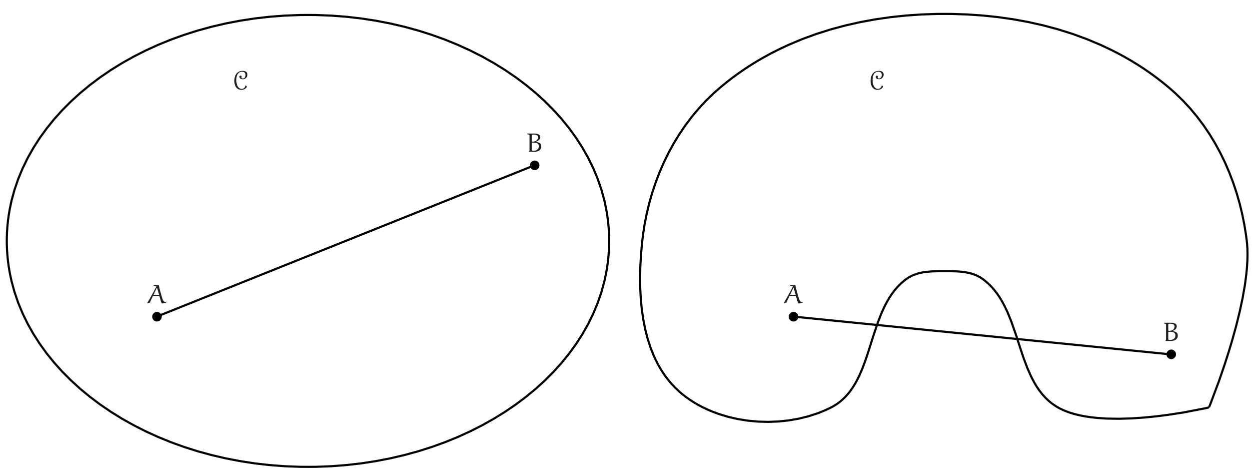
1.2.2 Parties convexes du plan

Définition 3. Soit $\mathcal{C}$ une partie non vide du plan. $\mathcal{C}$ est **convexe** si et seulement si pour tout $(A,B) \in \mathcal{C}^2$, $[A,B] \subset \mathcal{C}$.

Commentaire. Il revient au même de dire :

$$\mathcal{C} \text{ est convexe } \Leftrightarrow \forall (A,B) \in \mathcal{C}^2,\ \forall \lambda \in [0,1],\ \mathrm{bar}((A,1-\lambda),(B,\lambda)) \in \mathcal{C}.$$

Convention. $\varnothing$ est convexe.



$\mathcal{C}$ est convexe car
$\forall (A,B) \in \mathcal{C}^2$, $[A,B] \subset \mathcal{C}$

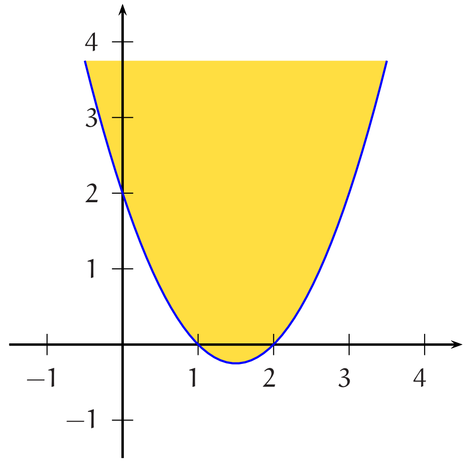
$\mathcal{C}$ n'est pas convexe car
$\exists (A,B) \in \mathcal{C}^2 /\ [A,B] \not\subset \mathcal{C}$

2 Fonctions convexes

2.1 Définition

Définition 4. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. L'**épigraphe** de f est $\{(x,y) \in I \times \mathbb{R} /\ y \geqslant f(x)\}$.

L'épigraphe de f est donc la partie du plan située au-dessus du graphe de f, bord compris.



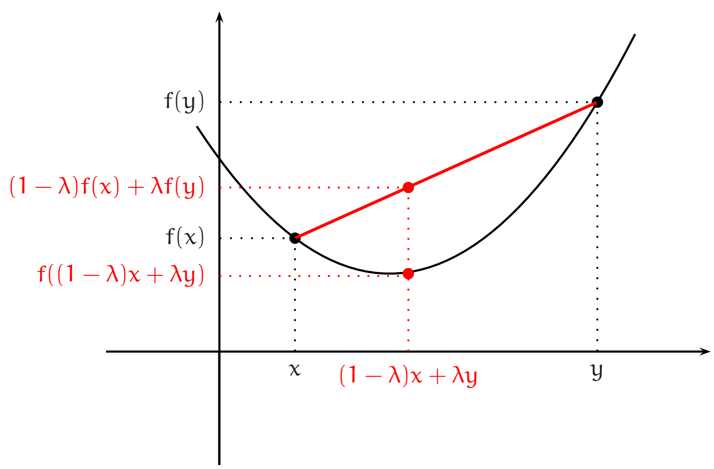
Définition 5. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

On dit que f est **convexe** sur I si et seulement si l'épigraphe de f est une partie convexe de $\mathbb{R}^2$.
On dit que f est **concave** sur I si et seulement si $-f$ est convexe sur I.

Théorème 3. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

f est convexe sur I si et seulement si $\forall (x, y) \in I^2$, $\forall \lambda \in [0, 1]$, $f((1-\lambda)x + \lambda y) \leqslant (1-\lambda)f(x) + \lambda f(y)$.
f est concave sur I si et seulement si $\forall (x, y) \in I^2$, $\forall \lambda \in [0, 1]$, $f((1-\lambda)x + \lambda y) \geqslant (1-\lambda)f(x) + \lambda f(y)$.

Voici un exemple de graphe de fonction convexe :



Démonstration .

