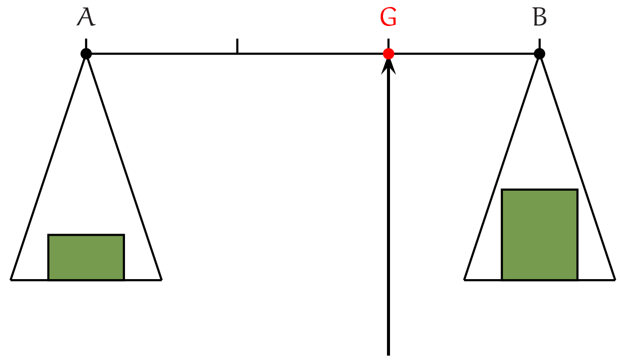
Convexité

Plan du chapitre

I - Parties convexes d'un espace vectoriel

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 $\overrightarrow{AG} = \frac{2}{3}\overrightarrow{AB}$ ou aussi $\overrightarrow{GA} + 2\overrightarrow{GB} = \overrightarrow{0}$. On rappelle alors une notation donnée en maths sup. Si $\overrightarrow{u} = \overrightarrow{AB}$, on peut écrire $\overrightarrow{u} = B - A$ ou aussi $B = A + \overrightarrow{u}$. Donc,

$$\overrightarrow{GA} + 2\overrightarrow{GB} = \overrightarrow{0} \Leftrightarrow (A - G) + 2(B - G) = \overrightarrow{0} \Leftrightarrow G = \frac{A + 2B}{1 + 2}.$$

Cette dernière égalité s'écrit aussi $G = \frac{1}{3}A + \frac{2}{3}B$. On généralise cette situation :

Définition 1. Soit E un $\mathbb{K}$-espace vectoriel. Soient $n \in \mathbb{N}^*$ puis $a_1, \dots, a_n$ n éléments de E. Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ tel que $\lambda_1 + \dots + \lambda_n \neq 0$.
Le **barycentre du système de points pondérés** $(a_1(\lambda_1), \dots, a_n(\lambda_n))$ est $g = \dfrac{\lambda_1 a_1 + \dots + \lambda_n a_n}{\lambda_1 + \dots + \lambda_n}$.

Notation. Le barycentre du système $(a_1(\lambda_1), \dots, a_n(\lambda_n))$ se note $\text{bar}(a_1(\lambda_1), \dots, a_n(\lambda_n))$.

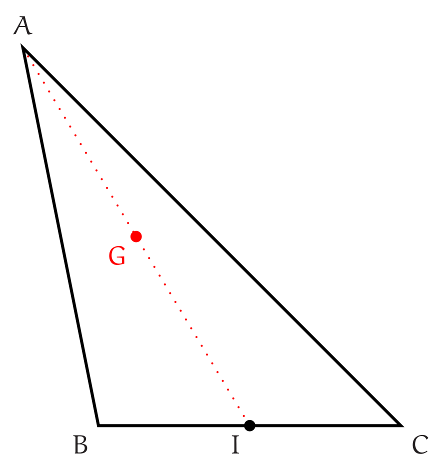
On a immédiatement

Théorème 1. g est l'unique point de E vérifiant $\lambda_1\overrightarrow{ga_1} + \dots + \lambda_n\overrightarrow{ga_n} = \overrightarrow{0}$.

Exemple. Soit ABC un triangle du plan. Construisons $G = \text{bar}(A(2), B(1), C(1))$. 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. Soit E un $\mathbb{K}$-espace vectoriel. Soient $n \in \mathbb{N}^*$ puis $a_1, \dots, a_n$ n éléments de E. Soit $(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ tel que $\lambda_1 + \dots + \lambda_n \neq 0$. Soit $k \in \mathbb{R} \setminus \{0\}$. Alors

$$\text{bar}(a_1(k\lambda_1), \dots, a_n(k\lambda_n)) = \text{bar}(a_1(\lambda_1), \dots, a_n(\lambda_n)).$$

