Chapitre 15

Fonctions convexes.

On se donne dans ce chapitre un intervalle I non vide et non réduit à un point. Le but de ce chapitre est de donner une définition plus rigoureuse des fonctions convexes (même si l'interprétation géométrique sera la même), de nous intéresser aux fonctions convexes générales (y compris non dérivables, que nous avons lâchement laissées de côté, les pauvres...) et de démontrer enfin les résultats admis au chapitre 2. Rappelons que l'on note $\overset{o}{I}$ l'intérieur de I, c'est-à-dire l'intervalle I privé de ses bornes éventuelles (ou, ce qui revient au même, l'intervalle ouvert ayant les mêmes bornes que I).

I Généralités

I.1 Définitions

Définition. Soit $f : I \longrightarrow \mathbb{R}$.

- f est convexe si :

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

- f est concave si $-f$ est convexe, c'est-à-dire si :

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

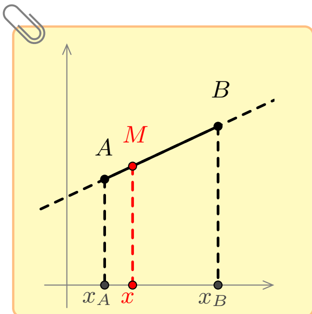
I.2 Interprétation géométrique

Soient $A(x_A, y_A)$ et $B(x_B, y_B)$ deux points du plan avec $x_A < x_B$. Soit $M(x,y)$ un point du plan. Alors :

$$M \in (AB) \qquad \iff \qquad \exists \lambda \in \mathbb{R}, \quad \left\{ \begin{array}{l} x = \lambda x_A + (1-\lambda) x_B \\ y = \lambda y_A + (1-\lambda) y_B \end{array} \right. .$$

C'est l'équation paramétrique de la droite (AB) (cf. programme de terminale). On suppose à présent que $M \in (AB)$. Il existe donc $\lambda \in \mathbb{R}$ tel que :

$$\left\{ \begin{array}{l} x = \lambda x_A + (1-\lambda) x_B \\ y = \lambda y_A + (1-\lambda) y_B \end{array} \right. .$$



Cherchons une CNS pour que $M \in [AB]$. Puisque M appartient déjà à la droite (AB), il appartient à $[AB]$ si et seulement si $x_A \leqslant x \leqslant x_B$. Dès lors,

$$M \in [AB] \quad \iff \quad \left\{ \begin{array}{l} x_A \leqslant \lambda x_A + (1-\lambda) x_B \\ \lambda x_A + (1-\lambda) x_B \leqslant x_B \end{array} \right. \quad \iff \quad \left\{ \begin{array}{c} (1-\lambda)(x_A - x_B) \leqslant 0 \\ \lambda (x_A - x_B) \leqslant 0 \end{array} \right. .$$

Finalement, puisque $x_A - x_B < 0$, alors $M \in [AB]$ si et seulement si $1 - \lambda \geqslant 0$ et $\lambda \geqslant 0$. En d'autres termes :

$$M \in [AB] \qquad \iff \qquad \exists \lambda \in [\,0\,;1\,], \qquad \left\{ \begin{array}{l} x = \lambda x_A + (1-\lambda) x_B \\ y = \lambda y_A + (1-\lambda) y_B \end{array} \right. .$$

En particulier on a le :

Théorème. Soient $x < y$ deux réels et $z \in \mathbb{R}$. Alors :

$$z \in [\,x\,;y\,] \iff \exists \lambda \in [\,0\,;1\,], z = \lambda x + (1-\lambda) y.$$

Démonstration. Appliquer ce qui précède à $M(z,0)$, $A(x,0)$ et $B(y,0)$: exo.