• Supposons que f est convexe sur I et donc que l'épigraphe $\mathcal{E}$ de f est un convexe de $\mathbb{R}^2$.

Soient $(x_1, x_2) \in I^2$ et $\lambda \in [0, 1]$. Les points $A_1 = (x_1, f(x_1))$ et $A_2 = (x_2, f(x_2))$ sont des points de $\mathcal{E}$. Donc, le segment $[A_1, A_2]$ est contenu dans $\mathcal{E}$. En particulier, le point d'affixe $(1-\lambda)a_1 + \lambda a_2$ (où a_1 et a_2 sont les affixes respectives des points A_1 et A_2) est un point de $\mathcal{E}$. L'ordonnée de ce point, à savoir $(1-\lambda)f(x_1) + \lambda f(x_2)$, est supérieure ou égale à l'image de son abscisse, à savoir $f((1-\lambda)x_1 + \lambda x_2)$.
On a montré que $\forall (x_1, x_2) \in I^2$, $\forall \lambda \in [0, 1]$ $f((1-\lambda)x_1 + \lambda x_2) \leqslant (1-\lambda)f(x_1) + \lambda f(x_2)$. Donc, f est convexe sur I.

• Supposons que $\forall (x_1, x_2) \in I^2$, $\forall \lambda \in [0, 1]$, $f((1-\lambda)x_1 + \lambda x_2) \leqslant (1-\lambda)f(x_1) + \lambda f(x_2)$. Montrons que $\mathcal{E}$ est un convexe de $\mathbb{R}^2$.

Soient $A_1 = (x_1, y_1)$ et $A_2 = (x_2, y_2)$ deux points de $\mathcal{E}$ et $\lambda \in [0, 1]$. On note a_1 et a_2 les affixes respectives des points A_1 et A_2. Soit M le point d'affixe $(1-\lambda)a_1 + \lambda a_2$. Posons $M = (x, y)$. Alors

$$y = (1-\lambda)y_1 + \lambda y_2 \geqslant (1-\lambda)f(x_1) + \lambda f(x_2) \geqslant f((1-\lambda)x_1 + \lambda x_2) = f(x),$$

et donc M est un point de $\mathcal{E}$. Ainsi, pour tout $(A_1, A_2) \in \mathcal{E}^2$ et tout point M de $[A_1, A_2]$, M est un point de $\mathcal{E}$. On a montré que $\mathcal{E}$ est un convexe de $\mathbb{R}^2$.

Enfin, f concave sur $I \Leftrightarrow -f$ convexe sur $I \Leftrightarrow \forall (x, y) \in I^2$, $\forall \lambda \in [0, 1]$, $-f((1-\lambda)x + \lambda y) \leqslant (1-\lambda)(-f(x)) + \lambda(-f(y)) \Leftrightarrow \forall (x, y) \in I^2$, $\forall \lambda \in [0, 1]$, $f((1-\lambda)x + \lambda y) \geqslant (1-\lambda)f(x) + \lambda f(y)$.

❏

Dorénavant, les équivalences du théorème 3 devient les définitions d'une fonction convexe ou concave sur un intervalle :

$$f \text{ est convexe sur } I \Leftrightarrow \forall (x,y) \in I^2,\ \forall \lambda \in [0,1],\ f((1-\lambda)x + \lambda y) \leqslant (1-\lambda)f(x) + \lambda f(y).$$
$$f \text{ est concave sur } I \Leftrightarrow \forall (x,y) \in I^2,\ \forall \lambda \in [0,1],\ f((1-\lambda)x + \lambda y) \geqslant (1-\lambda)f(x) + \lambda f(y).$$

Il existe des fonctions qui sont à la fois convexes et concaves : les fonctions affines. On peut montrer que ce sont les seules fonctions à être à la fois convexes et concaves :

Exercice 1. Soit f une fonction de $\mathbb{R}$ dans $\mathbb{R}$ telle que :

$$\forall (x,y) \in \mathbb{R}^2,\ \forall \lambda \in [0,1],\ f((1-\lambda)x + \lambda y) = (1-\lambda)f(x) + \lambda f(y).$$

Montrer que f est une fonction affine. Etablir la réciproque.

Solution 1. Montrons que pour tout réel x, $f(x) = ax + b$ où on a posé $a = \dfrac{f(1) - f(0)}{1 - 0} = f(1) - f(0)$ et $b = f(0)$. Soit x un réel.

1er cas. Si $0 \leqslant x \leqslant 1$, alors $x \in [0,1]$. Posons $\lambda = x$ ou encore $x = (1-\lambda) \times 0 + \lambda \times 1$ avec $\lambda \in [0,1]$. Par hypothèse

$$f(x) = f((1-\lambda) \times 0 + \lambda \times 1) = (1-\lambda)f(0) + \lambda f(1) = \lambda(f(1) - f(0)) + f(0) = ax + b.$$

2ème cas. Si $x > 1$, alors $1 \in [0,x]$. Dans ce cas, il existe $\lambda \in [0,1]$ tel que $1 = (1-\lambda) \times 0 + \lambda x$ (à savoir $\lambda = \dfrac{1}{x}$ et donc aussi $1 - \lambda = \dfrac{x-1}{x}$). Par hypothèse,

$$f(1) = f((1-\lambda) \times 0 + \lambda x) = (1-\lambda)f(0) + \lambda f(x) = \frac{x-1}{x}f(0) + \frac{f(x)}{x}$$

puis $f(x) = xf(1) - (x-1)f(0) = (f(1) - f(0))x + f(0) = ax + b$.