Démonstration. $k\lambda_1 + \ldots + k\lambda_n \neq 0$ et de plus

$$\text{bar}\left(a_1\left(k\lambda_1\right), \ldots, a_n\left(k\lambda_n\right)\right) = \frac{k\lambda_1 a_1 + \ldots + k\lambda_n a_n}{k\lambda_1 + \ldots + k\lambda_n} = \frac{\lambda_1 a_1 + \ldots + \lambda_n a_n}{\lambda_1 + \ldots + \lambda_n} = \text{bar}\left(a_1\left(\lambda_1\right), \ldots, a_n\left(\lambda_n\right)\right).$$

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

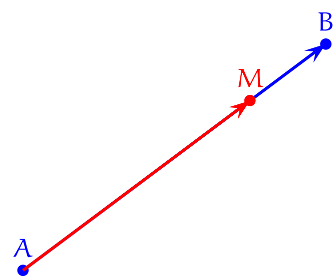
$$\sum_{i=1}^{n} \lambda_i' = \frac{1}{\lambda_1 + \ldots + \lambda_n} \sum_{i=1}^{n} \lambda_i = 1.$$

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

2) *Convexes*

2-a) *Segments*

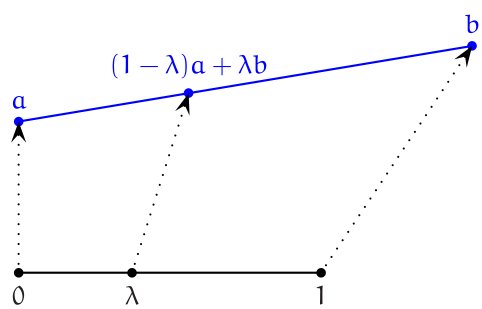
Soient A et B deux points d'un $\mathbb{K}$-espace vectoriel E. Le segment $[A, B]$ est l'ensemble des points M de E tel qu'il existe un réel $\lambda \in [0, 1]$ tel que $\overrightarrow{AM} = \lambda \overrightarrow{AB}$.



Cette dernière égalité s'écrit encore $M - A = \lambda(B - A)$ ou enfin $M = (1 - \lambda)A + \lambda B$. Donc

Définition 2. Soit E un $\mathbb{K}$-espace vectoriel. Soient a et b deux éléments de E.
Le segment $[a, b]$ est $\{(1 - \lambda)a + \lambda b,\ \lambda \in [0, 1]\}$ ou encore $\{\text{bar}(a(1 - \lambda), b\lambda),\ \lambda \in [0, 1]\}$.

Commentaire. La présentation d'un segment comme ensemble des $(1-\lambda)a+\lambda b$, $\lambda \in [0, 1]$, est meilleure que la présentation comme ensemble des $\lambda a + (1 - \lambda)b$. Quand λ croît de 0 à 1, $(1 - \lambda)a + \lambda b$ parcourt le segment $[a, b]$ de a à b alors que $\lambda a + (1 - \lambda)b$ 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 $[a, b]$ de E.



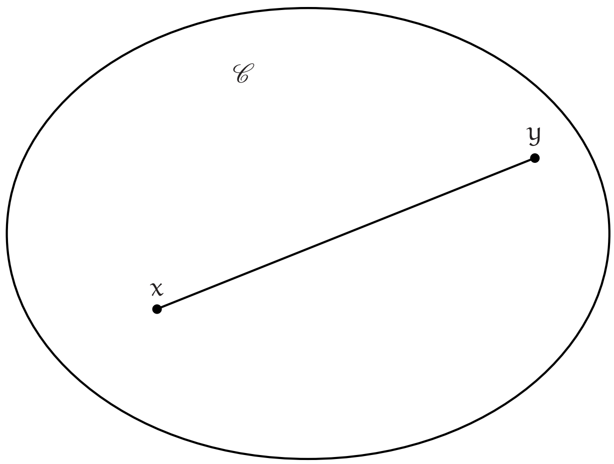
2-b) *Parties convexes d'un espace vectoriel*

Définition 3. Soit $\mathscr{C}$ une partie non vide d'un $\mathbb{K}$-espace vectoriel E. $\mathscr{C}$ est **convexe** si et seulement si pour tout $(x, y) \in \mathscr{C}^2$, $[x, y] \subset \mathscr{C}$.