Remarques :

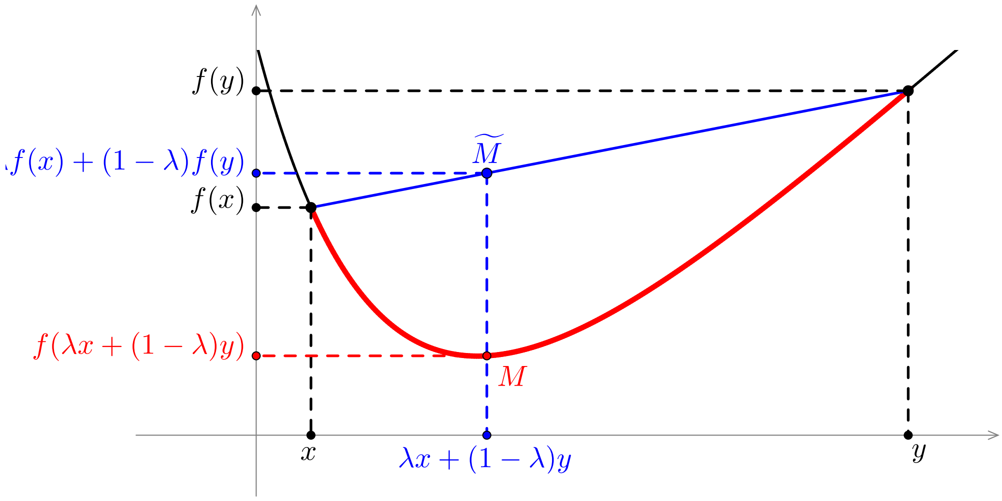
- Si $\lambda = 0$, alors $z = y$.
- Si $\lambda = 1$, alors $z = x$.
- Si $\lambda \neq 0, 1$, alors $z \in \,]\, x\,;y\,[$.

C'est pour cela que, quand on veut prouver qu'une fonction est convexe, il n'y a rien à prouver pour $\lambda = 0$ ou $\lambda = 1$, car l'inégalité de la définition devient alors $f(x) \leqslant f(x)$ ou $f(y) \leqslant f(y)$.

Interprétation géométrique de la convexité :

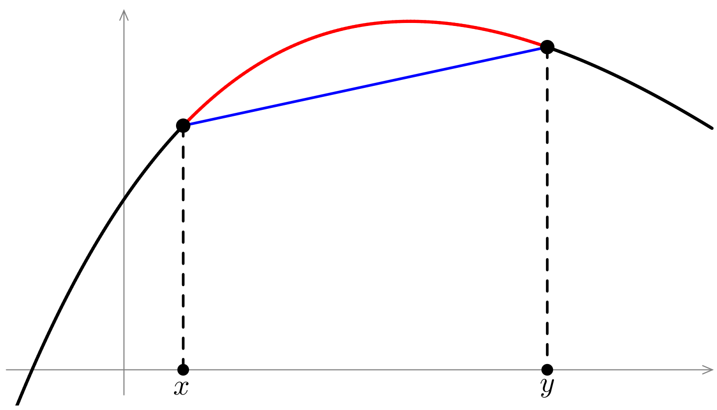
- D'après ce qui précède, les réels de la forme $\lambda x + (1-\lambda)y$, quand λ décrit $[\,0\,;1\,]$, sont tous les réels de l'intervalle $[\,x\,;y\,]$, et les réels de la forme $f(\lambda x + (1-\lambda)y)$ sont leurs images par f. Par conséquent, l'ensemble des points $M\begin{pmatrix} \lambda x + (1-\lambda)y \\ f(\lambda x + (1-\lambda)y) \end{pmatrix}$, pour $\lambda \in [\,0\,;1\,]$, est l'arc de la courbe $\mathscr{C}_f$ (en rouge ci-dessous) joignant les points d'abscisses x et y.
- L'ensemble des points $\widetilde{M}\begin{pmatrix} \lambda x + (1-\lambda)y \\ \lambda f(x) + (1-\lambda)f(y) \end{pmatrix} = \lambda \begin{pmatrix} x \\ f(x) \end{pmatrix} + (1-\lambda)\begin{pmatrix} y \\ f(y) \end{pmatrix}$, pour $\lambda \in [\,0\,;1\,]$, est le segment (appelé « corde », en bleu ci-dessous) joignant les points de $\mathscr{C}_f$ d'abscisses x et y.
- Ainsi, l'inégalité de convexité signifie que pout tout choix de x et de y dans I et de λ dans $[\,0\,;1\,]$, le point M est toujours en dessous du point $\widetilde{M}$ (car il a la même abscisse mais une ordonnée inférieure), c'est-à-dire que l'arc de courbe joignant les points d'abscisses x et y est situé en-dessous de la corde joignant ces mêmes points

Une fonction convexe est une fonction « qui sourit ».