3ème cas. Si $x < 0$, alors $0 \in [x,1]$. Dans ce cas, il existe $\lambda \in [0,1]$ tel que $0 = (1-\lambda) \times x + \lambda \times 1$ (à savoir $\lambda = \dfrac{x}{x-1}$ et donc aussi $1 - \lambda = \dfrac{1}{1-x}$). Par hypothèse,

$$f(0) = f((1-\lambda)x + \lambda \times 1) = (1-\lambda)f(x) + \lambda f(1) = \frac{1}{1-x}f(x) - \frac{x}{1-x}f(1)$$

puis $f(x) = (1-x)f(0) + xf(1) = (f(1) - f(0))x + f(0) = ax + b$.

On a monté qu'il existe deux réels a et b tels que, pour tout réel x, $f(x) = ax + b$. Donc, f est une fonction affine.

Inversement, soit f une fonction affine. Il existe deux réels a et b tels que, pour tout réel x, $f(x) = ax + b$. Soient alors $(x,y) \in \mathbb{R}^2$ et $\lambda \in [0,1]$,

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

Les fonctions affines sont les fonctions à la fois convexes et concaves.

Théorème 4. (Inégalité de Jensen)

Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

f est convexe sur I si et seulement si

$$\forall n \geqslant 2,\ \forall (x_i)_{1 \leqslant i \leqslant n} \in I^n,\ \forall (\lambda_i)_{1 \leqslant i \leqslant n} \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow f\left(\sum_{i=1}^{n} \lambda_i x_i \right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) \right).$$

Démonstration . $\Leftarrow$ / Supposons que $\forall n \geqslant 2,\ \forall (x_i)_{1 \leqslant i \leqslant n} \in I^n,\ \forall (\lambda_i)_{1 \leqslant i \leqslant n} \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow f\left(\sum_{i=1}^{n} \lambda_i x_i \right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) \right)$.
Quand $n = 2$, on obtient la définition d'une fonction convexe (en remplaçant x_1 et x_2 par x et y et en remplaçant λ_1 et λ_2 par λ et $1 - \lambda$ car $\lambda_1 + \lambda_2 = 1$).

$\Rightarrow$ / Supposons f convexe sur I. Montrons par récurrence que

$$\forall n \geqslant 2,\ \forall (x_i)_{1 \leqslant i \leqslant n} \in I^n,\ \forall (\lambda_i)_{1 \leqslant i \leqslant n} \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow f\left(\sum_{i=1}^{n} \lambda_i x_i \right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) \right) \quad (\mathcal{P}_n).$$

- $(\mathcal{P}_2)$ est vraie par définition d'une fonction convexe.

• Soit $n \geqslant 2$. Supposons $(\mathcal{P}_n)$.

Soient $(x_1, \ldots, x_n, x_{n+1}) \in I^{n+1}$ et $(\lambda_1, \ldots, \lambda_n, \lambda_{n+1}) \in [0,1]^{n+1}$ tel que $\displaystyle\sum_{i=1}^{n+1} \lambda_i = 1$.

1er cas. Si $\lambda_{n+1} = 1$, alors $\displaystyle\sum_{i=1}^{n} \lambda_i = 0$ puis, $\forall i \in [\![1, n]\!]$, $\lambda_i = 0$ (réels positifs de somme nulle). Dans ce cas, l'inégalité est immédiate (c'est une égalité).

2ème cas. Sinon $\lambda_{n+1} \in [0, 1[$ et en particulier, $\displaystyle\sum_{i=1}^{n} \lambda_i = 1 - \lambda_{n+1} > 0$. On commence par constater que

$$f(\lambda_1 x_1 + \ldots + \lambda_n x_n + \lambda_{n+1} x_{n+1}) = f\left((1 - \lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_{n+1}} x_i + \lambda_{n+1} x_{n+1}\right).$$

Pour $i \in [\![1, n]\!]$, posons $\lambda'_i = \dfrac{\lambda_i}{1 - \lambda_{n+1}}$. Tout d'abord, pour $i \in [\![1, n]\!]$, $0 \leqslant \lambda'_i = \dfrac{\lambda_i}{1 - \lambda_{n+1}} \leqslant \dfrac{\lambda_1 + \ldots + \lambda_n}{1 - \lambda_{n+1}} = 1$.

De plus, $\displaystyle\sum_{i=1}^{n} \lambda'_i = \frac{\displaystyle\sum_{i=1}^{n} \lambda_i}{1 - \lambda_{n+1}} = 1$.

Ensuite, en supposant avoir numéroté les réels $x_1, \ldots, x_n$ de telle sorte que par exemple $x_1 \leqslant x_2 \leqslant \ldots \leqslant x_n$, on a

$$\begin{aligned} x_1 &= x_1 \sum_{i=1}^{n} \lambda'_i = \sum_{i=1}^{n} \lambda'_i x_1 \\ &\leqslant \sum_{i=1}^{n} \lambda'_i x_i \\ &\leqslant \sum_{i=1}^{n} \lambda'_i x_n = x_n \sum_{i=1}^{n} \lambda'_i = x_n. \end{aligned}$$

Donc, $x_1 \leqslant \displaystyle\sum_{i=1}^{n} \lambda'_i x_i \leqslant x_n$ ou encore le réel $x = \displaystyle\sum_{i=1}^{n} \lambda'_i x_i$ est dans $[x_1, x_n]$. Puisque x_1 et x_n sont dans I et que I est un intervalle, on en déduit que $[x_1, x_n] \subset I$. Donc, le réel $x = \displaystyle\sum_{i=1}^{n} \lambda'_i x_i$ est un réel de I. Mais alors,

$$\begin{aligned} f\left((1 - \lambda_{n+1}) \sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_{n+1}} x_i + \lambda_{n+1} x_{n+1}\right) &= f((1 - \lambda_{n+1})\, x + \lambda_{n+1} x_{n+1}) \\ &\leqslant (1 - \lambda_{n+1})\, f(x) + \lambda_{n+1} f(x_{n+1}) \\ &\text{(puisque f est convexe sur I et que } x \text{ et } x_{n+1} \text{ sont dans I)} \\ &= (1 - \lambda_{n+1})\, f\left(\sum_{i=1}^{n} \lambda'_i x_i\right) + \lambda_{n+1} f(x_{n+1}) \\ &\leqslant (1 - \lambda_{n+1}) \sum_{i=1}^{n} \lambda'_i f(x_i) + \lambda_{n+1} f(x_{n+1}) \\ &\text{(par hypothèse de récurrence en tenant compte entre autre de } \sum_{i=1}^{n} \lambda'_i = 1\text{)} \\ &= \sum_{i=1}^{n} \lambda_i f(x_i) + \lambda_{n+1} f(x_{n+1}) = \sum_{i=1}^{n+1} \lambda_i f(x_i) \end{aligned}$$

Le résultat est démontré par récurrence.