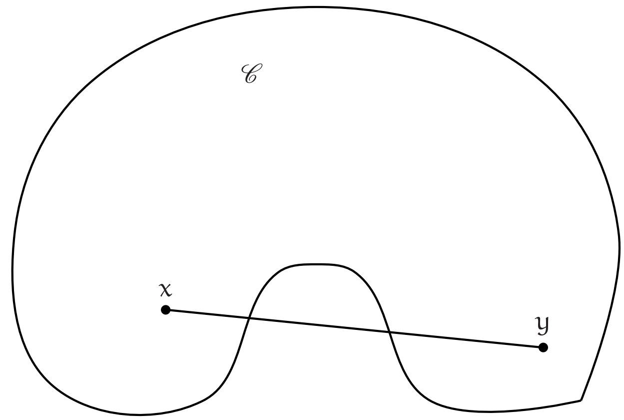
Commentaire. Il revient au même de dire :

$$\mathscr{C} \text{ est convexe } \Leftrightarrow \forall (x,y) \in \mathscr{C}^2,\ \forall \lambda \in [0,1],\ (1-\lambda)x + \lambda y \in \mathscr{C}.$$

Convention. $\varnothing$ est convexe.



$\mathscr{C}$ est convexe car
$\forall (x,y) \in \mathscr{C}^2$, $[x,y] \subset \mathscr{C}$



$\mathscr{C}$ n'est pas convexe car
$\exists (x,y) \in \mathscr{C}^2 /\ [x,y] \not\subset \mathscr{C}$

Théorème 3. Soit $\mathscr{C}$ une partie non vide d'un $\mathbb{K}$-espace vectoriel E. $\mathscr{C}$ est **convexe** si et seulement si

$$\forall n \in \mathbb{N}^*,\ \forall (x_1, \dots, x_n) \in \mathscr{C}^n,\ \forall (\lambda_1, \dots, \lambda_n) \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow \sum_{i=1}^{n} \lambda_i x_i \in \mathscr{C}\right).$$

Démonstration. $\Longleftarrow$ / Supposons que $\forall n \geqslant 2,\ \forall (x_1, \dots, x_n) \in \mathscr{C}^n,\ \forall (\lambda_1, \dots, \lambda_n) \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow \sum_{i=1}^{n} \lambda_i x_i \in \mathscr{C}\right)$.

En particulier, pour $n = 2$, on a

$$\forall (x_1, x_2) \in \mathscr{C}^2,\ \forall (\lambda_1, \lambda_2) \in [0,1]^2,\ (\lambda_1 + \lambda_2 = 1 \Rightarrow \lambda_1 x_1 + \lambda_1 x_2 \in \mathscr{C}).$$

Ceci s'écrit encore $\forall (x_1, x_2) \in \mathscr{C}^2,\ \forall \lambda \in [0,1],\ (1-\lambda)x_1 + \lambda x_2 \in \mathscr{C}$ ce qui est la définition d'un convexe.

$\Longrightarrow$ / Supposons que $\mathscr{C}$ est convexe.

Montrons par récurrence que $\forall n \geqslant 2,\ \forall (x_1, \dots, x_n) \in \mathscr{C}^n,\ \forall (\lambda_1, \dots, \lambda_n) \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow \sum_{i=1}^{n} \lambda_i x_i \in \mathscr{C}\right)$ (le résultat étant immédiat quand $n = 1$).

- La propriété est vraie quand $n = 2$ par définition d'un convexe.
- Soit $n \geqslant 2$. Supposons que $\forall (x_1, \dots, x_n) \in \mathscr{C}^n,\ \forall (\lambda_1, \dots, \lambda_n) \in [0,1]^n,\ \left(\sum_{i=1}^{n} \lambda_i = 1 \Rightarrow \sum_{i=1}^{n} \lambda_i x_i \in \mathscr{C}\right)$.

