Chapitre 28

Fonctions convexes

Objectifs

- Définir la notion de fonction convexe.
- Étudier différentes caractérisations de fonctions convexes.
- Étudier la régularité des fonctions convexes.
- Donner des exemples d'inégalités de convexité.

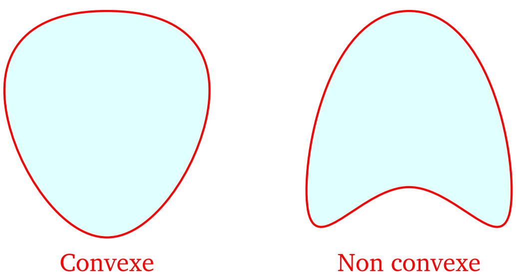
Sommaire

Dans ce chapître on désigne par I un intervalle non trivial de $\mathbb{R}$, et toutes les fonctions considérées vont de I vers $\mathbb{R}$.

I) Définition

1) Origine géométrique

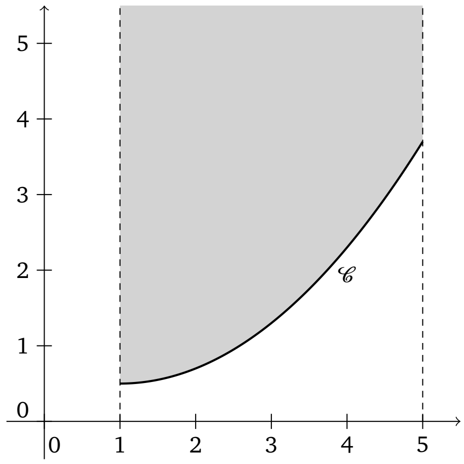
Nous avons déjà défini la notion de convexité pour les parties d'un espace euclidien E : une partie $\mathscr{V}$ non vide de E est *convexe* lorsque pour tous points A et B de $\mathscr{V}$, le segment $[A,B]$ est inclus dans $\mathscr{V}$.



Soit $f : I \to \mathbb{R}$, on dira que f est convexe sur I lorsque la partie du plan située **au-dessus de la courbe** est une partie convexe, on dira que f est concave sur I lorsque $-f$ est convexe.

DÉFINITION 28.1

$f : I \mapsto \mathbb{R}$ est dite convexe lorsque $\{(x, y) \ / \ x \in I \text{ et } y \geqslant f(x)\}$ est une partie convexe de $\mathbb{R}^2$.

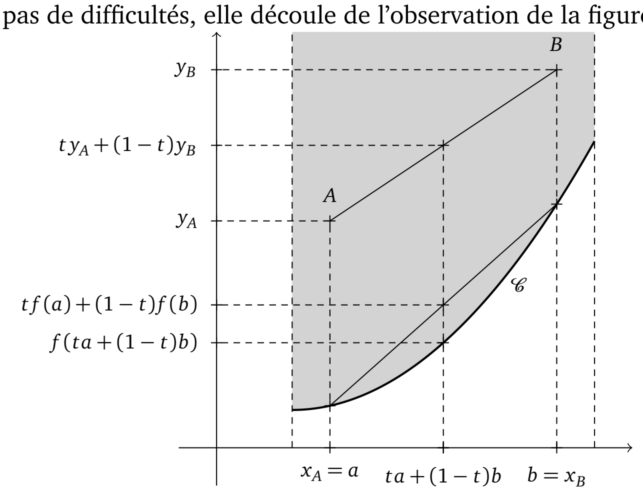


THÉORÈME 28.1

$f : I \longrightarrow \mathbb{R}$ est convexe si et seulement si pour tout réels a, b de I et tout réel t de $[0;1]$ on a :

$$f(ta+(1-t)b) \leqslant tf(a)+(1-t)f(b)$$

Preuve: Celle-ci ne pose pas de difficultés, elle découle de l'observation de la figure suivante :



□

Remarques:
- Le théorème ci-dessus signifie que sur l'intervalle I, « chaque arc est sous sa corde ».
- f est concave sur I si et seulement si :

$$\forall x, y \in I, \ \forall t \in [0;1], \ f(ta+(1-t)b) \geqslant tf(a)+(1-t)f(b)$$

- La convexité est une notion globale, une fonction peut être convexe sur un intervalle et pas sur un autre. De même, une fonction peut être ni concave ni convexe sur un intervalle.