❏

On mettra en œuvre plus loin l'inégalité de Jensen dans l'exercice sur la comparaison entre la moyenne arithmétique et la moyenne géométrique de n réels positifs.

On peut affiner la définition d'une fonction convexe :

DÉFINITION 6. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

On dit que f est **strictement convexe** sur I si et seulement si si et seulement si $\forall (x,y) \in I^2$ tel que $x \neq y$, $\forall \lambda \in]0,1[$, $f((1-\lambda)x + \lambda y) < (1-\lambda)f(x) + \lambda f(y)$.
On dit que f est **strictement convave** sur I si et seulement si si et seulement si $\forall (x,y) \in I^2$ tel que $x \neq y$, $\forall \lambda \in]0,1[$, $f((1-\lambda)x + \lambda y) > (1-\lambda)f(x) + \lambda f(y)$.

2.2 Caractérisation par la « fonction pente »

Théorème 5. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

f est convexe sur I si et seulement si

$$\forall x_0 \in I, \text{ « la fonction pente en } x_0 \text{ » } \quad \varphi_{x_0} : \begin{array}{ccc} I \setminus \{x_0\} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x) - f(x_0)}{x - x_0} \end{array} \quad \text{est croissante sur } I \setminus \{x_0\}.$$

f est strictement convexe sur I si et seulement si la fonction pente en tout x_0 de I est strictement croissante sur I.

On a un théorème analogue pour les fonctions concaves en remplaçant le mot « croissante » par « décroissante ».

DÉMONSTRATION .
• Supposons que pour tout x_0 de I, la fonction φ_{x_0} est croissante sur $I \setminus \{x_0\}$.

Soit $(x_1, x_2) \in I^2$ tel que $x_1 < x_2$ et $\lambda \in]0,1[$. On a donc $x_1 < (1-\lambda)x_1 + \lambda x_2 < x_2$. La fonction φ_{x_1} est croissante sur $I \setminus \{x_1\}$. Donc, $\varphi_{x_1}((1-\lambda)x_1 + \lambda x_2) \leqslant \varphi_{x_1}(x_2)$. Cette inégalité s'écrit explicitement

$$\frac{f((1-\lambda)x_1 + \lambda x_2) - f(x_1)}{((1-\lambda)x_1 + \lambda x_2) - x_1} \leqslant \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

ou encore $\dfrac{f((1-\lambda)x_1 + \lambda x_2) - f(x_1)}{\lambda(x_2 - x_1)} \leqslant \dfrac{f(x_2) - f(x_1)}{x_2 - x_1}$ puis $f((1-\lambda)x_1 + \lambda x_2) - f(x_1) \leqslant \lambda(f(x_2) - f(x_1))$
(car $x_2 - x_1 > 0$) et finalement

$$f((1-\lambda)x_1 + \lambda x_2) \leqslant (1-\lambda)f(x_1) + \lambda f(x_2).$$

Cette inégalité reste claire quand $\lambda = 0$ ou $\lambda = 1$ ou $x_1 = x_2$ et on a donc montré que f est convexe. Enfin, si pour tout x_0 de I, la fonction φ_{x_0} est strictement croissante sur $I \setminus \{x_0\}$, on remplace ci-dessus les inégalités larges par des inégalités strictes et on obtient le fait que f est strictement convexe sur I.

• Supposons f convexe sur I. Soit $x_0 \in I$. On supposera dans ce qui suit que x_0 n'est pas une borne de I, la démonstration s'adaptant facilement dans le cas contraire. Soit $(x_1, x_2) \in I^2$ tel que $x_1 < x_2$. Trois cas de figure se présentent :

1er cas. Supposons $x_0 < x_1 < x_2$.
Soit $\lambda = \dfrac{x_1 - x_0}{x_2 - x_0}$. λ est un réel de $]0,1[$ tel que $x_1 = (1-\lambda)x_0 + \lambda x_2$. Puisque f est convexe sur I,

$$f(x_1) = f((1-\lambda)x_0 + \lambda x_2) \leqslant (1-\lambda)f(x_0) + \lambda f(x_2) = \frac{x_2 - x_1}{x_2 - x_0} f(x_0) + \frac{x_1 - x_0}{x_2 - x_0} f(x_2)$$

puis

$$f(x_1) - f(x_0) \leqslant \frac{-(x_1 - x_0)}{x_2 - x_0} f(x_0) + \frac{x_1 - x_0}{x_2 - x_0} f(x_2) = \frac{x_1 - x_0}{x_2 - x_0}(f(x_2) - f(x_0))$$

et donc $\varphi_{x_0}(x_1) \dfrac{f(x_1) - f(x_0)}{x_1 - x_0} \leqslant \dfrac{f(x_2) - f(x_0)}{x_2 - x_0} = \varphi_{x_0}(x_2)$ après division des deux membres de l'inégalité par $x_1 - x_0 > 0$.