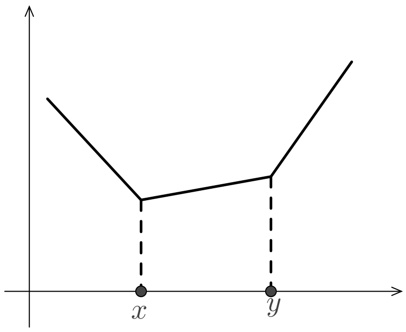


Morale de l'histoire : Une fonction est convexe lorsque son graphe est situé « en dessous de ses cordes ». De même, une fonction est concave lorsque son graphe est situé « au-dessus de ses cordes ».

Une fonction concave est une fonction « qui fait la tête ».

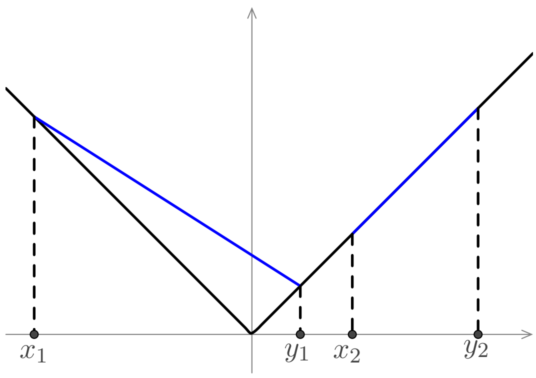


Remarque : Attention, les inégalités sont larges (le graphe peut être confondu avec la corde).



Exemples :

- La valeur absolue est convexe. Ci-dessous le graphe avec deux cordes : on voit que le graphe est en-dessous (au sens large) à chaque fois.



  En effet, soient $(x, y) \in \mathbb{R}^2$ et $\lambda \in [\,0\,;1\,]$. D'après l'inégalité triangulaire,

$$|\lambda x + (1-\lambda)y| \leqslant |\lambda x| + |(1-\lambda)y| \leqslant \lambda|x| + (1-\lambda)|y|$$

  car λ et $1-\lambda$ sont positifs. Ainsi, la valeur absolue est convexe.

- Les fonctions affines sont à la fois convexes et concaves. Ce sont même les seules : si une fonction f est convexe et concave, alors elle est au-dessus et en dessous de la corde, donc est confondue avec la corde, donc est affine.

I.3 Premières propriétés

Proposition. Soient $n \geqslant 1, f_1, \dots, f_n$ convexes sur I et $\alpha_1, \dots, \alpha_n$ des réels **positifs**. Alors $\alpha_1 f_1 + \dots + \alpha_n f_n$ est convexe sur I. En d'autres termes, une CL à coefficients positifs de fonctions convexes est une fonction convexe.

Démonstration. Soient $(x, y) \in I^2$ et $\lambda \in [\,0\,;1\,]$. Soit $i \in [\![\,1\,;\,n\,]\!]$. La fonction f_i étant convexe,

$$f_i(\lambda x + (1-\lambda)y) \leqslant \lambda f_i(x) + (1-\lambda) f_i(y)$$

De plus, $\alpha_i \geqslant 0$ donc $\alpha_i f_i(\lambda x + (1-\lambda)y) \leqslant \lambda \alpha_i f_i(x) + (1-\lambda)\alpha_i f_i(y)$. Par somme :

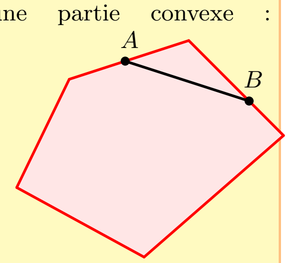
$$\left(\sum_{i=1}^n \alpha_i f_i\right)(\lambda x + (1-\lambda y)) \leqslant \lambda \left(\sum_{i=1}^n \alpha_i f_i\right)(x) + (1-\lambda)\left(\sum_{i=1}^n \alpha_i f_i\right)(y)$$

□

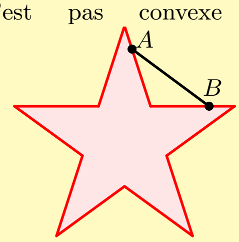
ce qui est le résultat voulu.