Exercices:
- Montrer que les fonctions affines sur $\mathbb{R}$ sont convexes.
- Montrer que la fonction $x \mapsto x^2$ est convexe sur $\mathbb{R}$.

Par récurrence sur $n \geqslant 2$, et en utilisant l'associativité du barycentre, on établit la généralisation suivante :

THÉORÈME 28.2

$f : I \to \mathbb{R}$ est convexe si et seulement si pour tous réels $a_1, \dots, a_n$ de I et tous réels $t_1, \dots, t_n$ de $[0;1]$ tels que $t_1 + \cdots + t_n = 1$, on a :

$$f(t_1a_1 + \cdots t_na_n) \leqslant t_1f(a_1) + \cdots + t_nf(a_n)$$

2) Convexité et pentes

THÉORÈME 28.3

$f : I \to \mathbb{R}$ est convexe sur I si et seulement si pour tout $a \in I$, la fonction $t_a : x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $I \setminus \{a\}$.

Preuve: Si f est convexe sur I, soit $a \in I$, envisageons plusieurs cas :

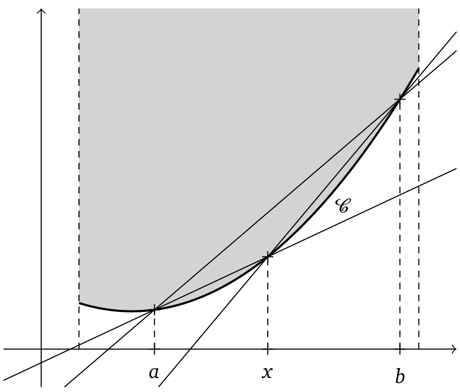
- Si $a < x < y$ sont dans I, l'arc sur $[a;y]$ est sous la corde, donc $f(x) \leqslant \frac{f(y)-f(a)}{y-a}(x-a) + f(a)$ ce qui entraîne $\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(y)-f(a)}{y-a}$ car $x - a > 0$.
- Si $x < y < a$ sont dans I, le même argument donne $f(y) \leqslant \frac{f(x)-f(a)}{x-a}(y-a) + f(a)$ ce qui entraîne $\frac{f(y)-f(a)}{y-a} \geqslant \frac{f(x)-f(a)}{x-a}$ car $y - a < 0$.
- Si $x < a < y$ sont dans I alors $f(a) \leqslant \frac{f(y)-f(x)}{y-x}(a-x) + f(x)$ d'où $\frac{f(a)-f(x)}{a-x} \leqslant \frac{f(y)-f(x)}{y-x}$ car $a - x > 0$. En écrivant $a = tx + (1-t)y$ (avec $t \in]0;1[$) on a $y - a = t(y-x)$ et $f(a) \leqslant tf(x) + (1-t)f(y)$ donc $f(y) - f(a) \geqslant t(f(y) - f(x))$ on en déduit que $\frac{f(y)-f(x)}{y-x} = \frac{t(f(y)-f(x))}{t(y-x)} \leqslant \frac{f(y)-f(a)}{y-a}$, par conséquent $\frac{f(a)-f(x)}{a-x} \leqslant \frac{f(y)-f(a)}{y-a}$.

La fonction $t_a : x \mapsto \frac{f(x)-f(a)}{x-a}$ est donc croissante sur $I \setminus \{a\}$.

Réciproquement, si la fonction $t_a : x \mapsto \frac{f(x)-f(a)}{x-a}$ est croissante sur $I \setminus \{a\}$ pour tout $a \in I$. Soit $a < x < b$ dans I, alors $\frac{f(x)-f(a)}{x-a} \leqslant \frac{f(b)-f(a)}{b-a}$ donc $f(x) \leqslant \frac{f(b)-f(a)}{b-a}(x-a) + f(a)$, ce qui signifie que sur $[a;b]$, l'arc est sous la corde, f est donc convexe. □

Remarques:

- La fonction t_a est la fonction taux d'accroissement de f en a.
- Si f est convexe sur I, alors l'arc sur $[a;b]$ est sous la corde, mais en dehors de $[a;b]$ la courbe est au-dessus.
- Si f est convexe sur I et si $a < b$ sont dans I, alors $t_a \leqslant t_b$.



Application: par exemple, sachant que la fonction ln est concave sur $]0;+\infty[$, on peut affirmer que la fonction $x \mapsto \frac{\ln(x)}{x-1}$ est décroissante sur son ensemble de définition, car c'est le taux d'accroissement en 1 de la fonction ln.