2ème cas. Supposons $x_1 < x_2 < x_0$.
Soit $\lambda = \dfrac{x_2 - x_1}{x_0 - x_1}$. λ est un réel de $]0,1[$ tel que $x_2 = (1-\lambda)x_1 + \lambda x_0$. Puisque f est convexe sur I,

$$f(x_2) \leqslant (1-\lambda)f(x_1) + \lambda f(x_0) = \frac{x_0 - x_2}{x_0 - x_1} f(x_1) + \frac{x_2 - x_1}{x_0 - x_1} f(x_0)$$

puis

$$f(x_2) - f(x_0) \leqslant \frac{x_0 - x_2}{x_0 - x_1} f(x_1) + \frac{x_2 - x_0}{x_0 - x_1} f(x_0) = \frac{x_2 - x_0}{x_1 - x_0}(f(x_1) - f(x_0))$$

et donc encore $\varphi_{x_0}(x_1) \leqslant \varphi_{x_0}(x_2)$ après division des deux membres de l'inégalité par $x_2 - x_0 < 0$.

3ème cas. Supposons $x_1 < x_0 < x_2$. D'après les deux premiers cas,

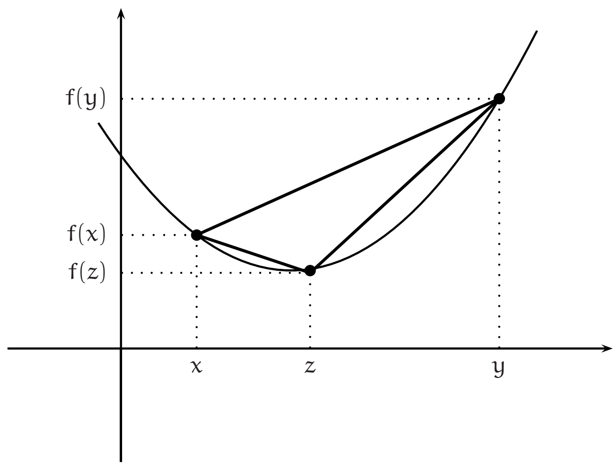
$$\frac{f(x_1) - f(x_0)}{x_1 - x_0} = \varphi_{x_0}(x_1) = \varphi_{x_1}(x_0) \leqslant \varphi_{x_1}(x_2) = \varphi_{x_2}(x_1) \leqslant \varphi_{x_2}(x_0) = \frac{f(x_2) - f(x_0)}{x_2 - x_0}.$$

Ceci montre que la fonction pente en x_0 est croissante sur $I\setminus\{x_0\}$. D'autre part, si f est strictement convexe sur I, on remplace toutes les inégalités larges précédentes par des inégalités strictes et on obtient le fait que la fonction pente en x_0 est strictement croissante sur $I\setminus\{x_0\}$.

❑

Notons au passage cette configuration usuelle concernant les fonctions convexes : si $x < z < y$,

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



2.3 Caractérisation des fonctions convexes dérivables

Théorème 6. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. On suppose de plus que f est dérivable sur I.

f est convexe sur I si et seulement si f' est une fonction croissante sur I.
f est concave sur I si et seulement si f' est une fonction décroissante sur I.

Démonstration .

• Supposons f' croissante sur I. Soit $(x,y)\in I^2$ tel que $x<y$.
Pour $\lambda\in[0,1]$, posons $g(\lambda)=(1-\lambda)f(x)+\lambda f(y)-f((1-\lambda)x+\lambda y)$. La fonction $\lambda\mapsto(1-\lambda)x+\lambda y$ est dérivable sur $[0,1]$ à valeurs dans $[x,y]\subset I$ et la fonction f est dérivable sur I. Donc la fonction $\lambda\mapsto f((1-\lambda)x+\lambda y)$ est dérivable sur $[0,1]$ et il en est de même de la fonction g. De plus, pour $\lambda\in[0,1]$,

$$g'(\lambda)=f(y)-f(x)-(y-x)f'((1-\lambda)x+\lambda y).$$

D'après le théorème des accroissements finis (puisque f est continue sur $[x,y]\subset I$ et dérivable sur $]x,y[\subset I$), il existe un réel $c\in]x,y[$ tel que $f(y)-f(x)=(y-x)f'(c)$ ou encore, il existe un réel $\lambda_0\in]0,1[$, à savoir $\lambda_0=\dfrac{c-x}{y-x}$, tel que $f(y)-f(x)=(y-x)f'\left((1-\lambda_0)x+\lambda_0 y\right)$.
Donc,

$$\forall\lambda\in[0,1],\ g'(\lambda)=(y-x)\left[f'\left((1-\lambda_0)x+\lambda_0 y\right)-f'\left((1-\lambda)x+\lambda y\right)\right].$$

La fonction affine $\lambda\mapsto(1-\lambda)x+\lambda y=\lambda(y-x)-x$ est croissante sur $[0,1]$ (car $x<y$) et donc la fonction $\lambda\mapsto f'\left((1-\lambda)x+\lambda y\right)$ est croissante sur $[0,1]$ puis la fonction g' est décroissante sur $[0,1]$. Puisque $g'(\lambda_0)=0$, on en déduit que g' est positive sur $[0,\lambda_0]$ et négative sur $[\lambda_0,1]$.
Ainsi, la fonction g est croissante sur $[0,\lambda_0]$ et décroissante sur $[\lambda_0,1]$. Puisque $g(0)=g(1)=0$, la fonction g est positive sur $[0,1]$ ou encore

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

Donc, la fonction f est convexe sur I.