  Soient $(x_1, \dots, x_n, x_{n+1}) \in \mathscr{C}^{n+1}$ puis $(\lambda_1, \dots, \lambda_n, \lambda_{n+1}) \in [0,1]^{n+1}$ tel que $\lambda_1 + \ldots + \lambda_n + \lambda_{n+1} = 1$.

  Si $\lambda_{n+1} = 1$, alors $\forall i \in [\![1,n]\!]$, $\lambda_i = 0$ et donc $\sum_{i=1}^{n+1} \lambda_i x_i = x_{n+1} \in \mathscr{C}$.

  Sinon, $\lambda_{n+1} \in [0,1[$ et on peut écrire

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

  Pour tout $i \in [\![1,n]\!]$, $\dfrac{\lambda_i}{1 - \lambda_{n+1}} \geqslant 0$. De plus,

$$\sum_{i=1}^{n} \frac{\lambda_i}{1 - \lambda_{n+1}} = \frac{1}{1 - \lambda_{n+1}} \sum_{i=1}^{n} \lambda_i = \frac{1 - \lambda_{n+1}}{1 - \lambda_{n+1}} = 1.$$

Les réels $\dfrac{\lambda_i}{1-\lambda_{n+1}}$, $1 \leqslant i \leqslant n$, sont donc n réels de $[0,1]$ et de somme égale à 1. Par hypothèse de récurrence, $\displaystyle\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i \in \mathscr{C}$ puis, d'après le cas $n=2$, $\displaystyle(1-\lambda_{n+1})\left(\sum_{i=1}^{n} \frac{\lambda_i}{1-\lambda_{n+1}} x_i\right) + \lambda_{n+1}x_{n+1} \in \mathscr{C}$.

Le résultat est démontré par récurrence.

Théorème 4. Soit E un $\mathbb{K}$-espace vectoriel. Soient I un ensemble non vide d'indices puis $(\mathscr{C}_i)_{i\in I}$ une famille de convexes de E indexée par I. Alors $\displaystyle\bigcap_{i\in I} \mathscr{C}_i$ est un convexe de E.

Démonstration. Posons $\mathscr{C} = \displaystyle\bigcap_{i\in I} \mathscr{C}_i$. Si $\mathscr{C}$ est vide, alors $\mathscr{C}$ est un convexe de E par convention.

Supposons $\mathscr{C} \neq \varnothing$. Soit $(x,y) \in E^2$.

$$(x,y) \in \mathscr{C}^2 \Rightarrow \forall i \in I,\ (x,y) \in \mathscr{C}_i^2 \Rightarrow \forall i \in I,\ [x,y] \subset \mathscr{C}_i \Rightarrow [x,y] \subset \mathscr{C}$$

et donc $\mathscr{C}$ est convexe.

2-c) Enveloppe convexe

Théorème 5. Soit E un $\mathbb{K}$-espace vectoriel. Soit $\mathscr{A}$ une partie de E. Il existe une plus petite partie convexe (au sens de l'inclusion) contenant $\mathscr{A}$.

Démonstration. Il existe au moins un convexe contenant $\mathscr{A}$ à savoir l'espace E tout entier. Soit $\mathscr{C}$ l'intersection de tous les convexes de E contenant $\mathscr{A}$.
$\mathscr{C}$ contient $\mathscr{A}$ en tant qu'intersection de parties contenant $\mathscr{A}$ et $\mathscr{C}$ est un convexe en tant qu'intersection de convexes. Donc, $\mathscr{C}$ est un convexe contenant $\mathscr{A}$. Enfin, tout convexe contenant $\mathscr{A}$ contient l'intersection de tous les convexes contenant $\mathscr{A}$ c'est-à-dire $\mathscr{C}$.

Définition 4. Soit $\mathscr{A}$ une partie de E. Le plus petit convexe (au sens de l'inclusion) de E contenant $\mathscr{A}$ s'appelle l'**enveloppe convexe** de $\mathscr{A}$ et se note $\text{conv}(\mathscr{A})$.