THÉORÈME 28.4

Si f est convexe sur I et admet un minimum local en a, alors c'est un minimum global.

Preuve: Au voisinage à gauche de a la fonction t_a est négative, comme cette fonction est croissante on en déduit que pour $x < a$ on a $\frac{f(x)-f(a)}{x-a} \leqslant 0$ et donc $f(x) \geqslant f(a)$. Au voisinage à droite de a la fonction t_a est positive, comme cette fonction est croissante on en déduit que pour $x > a$ on a $\frac{f(x)-f(a)}{x-a} \geqslant 0$ et donc $f(x) \geqslant f(a)$. □

II) Convexité et régularité

1) Dérivabilité à gauche et à droite

THÉORÈME 28.5

*Soit f une fonction convexe sur I et a un point **intérieur** à I, alors f est dérivable à gauche et à droite en a. De plus, si $x < a < y$ alors :*

$$\frac{f(x)-f(a)}{x-a} \leqslant f'_g(a) \leqslant f'_d(a) \leqslant \frac{f(y)-f(a)}{y-a}.$$

En particulier, si $a < b$ sont intérieurs à I, alors $f'_d(a) \leqslant f'_g(b)$.

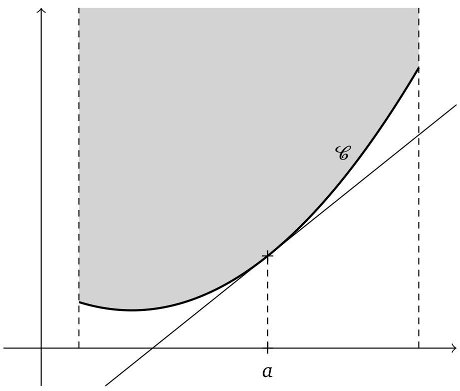
Preuve: La fonction t_a est croissante sur $I \setminus \{a\}$ donc elle majorée sur $I\cap]-\infty;a[$ par $t_a(y)$, par conséquent elle admet une limite finie à gauche en a qui est $\sup\limits_{x<a} t_a(x)$, cela signifie que f est dérivable à gauche en a et $f'_g(a) = \sup\limits_{x<a} t_a(x) \leqslant t_a(y)$.

Sur $I\cap]a;+\infty[$ la fonction t_a est croissante minorée par $f'_g(a)$, elle admet une limite finie à droite en a qui est $\inf\limits_{x>a} t_a(x)$, cela signifie que f est dérivable à droite en a et $f'_g(a) \leqslant f'_d(a) = \inf\limits_{x>a} t_a(x) \leqslant t_a(y)$. □

- *La fonction f n'a aucune raison d'être dérivable aux bornes I, par exemple, la fonction* arcsin *est convexe sur $[0;1]$ mais non dérivable à gauche en 1.*
- *Le théorème ne dit pas que f est dérivable en a ! Considérer par exemple $f(t) = |t|$ sur $[-1;1]$, elle est convexe mais non dérivable en 0.*

Application: Si f une fonction convexe sur I et a un point **intérieur** à I, alors :

- sur $I\cap]-\infty;a[$ la courbe de f est au-dessus de la droite d'quation $y = f'_g(a)(x-a)+f(a)$ (tangente à la courbe à gauche au point d'abscisse c). En effet, pour $x < a$, l'inégalité $t_a(x) \leqslant f'_g(a)$ équivaut à $f(x) \geqslant f'_g(a)(x-a)+f(a)$.
- sur $I\cap]a;+\infty[$ la courbe de f est au-dessus de la droite d'quation $y = f'_d(a)(x-a)+f(a)$ (tangente à la courbe à droite au point d'abscisse c). En effet, pour $x > a$, l'inégalité $f'_d(a) \leqslant t_a(x)$ équivaut à $f(x) \geqslant f'_d(a)(x-a)+f(a)$.
- si f est dérivable en a alors la courbe de f est au-dessus de la tangente au point d'abscisse a.



Il découle du théorème précédent :

THÉORÈME 28.6

*Si f est convexe sur I alors f est continue en tout point **intérieur** à I.*

La fonction f n'a aucune raison d'être continue aux bornes I, par exemple, la fonction f définie sur $[0;1]$ par $f(t)=0$ si $t\in[0;1[$ et $f(1)=1$ est convexe mais discontinue en 1.