• Supposons f convexe sur I. Soit $(x,y)\in I^2$ tel que $x<y$. Puisque la fonction $\varphi_x\ :\ t\mapsto\dfrac{f(t)-f(x)}{t-x}$ est croissante sur $I\setminus\{x\}$ (d'après le théorème 5), pour tout $t\in I\cap]x,+\infty[$, $\varphi_x(t)\geqslant\lim\limits_{\substack{u\to x\\ u>x}}\varphi_x(u)=f'(x)$ et en particulier, $\dfrac{f(y)-f(x)}{y-x}\geqslant f'(x)$.
De même, la fonction $\varphi_y\ :\ t\mapsto\dfrac{f(t)-f(y)}{t-y}$ est croissante sur $I\setminus\{y\}$ et donc $\varphi_y(x)\leqslant\lim\limits_{\substack{u\to y\\ u<y}}\varphi_y(u)=f'(y)$ et en particulier, $\dfrac{f(y)-f(x)}{y-x}\leqslant f'(y)$.

On a donc $f'(x) \leqslant \dfrac{f(y)-f(x)}{y-x} \leqslant f'(y)$ et en particulier $f'(x) \leqslant f'(y)$. On a montré que f' est croissante sur I.

❏

On démontre de manière analogue le théorème suivant :

Théorème 7. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. On suppose de plus que f est dérivable sur I.

f est strictement convexe sur I si et seulement si f' est une fonction strictement croissante sur I.
f est strictement concave sur I si et seulement si f' est une fonction strictement décroissante sur I.

et on en déduit :

Théorème 8. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. On suppose de plus que f est deux fois dérivable sur I.

f est convexe sur I si et seulement si $f'' \geqslant 0$.
f est concave sur I si et seulement si $f'' \leqslant 0$.

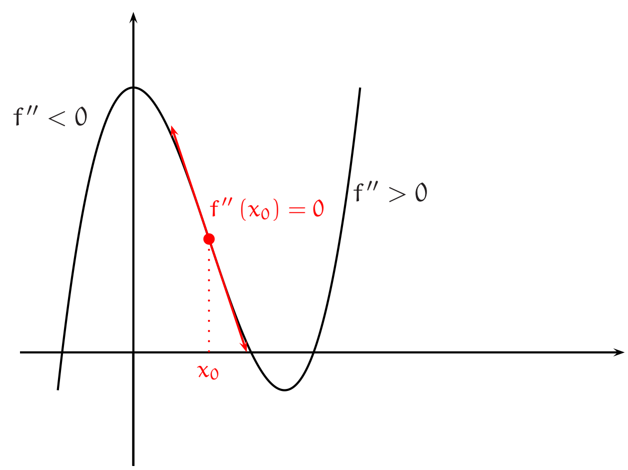
Si $f'' > 0$ sur I sauf peut-être en un nombre fini de points, alors f est strictement convexe sur I.
Si $f'' < 0$ sur I sauf peut-être en un nombre fini de points, alors f est strictement concave sur I.

Démonstration. On sait que f' est croissante sur I si et seulement si $(f')' = f''$ est positive sur I.

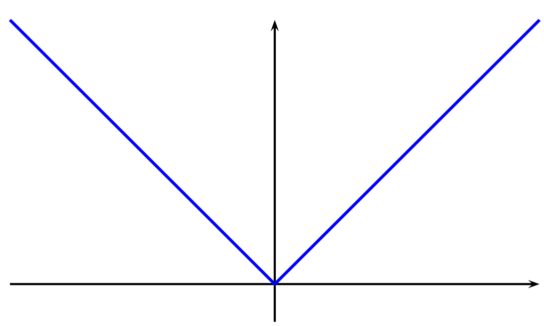
Définition 7. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$, deux fois dérivable sur I.

Si x_0 est un réel de I en lequel f'' s'annule en changeant de signe, on dit que le point $(x_0, f(x_0))$ est un **point d'inflexion** de la courbe représentative de f.

Un point d'inflexion est un point de la courbe en lequel la concavité change de sens. On verra aussi qu'en un point d'inflexion, la courbe « traverse sa tangente ».



Dans les théorèmes de ce paragraphe, nous avons rajouté des hypothèses du genre « on suppose f dérivable ou deux fois dérivables sur I ». Le fait que la fonction f soit convexe ou concave n'entraine absolument pas que la fonction f soit dérivable sur I. Par exemple, la fonction $x \mapsto |x|$ est convexe sur $\mathbb{R}$ mais n'est pas dérivable sur $\mathbb{R}$.



Néanmoins,

Exercice 2. Soit f une fonction convexe sur un intervalle **ouvert** I de $\mathbb{R}$.

1) Montrer que f est dérivable à droite et à gauche en tout réel de I.

2) Montrer que f est continue sur I.

Solution 2.

1) Soit x_0 un réel de I. Puisque I est un intervalle ouvert, $I \cap]-\infty, x_0[$ et $I \cap]x_0, +\infty[$ sont des intervalles ouverts non vides. Puisque f est convexe sur I, la fonction pente en x_0 (à savoir $\varphi_{x_0} : x \mapsto \dfrac{f(x) - f(x_0)}{x - x_0}$) est croissante sur $I \setminus \{x_0\}$. On en déduit que cette fonction admet des limites réelles à droite et à gauche en x_0. f est donc dérivable à droite et à gauche en x_0.

2) Soit x_0 un réel de I. D'après 1), f est dérivable à droite et à gauche en x_0 et en particulier, f est continue à droite et à gauche en x_0. Mais alors f est continue en x_0. Ainsi, f est continue en chaque réel x_0 de I et donc f est continue sur I.

2.4 Tangentes au graphe d'une fonction convexe

Soit f une fonction dérivable et convexe sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. Soit a un réel. On va montrer que la courbe représentative $\mathcal{C}_f$ de f est au-dessus de sa tangente (T_a) en son point d'abscisse a sur I. Une équation de la tangente à $\mathcal{C}_f$ en le point $(a, f(a))$ est

$$y = f'(a)(x - a) + f(a).$$

Pour $x \in I$, posons $g(x) = f(x) - (f'(a)(x - a) + f(a))$. g est dérivable sur I et pour tout x de I, $g'(x) = f'(x) - f'(a)$. Puisque f est convexe sur I, f' est croissante sur I. Puisque $g'(a) = 0$, on en déduit que g' est négative sur $I \cap]-\infty, a]$ et positive sur $I \cap [a, +\infty[$. La fonction g admet donc un minimum en a égal à $g(a) = 0$. Ceci montre que la fonction g est positive sur I et donc que

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

On a montré que $\mathcal{C}_f$ est au-dessus de (T_a) sur I.