Théorème 6. Soit E un $\mathbb{K}$-espace vectoriel. Soit $\mathscr{A}$ une partie non vide de E. L'enveloppe convexe de $\mathscr{A}$ est l'ensemble des $\displaystyle\sum_{i=1}^{n} \lambda_i x_i$ où $n \in \mathbb{N}^*$, $(x_1,\ldots,x_n) \in \mathscr{A}^n$, $(\lambda_1,\ldots,\lambda_n) \in [0,1]^n$ et $\displaystyle\sum_{i=1}^{n} \lambda_i = 1$.

Dit autrement, $\text{conv}(\mathscr{A})$ est l'ensemble des barycentres à coefficients positifs d'éléments de $\mathscr{A}$.

Démonstration. Notons $\mathscr{C}$ l'ensemble des barycentres à coefficients positifs d'éléments de $\mathscr{A}$.

$\text{conv}(\mathscr{A})$ contient les éléments de $\mathscr{A}$ et donc les barycentres à coefficients positifs d'éléments de $\mathscr{A}$ d'après le théorème 3, page 3, et donc $\mathscr{C} \subset \text{conv}(\mathscr{A})$.

Inversement, un élément de $\mathscr{A}$ est un barycentre à coefficients positifs d'éléments de $\mathscr{A}$ et donc $\mathscr{C}$ contient $\mathscr{A}$. De plus, soit $x = \displaystyle\sum_{i=1}^{n} \alpha_i x_i$ où les x_i sont dans $\mathscr{A}$ et les α_i sont dans $[0,1]$ et tels que $\displaystyle\sum_{i=1}^{n} \alpha_i = 1$ et soit $y = \displaystyle\sum_{i=1}^{p} \beta_i y_i$ où les y_i sont dans $\mathscr{A}$ et les β_i sont dans $[0,1]$ et tels que $\displaystyle\sum_{i=1}^{p} \beta_i = 1$. Alors, pour $\lambda \in [0,1]$,

$$(1-\lambda)x + \lambda y = (1-\lambda)\alpha_1 x_1 + \ldots + (1-\lambda)\alpha_n x_n + \lambda\beta_1 y_1 + \ldots \lambda\beta_p y_p,$$

avec $\forall i \in [\![1,n]\!]$, $0 \leqslant (1-\lambda)\alpha_i \leqslant 1-\lambda \leqslant 1$, $\forall j \in [\![1,p]\!]$, $0 \leqslant \lambda\beta_j \leqslant \lambda \leqslant 1$ et enfin

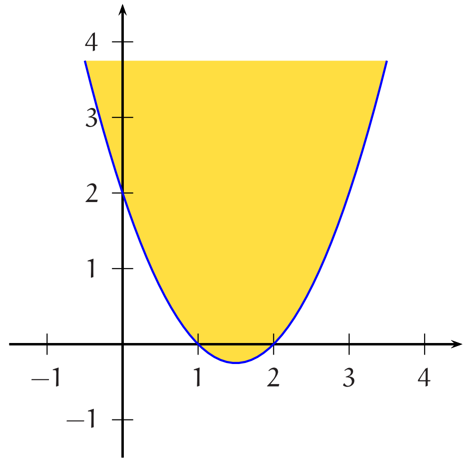
$$\sum_{i=1}^{n}(1-\lambda)\alpha_i + \sum_{j=1}^{p}\lambda\beta_j = (1-\lambda)\sum_{i=1}^{n}\alpha_i + \lambda\sum_{j=1}^{p}\beta_j = 1-\lambda+\lambda = 1.$$

Donc $(1-\lambda)x + \lambda y \in \mathscr{C}$. Ainsi, $\mathscr{C}$ est un convexe contenant $\mathscr{A}$ et donc $\text{conv}(\mathscr{A}) \subset \mathscr{C}$. On a montré que $\text{conv}(\mathscr{A}) = \mathscr{C}$.

II - Fonctions convexes

1) *Définition*

DÉFINITION 5. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$. L'**épigraphe** de f est $\{(x,y) \in I \times \mathbb{R} / \ y \geqslant f(x)\}$.



DÉFINITION 6. 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.