THÉORÈME 28.7

*On sait qu'une fonction f, **continue** et **convexe** sur un **segment** $[a;b]$, admet un maximum global. Si ce maximum global est atteint à l'interieur de I, alors f est constante. On en déduit que si f n'est pas constante, le maximum est atteint uniquement en a ou en b (voire les deux).*

Preuve: Soit $M=f(c)$ ce maximum avec $c\in]a;b[$, à gauche de c on a $t_c(x)\geqslant 0$, alors qu'à droite de c on a $t_c(x)\leqslant 0$, or la fonction t_c est croissante sur $[a;b]\setminus\{c\}$, par conséquent la fonction t_c est nulle, ce qui signifie que f est constante. □

2) Convexité des fonctions dérivables

THÉORÈME 28.8

Une fonction $f: I\to\mathbb{R}$, dérivable sur I, est convexe si et seulement si f' est croissante sur I.

Preuve: Si f est convexe : on sait que pour tous points $a<x<b$ de I, on a $t_a(x)\leqslant t_a(b)$, en faisant tendre x vers a on obtient $f'(a)\leqslant t_a(b)$. De même $t_b(a)\leqslant t_b(x)$, en faisant tendre x vers b on obtient $t_a(b)\leqslant f'(b)$, ce qui entraîne $f'(a)\leqslant f'(b)$ car $t_(b)=t_b(a)$.

Réciproquement, si f' est croissante : soient $a<b$ dans I, on pose $g(x)=f(x)-\frac{f(b)-f(a)}{b-a}(x-a)-f(a)$ sur $[a;b]$, cette fonction est dérivable et $g'(x)=f'(x)-\frac{f(b)-f(a)}{b-a}$, or il existe $c\in]a;b[$ tel que $f'(c)=\frac{f(b)-f(a)}{b-a}$ (égalité des accroissements finis), comme la fonction g' est croissante on a g décroissante sur $[a;c]$ puis croissante sur $[c;b]$, comme $g(a)=g(b)=0$ on a $g(x)\leqslant 0$ sur $[0;1]$ et donc $f(x)\leqslant\frac{f(b)-f(a)}{b-a}(x-a)+f(a)$, c'est à dire que f est convexe sur I.

x	a		c		b
g'		$-$	0	$+$	
g	0	$\searrow$	$g(c)$	$\nearrow$	0

□

Il en découle qu'une fonction f deux fois dérivable sur un intervalle I est convexe si et seulement si sa dérivée seconde et positive sur I (et concave si et seulement si $f''\leqslant 0$ sur I).

Exemples:
- La fonction exp est convexe sur $\mathbb{R}$.
- La fonction ln est concave sur $]0;+\infty[$.
- La fonction sin est convexe sur $[-\pi;0]$ et concave sur $[0;\pi]$.
- La fonction $x\mapsto x^n$ avec $n\in\mathbb{N}^*$ est convexe sur $\mathbb{R}$ lorsque n est pair, sur $[0;+\infty[$ seulement si n est impair.

III) Inégalités de convexité

1) Tangentes ou sécantes

Lorsqu'une fonction f est dérivable et convexe sur I, on sait que la courbe est au-dessus chacune de ses tangentes :

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

Pour les fonctions concaves, l'inégalité est inversée.

Exemples:

– La fonction exp est convexe sur $\mathbb{R}$ donc :

$$\forall a,\ x \in \mathbb{R},\ e^x \geqslant e^a(x-a)+e^a$$

En particulier avec $a=0$ on retrouve l'inégalité classique : $e^x \geqslant x+1$.

– La fonction ln est concave sur $]0;+\infty[$ donc :

$$\forall a,\ x \in]0;+\infty[,\ \ln(x) \leqslant \frac{x-a}{a}+\ln(a)$$

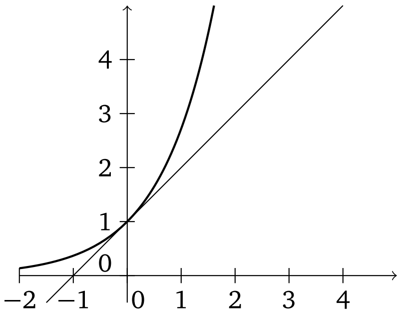
En particulier avec $a=1$ on retrouve l'inégalité classique : $\forall x>0,\ \ln(x) \leqslant x-1$.