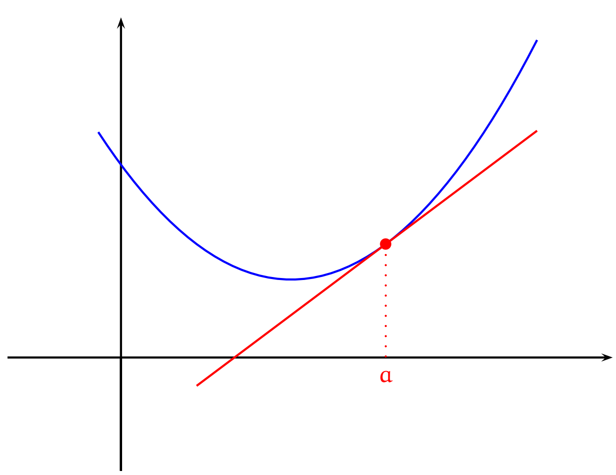
Commentaire 1. Si de plus, f est de classe C^1 sur I, on peut aussi constater que pour $x \in I$,

$$f(x) - (f'(a)(x - a) + f(a)) = (f(x) - f(a)) - f'(a)(x - a) = \int_a^x (f'(t) - f'(a))\ dt$$

et vérifier que cette intégrale est positive en discutant suivant le fait que $x \geqslant a$ ou $x \leqslant a$.

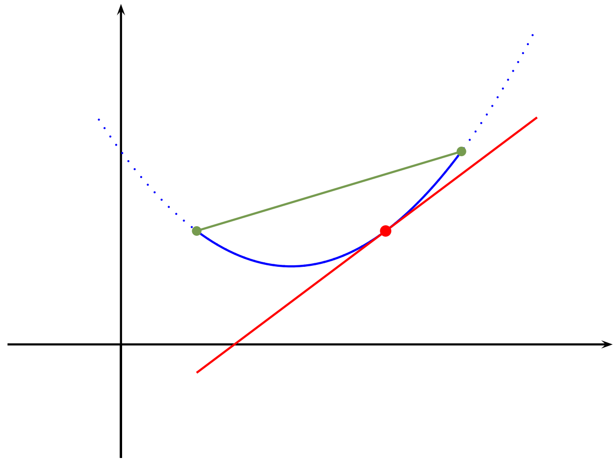
Commentaire 2. Dans le cas où f est strictement convexe, en adaptant la démonstration ci-dessus, on montre que $\mathcal{C}_f$ est strictement au-dessus de (T_a) sur $I \setminus \{a\}$.

Commentaire 3. Enfin, si f est concave sur I (resp. strictement concave sur I), $\mathcal{C}_f$ est au-dessous de (T_a) sur I (resp. strictement au-dessous de (T_a) sur $I \setminus \{a\}$).



2.5 Inégalités de convexité

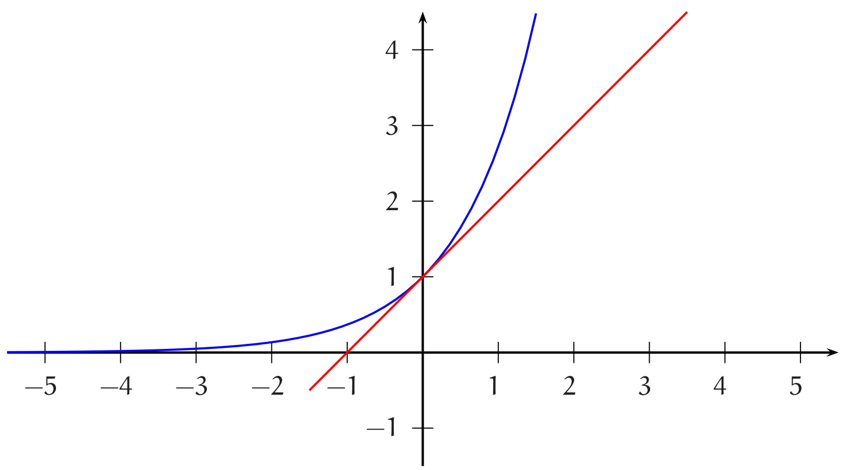
De ce qui précède, on a l'habitude de déduire que « le graphe d'une fonction convexe est au-dessus de ses tangentes et au-dessous de ses cordes » (étant entendu que le graphe est au-dessus de sa tangente sur I tout entier et au-dessous de la corde joignant les points $(a, f(a))$ et $(b, f(b))$ sur le segment $[a, b]$ uniquement).



Cette constatation explicitement utilisée fournit des inégalités appelées « inégalités de convexité ». On donne ci-dessous un certain nombre d'inégalités de convexité classique à connaître.

• La fonction exponentielle est strictement convexe sur $\mathbb{R}$ car sa dérivée seconde, à savoir $x \mapsto e^x$, est strictement positive sur $\mathbb{R}$. Son graphe est donc au-dessus de sa tangente en $(0, e^0) = (0, 1)$ sur $\mathbb{R}$ et strictement au-dessus sur $\mathbb{R}^*$. Ceci fournit

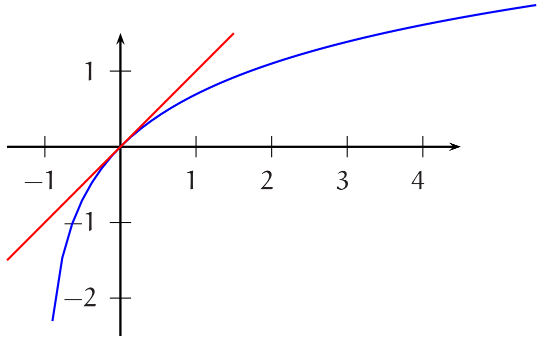
$$\boxed{\forall x \in \mathbb{R},\ e^x \geqslant 1 + x \text{ et } \forall x \in \mathbb{R}^*,\ e^x > 1 + x.}$$



En constatant que le graphe de la fonction exponentielle est au-dessus de sa tangente au point $(1, e)$, on obtient aussi : $\forall x \in \mathbb{R}$, $e^x \geqslant ex$.