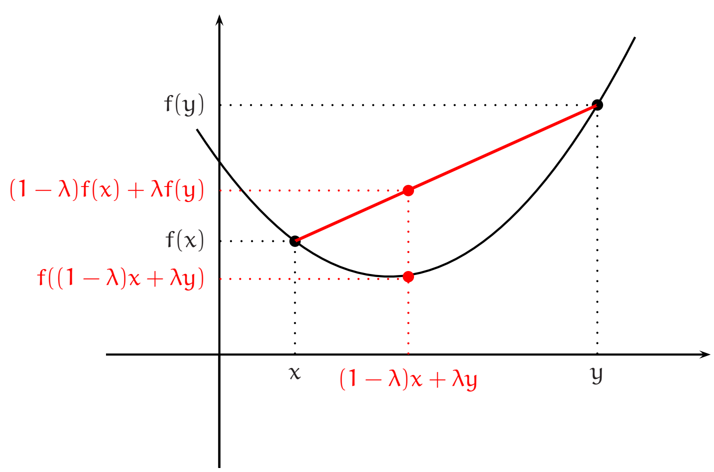
Théorème 7. Soit f une fonction définie sur un intervalle I de $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

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

Commentaire. 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.

DÉFINITION 7. 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 $x \neq y$, $\forall \lambda \in]0,1[$, $f((1-\lambda)x + \lambda y) > (1-\lambda)f(x) + \lambda f(y)$.



Démonstration du théorème 7.

• Supposons que f est convexe sur I et donc que l'épigraphe $\mathscr{E}$ 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 $\mathscr{E}$. Donc, le segment $[A_1, A_2]$ est contenu dans $\mathscr{E}$. En particulier, le point $(1-\lambda)A_1 + \lambda A_2$ est un point de $\mathscr{E}$. 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 $f\left((1-\lambda)x_1+\lambda x_2\right) \leqslant (1-\lambda)f\left(x_1\right)+\lambda f\left(x_2\right)$.

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

Soient $A_1=(x_1,y_1)$ et $A_2=(x_2,y_2)$ deux points de $\mathscr{E}$ et $\lambda \in [0,1]$. Soit $M=(1-\lambda)A_1+\lambda A_2$. Posons $M=(x,y)$. Alors

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

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

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

Théorème 8. 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

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

Démonstration.

Si $\forall n \in \mathbb{N}^*$, $\forall\left(x_i\right)_{1\leqslant i\leqslant n} \in I^n$, $\forall\left(\lambda_i\right)_{1\leqslant i\leqslant n} \in [0,1]^n$, $\left(\sum_{i=1}^{n}\lambda_i=1 \Rightarrow f\left(\sum_{i=1}^{n}\lambda_i x_i\right) \leqslant \sum_{i=1}^{n}\lambda_i f\left(x_i\right)\right)$, quand $n=2$, on obtient la définition d'une fonction convexe.

Réciproquement, si f est convexe, son épigraphe est convexe et le théorème 3, page 3, montre le résultat.

---

2) *Caractérisation par la « fonction pente »*

Théorème 9. 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

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

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

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, $\varphi_{x_1}\left((1-\lambda)x_1+\lambda x_2\right) \leqslant \varphi_{x_1}\left(x_2\right)$. Cette inégalité s'écrit explicitement

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

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

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

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

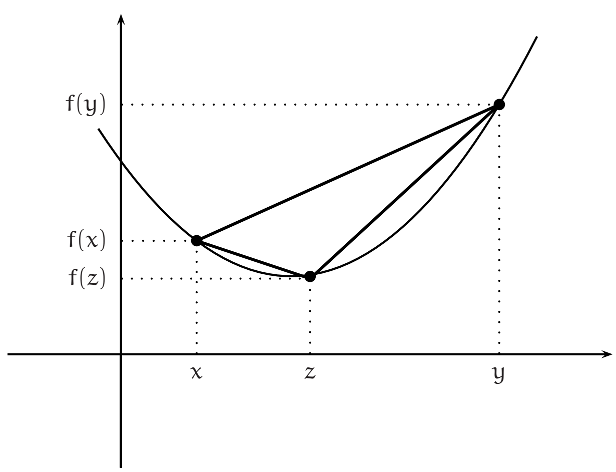
3ème cas. Supposons $x_1 < x_0 < x_2$. D'après les deux premiers cas,

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

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 $x < z < y$,

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



3) *Caractérisation des fonctions convexes dérivables*

Théorème 10. 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, 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 9), 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 11. 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 :

Corollaire. 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 $f'' \leqslant 0$.