– La fonction sin est concave sur $[0;\pi]$ donc :

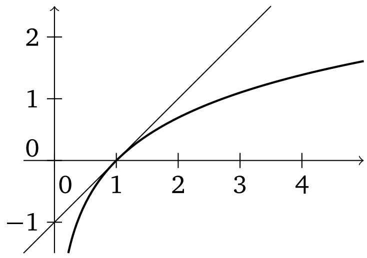
$$\forall a,\ x \in [0;\pi],\ \sin(x) \leqslant \cos(a)(x-a)+\sin(a)$$

En particulier avec $a=0$ on retrouve l'inégalité classique : $\forall x \in [0;\pi],\ \sin(x) \leqslant x$.

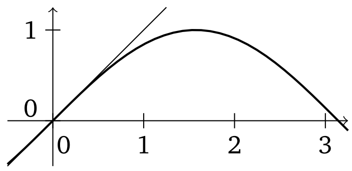
– ...



$e^x \geqslant x+1$



$\ln(x) \leqslant x-1$



$\sin(x) \leqslant x$

Lorsqu'une fonction f est convexe sur I, on sait que sur tout segment $[a;b] \subset I$ l'arc est sous la corde :

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

Pour les fonctions concaves, l'inégalité est inversée.

Exemples:

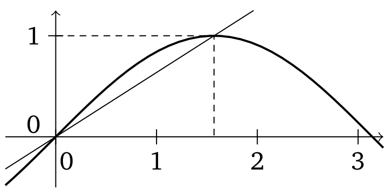
– La fonction sin est concave sur $[0;\frac{\pi}{2}]$ donc :

$$\forall x \in [0;\frac{\pi}{2}],\ \sin(x) \geqslant \frac{2x}{\pi}$$

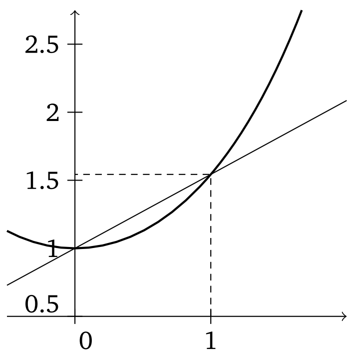
– La fonction ch est convexe sur $[0;a]$ $(a>0)$ donc :

$$\forall x \in [0;a],\ \mathrm{ch}(x) \leqslant \frac{\mathrm{ch}(a)-1}{a}x+1$$

– ...



$\sin(x) \geqslant \frac{2x}{\pi}$ sur $[0;\frac{\pi}{2}]$



$\mathrm{ch}(x) \leqslant \frac{\mathrm{ch}(a)-1}{a}x+1$ sur $[0;1]$

2) Inégalités de moyennes

Soient $x_1, \dots, x_n$ des réels strictement positifs, la fonction ln étant concave, on a :

$$\ln\left(\frac{x_1 + \dots + x_n}{n}\right) \geqslant \frac{\ln(x_1) + \dots \ln(x_n)}{n} = \frac{1}{n} \ln\left(\prod_{i=1}^{n} x_i\right)$$

en passant à l'exponentielle on obtient une inégalité entre la moyenne géométrique et la moyenne arithmétique :

$$\sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{x_1 + \cdots x_n}{n}$$

en remplaçant x_i par son inverse, inégalité entre la moyenne harmonique et la moyenne géométrique :

$$\frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}} \leqslant \sqrt[n]{\prod_{i=1}^{n} x_i}$$

et finalement :

$$\boxed{\frac{n}{\frac{1}{x_1} + \dots + \frac{1}{x_n}} \leqslant \sqrt[n]{\prod_{i=1}^{n} x_i} \leqslant \frac{x_1 + \cdots x_n}{n}}$$

3) Inégalités de Hölder et Minkowski

Soient u, v, p et q strictement positifs avec $\frac{1}{p} + \frac{1}{q} = 1$, alors la conacvité de la fonction ln permet d'écrire : $\ln(\frac{1}{p}u^p + \frac{1}{q}v^q) \geqslant \frac{1}{p}\ln(u^p) + \frac{1}{a}\ln(v^q) = \ln(uv)$, en passant à l'expoentielle on obtient :