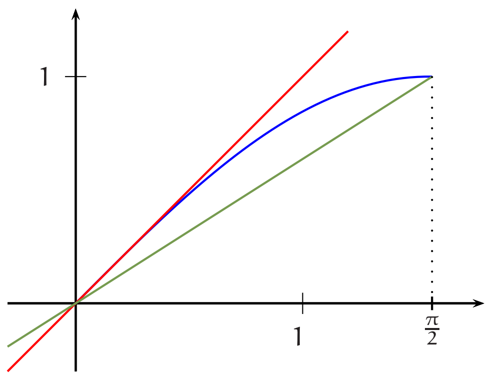
• La fonction $x \mapsto \ln(1 + x)$ est strictement concave sur $] - 1, +\infty[$ car sa dérivée seconde, à savoir $x \mapsto -\dfrac{1}{(1 + x)^2}$, est strictement négative sur $] - 1, +\infty[$. Son graphe est donc au-dessous de sa tangente en $(0, \ln 1) = (0, 0)$ sur $] - 1, +\infty[$ et strictement au-dessous sur $] - 1, 0[\cup]0, +\infty[$. Ceci fournit

$$\boxed{\forall x \in] - 1, +\infty[,\ \ln(1 + x) \leqslant x \text{ et } \forall x \in] - 1, 0[\cup]0, +\infty[,\ \ln(1 + x) < x.}$$



• La fonction $x \mapsto \sin x$ est strictement concave sur $\left[0, \frac{\pi}{2}\right]$ car sa dérivée seconde à savoir $x \mapsto -\sin x$, est strictement négative sur $\left]0, \frac{\pi}{2}\right]$. Son graphe est donc au-dessous de sa tangente en $(0, \sin 0) = (0, 0)$ sur $\left[0, \frac{\pi}{2}\right]$ et strictement au-dessous sur $\left]0, \frac{\pi}{2}\right]$ et son graphe est au-dessus de sa corde joignant les points $(0, 0)$ et $\left(\frac{\pi}{2}, 1\right)$ sur $\left[0, \frac{\pi}{2}\right]$. Ceci fournit entre autre

$$\boxed{\forall x \in \left[0, \frac{\pi}{2}\right],\ \frac{2}{\pi}x \leqslant \sin x \leqslant x.}$$



Exercice 3. (comparaison entre la moyenne géométrique et la moyenne arithmétique de n réels positifs)

Soient $n \geqslant 2$ puis $(x_1, \ldots, x_n) \in [0, +\infty[^n$. Montrer que

$$\forall n \geqslant 2,\ \forall (x_1, \ldots, x_n) \in [0, +\infty[^n,\ \sqrt[n]{x_1 \ldots x_n} \leqslant \frac{x_1 + \ldots + x_n}{n}.$$

Solution 3. Soit $(x_1, \ldots, x_n) \in [0, +\infty[^n$.

Si il existe $i \in [\![1, n]\!]$ tel que $x_i = 0$, l'inégalité est immédiate car dans ce cas, $\sqrt[n]{x_1 \ldots x_n} = 0$.

Supposons donc : $\forall i \in [\![1, n]\!]$, $x_i > 0$. La fonction $\ln$ est concave sur $]0, +\infty[$ car sa dérivée seconde, à savoir $x \mapsto -\frac{1}{x^2}$, est négative sur $]0, +\infty[$. Les n nombres $\lambda_1 = \frac{1}{n}$, $\lambda_2 = \frac{1}{n}$, ..., $\lambda_n = \frac{1}{n}$, sont n réels positifs de somme 1. D'après l'inégalité de Jensen

$$\frac{1}{n}\ln(x_1) + \ldots + \frac{1}{n}\ln(x_n) \leqslant \ln\left(\frac{1}{n}x_1 + \ldots + \frac{1}{n}x_n\right)$$

ce qui s'écrit encore

$$\ln\left(\sqrt[n]{x_1 \ldots x_n}\right) \leqslant \ln\left(\frac{x_1 + \ldots + x_n}{n}\right)$$

et finalement

$$\sqrt[n]{x_1 \ldots x_n} \leqslant \frac{x_1 + \ldots + x_n}{n}.$$

$$\boxed{\forall n \geqslant 2,\ \forall (x_1, \ldots, x_n) \in [0, +\infty[^n,\ \sqrt[n]{x_1 \ldots x_n} \leqslant \frac{x_1 + \ldots + x_n}{n}.}$$

La moyenne géométrique de n nombres est toujours inférieure ou égale à la moyenne arithmétique de ces n nombres.

En particulier, quand $n = 2$, on obtient : $\forall (a, b) \in [0, +\infty[^2$, $\sqrt{ab} \leqslant \dfrac{a+b}{2}$. On rappelle que cette inégalité peut s'obtenir directement par un calcul algébrique sans passer par l'analyse (dérivées, fonction logarithme, ...) :

$$\frac{a+b}{2} - \sqrt{ab} = \frac{1}{2}\left(a - 2\sqrt{a}\sqrt{b} + b\right) = \frac{1}{2}\left(\sqrt{a} - \sqrt{b}\right)^2 \geqslant 0.$$