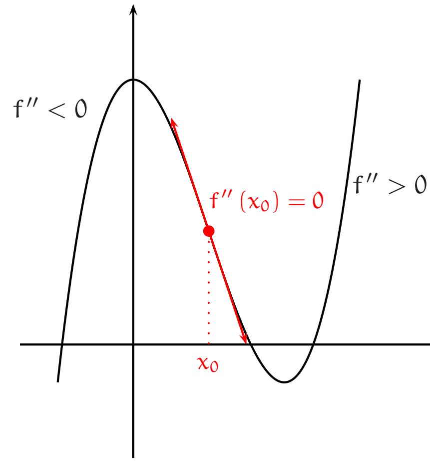
Si $f'' > 0$ sur I sauf peut-être en un nombre fini de points, alors f est strictement convexe sur I.
Si $f'' < 0$ sur I sauf peut-être en un nombre fini de points, alors f est strictement concave sur I.

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

Définition 8. 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 ».



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 $\mathscr{C}_f$ de f est au-dessus de sa tangente (T_a) en son point d'abscisse a sur I. Une équation de la tangente à $\mathscr{C}_f$ 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 $\mathscr{C}_f$ est au-dessus de (T_a) sur I.

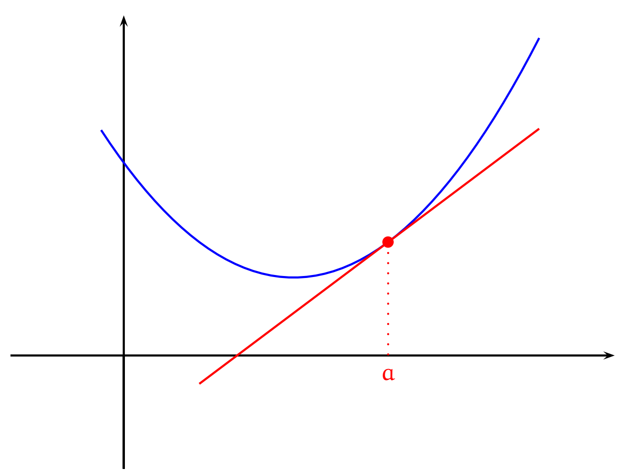
Commentaire 1. Si de plus, f est de classe C^1 sur I, on peut aussi constater que pour $x \in I$,

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

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

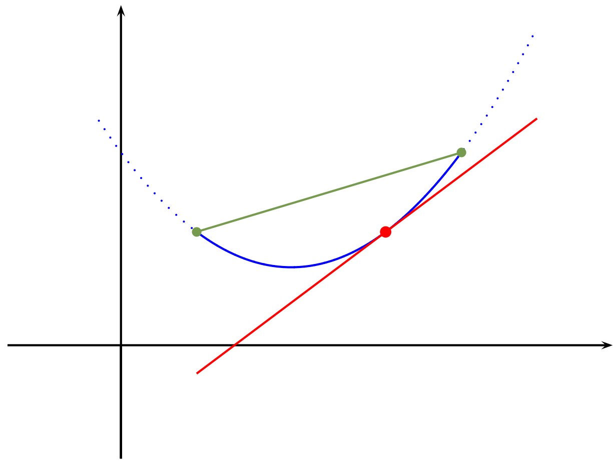
Commentaire 2. Dans le cas où f est strictement convexe, en adaptant la démonstration ci-dessus, on montre que $\mathscr{C}_f$ est strictement au-dessus de (T_a) sur $I \setminus \{a\}$.