$$\boxed{uv \leqslant \frac{1}{p}u^p + \frac{1}{q}v^q}$$

Soient $a_i, \dots, a_n,\ b_1, \dots, b_n$ strictement positifs, $p,\ q > 0$ tels que $\frac{1}{p} + \frac{1}{q} = 1$. Soit $A = \left(\sum_{i=1}^{n} a_i^p\right)^{1/p}$ et $B = \left(\sum_{i=1}^{n} b_i^q\right)^{1/q}$, on a $\frac{a_i}{A}\frac{b_i}{B} \leqslant \frac{1}{p}\frac{a_i^p}{A^p} + \frac{1}{q}\frac{b_i^q}{B^q}$, en sommant de 1 à n on obtient : $\frac{\sum_{i=1}^{n} a_i b_i}{AB} \leqslant \frac{1}{p} + \frac{1}{q} = 1$, d'où $\sum_{i=1}^{n} a_i b_i \leqslant AB$, c'est à dire :

$$\boxed{\sum_{i=1}^{n} a_i b_i \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} b_i^q\right)^{1/q} \quad \text{(inégalité de Hölder)}}$$

Soient $a_i, \dots, a_n,\ b_1, \dots, b_n$ strictement positifs, $p > 1$, on pose $q = \frac{1}{1 - \frac{1}{p}} = \frac{p}{p-1}$ de tel sorte que $q > 0$ et $\frac{1}{p} + \frac{1}{q} = 1$. L'inégalité de Hölder donne $\sum_{i=1}^{n} a_i(a_i + b_i)^{p-1} \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} \left(\sum_{i=1}^{n} (a_i + b_i)^{(p-1)q}\right)^{1/q}$, de même, on a $\sum_{i=1}^{n} b_i(a_i + b_i)^{p-1} \leqslant \left(\sum_{i=1}^{n} b_i^p\right)^{1/p} \left(\sum_{i=1}^{n} (a_i + b_i)^{(p-1)q}\right)^{1/q}$, en ajoutant ces deux inégalités, obtient $\sum_{i=1}^{n} (a_i + b_i)^p \leqslant \left[\left(\sum_{i=1}^{n} a_i^p\right)^{1/p} + \left(\sum_{i=1}^{n} b_i^p\right)^{1/p}\right] \left(\sum_{i=1}^{n} (a_i + b_i)^{(p-1)q}\right)^{1/q}$, c'est à dire :

$$\sum_{i=1}^{n} (a_i + b_i)^p \leqslant \left[\left(\sum_{i=1}^{n} a_i^p\right)^{1/p} + \left(\sum_{i=1}^{n} b_i^p\right)^{1/p}\right] \left(\sum_{i=1}^{n} (a_i + b_i)^p\right)^{1-1/p}$$

d'où :

$$\boxed{\left(\sum_{i=1}^{n} (a_i + b_i)^p\right)^{1/p} \leqslant \left(\sum_{i=1}^{n} a_i^p\right)^{1/p} + \left(\sum_{i=1}^{n} b_i^p\right)^{1/p} \quad \text{(inégalité de Minkowski)}}$$

IV) Exercices

★Exercice 28.1

Soit $\mathscr{C}$ l'ensemble des fonctions convexes sur l'intervalle I. Étudier la structure de cet ensemble pour les opérations usuelles sur les fonctions.

★Exercice 28.2

Soit $f : I \to \mathbb{R}$ une fonction convexe sur l'intervalle **ouvert** I. Montrer que si f possède un maximum, alors f est constante. Ceci reste-t'il vrai si I n'est plus supposé ouvert ?

★Exercice 28.3

Soit f une fonction convexe sur $\mathbb{R}$ et majorée. Montrer que f est constante. Ceci reste-t'il vrai sur un intervalle I différent de $\mathbb{R}$?

★Exercice 28.4

a) Montrer que la fonction $f : x \mapsto \ln(1+\frac{1}{x})$ est convexe sur $]0;1]$, et que $x \mapsto \sin(x)$ est concave sur $]0;\frac{\pi}{2}]$. En déduire que la fonction $g : x \mapsto \ln(1+\frac{1}{\sin(x)})$ est convexe sur $]0;\frac{\pi}{2}]$.

b) Soit (ABC) un triangle non aplati, on note a l'angle géométrique au sommet A (respectivement b et c). Montrer que :

$$\left(1+\frac{1}{\sin(\frac{a}{2})}\right)\left(1+\frac{1}{\sin(\frac{b}{2})}\right)\left(1+\frac{1}{\sin(\frac{c}{2})}\right) \geqslant 27$$

Quand a-t-on l'égalité ?