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



5) Inégalités de convexité

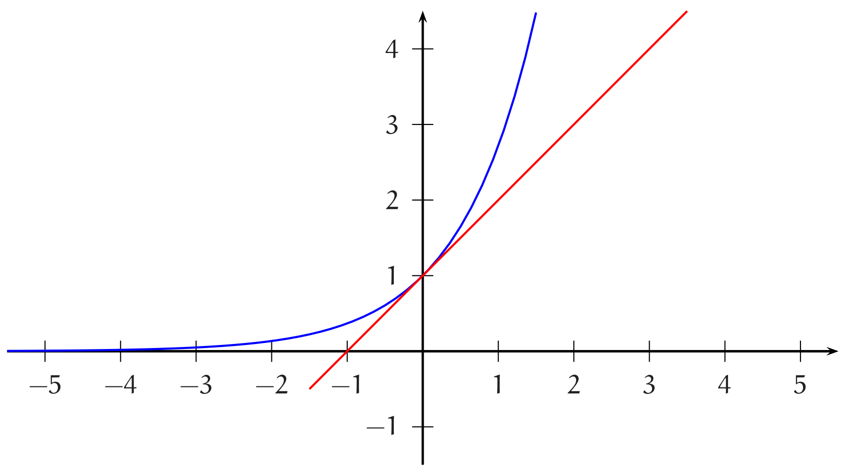
De ce qui précède, on a l'habitude de déduire que « le graphe d'une fonction convexe est au-dessus de ses tangentes et au-dessous de ses cordes » (étant entendu que le graphe est au-dessus de sa tangente sur I tout entier et au-dessous de la corde joignant les points $(a, f(a))$ et $(b, f(b))$ sur le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

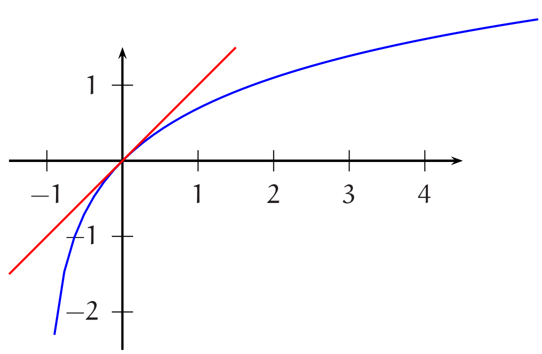
• La fonction exponentielle est strictement convexe sur $\mathbb{R}$ car sa dérivée seconde à savoir $x \mapsto e^x$, est strictement positive sur $\mathbb{R}$. Son graphe est donc au-dessus de sa tangente en $(0, e^0) = (0, 1)$ sur $\mathbb{R}$ et strictement au-dessus sur $\mathbb{R}^*$. Ceci fournit

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



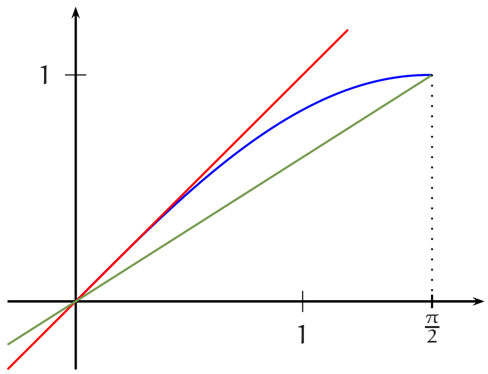
• La fonction $x \mapsto \ln(1 + x)$ est strictement concave sur $]-1, +\infty[$ car sa dérivée seconde à savoir $x \mapsto -\dfrac{1}{(1 + x)^2}$, est strictement négative sur $]-1, +\infty[$. Son graphe est donc au-dessous de sa tangente en $(0, \ln 1) = (0, 0)$ sur $]-1, +\infty[$ et strictement au-dessous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

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



• Signalons enfin l'inégalité usuelle entre la moyenne géométrique et la moyenne arithmétique de n nombres positifs. Soient $n \geqslant 2$ puis $(x_1, \dots, x_n) \in]0, +\infty[^n$. La fonction $\ln$ est concave sur $]0, +\infty[$ et donc

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

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

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

Cette inégalité reste claire quand l'un des x_i est nul et donc

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

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