Une fonction convexe n'est pas forcément monotone ni dérivable. Cependant, nous verrons dans le paragraphe II.2 qu'une fonction convexe ne peut pas être « trop irrégulière ».

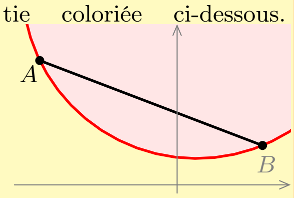
Pourquoi « convexe » ? Une partie C du plan est dite « convexe » si, pour tous points A et B de C, le segment $[AB]$ est inclus dans C. Ci-dessous une partie convexe :



Ci-dessous une partie du plan qui n'est pas convexe :



Si $f : I \longrightarrow \mathbb{R}$ est une fonction, on appelle épigraphe l'ensemble $\text{épi}(f) = \{(x, y) \in I \times \mathbb{R} \mid y \geqslant f(x)\}$. C'est l'ensemble des points du plan situés au-dessus du graphe de f : partie coloriée ci-dessous.



Alors on peut montrer (cf. exercice 7) que f est une fonction convexe si et seulement si son épigraphe est une partie convexe du plan. D'où le nom de « fonction convexe »

Corollaire. Une somme de fonctions convexes est convexe.

La proposition ci-dessus et le corollaire ci-contre sont encore valables avec des fonctions concaves.

Remarques :

- Le résultat de la proposition est faux si les coefficients ne sont pas positifs ! En effet, si f est convexe et si $\alpha < 0$ alors αf est concave donc n'est pas convexe (sauf si f est affine).
- Attention, un produit, un quotient, une composée de fonctions convexes n'est pas forcément convexe. Voir des contre-exemples dans la suite du cours.

Théorème (Inégalité de Jensen). Soient $f : I \longrightarrow \mathbb{R}$ convexe, $n \geqslant 2$, $(\lambda_1, \dots, \lambda_n)$ des réels **positifs de somme** 1, et $(x_1, \dots, x_n) \in I^n$. Alors :

$$f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leqslant \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n).$$

En particulier, si $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1/n$, alors

$$f\left(\frac{x_1 + \cdots + x_n}{n}\right) \leqslant \frac{f(x_1) + \cdots + f(x_n)}{n}.$$

Ce théorème est toujours valable pour une fonction concave, mais en changeant le sens des inégalités.

Démonstration. Par récurrence sur n.

- Si $n \geqslant 2$, on note H_n : « Pour tous $(\lambda_1, \dots, \lambda_n)$ positifs de somme 1, et $(x_1, \dots, x_n) \in I^n$, on a : $f(\lambda_1 x_1 + \cdots + \lambda_n x_n) \leqslant \lambda_1 f(x_1) + \cdots + \lambda_n f(x_n)$. »
- H_2 est vraie par définition d'une fonction convexe. En effet, si λ_1 et λ_2 sont positifs de somme 1 alors $\lambda_1 \in [\,0\,;1\,]$ et $\lambda_2 = 1 - \lambda_1$ et on a le résultat voulu.
- Soit $n \geqslant 2$. Supposons H_n vraie et montrons que H_{n+1} est vraie. Soient donc $\lambda_1, \dots, \lambda_{n+1}$ positifs de somme 1 et $(x_1, \dots, x_{n+1}) \in I^{n+1}$. Supposons qu'il existe j tel que $\lambda_j = 0$. Alors

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) = f\left(\sum_{i \neq j} \lambda_i x_i\right)$$

puisque $\lambda_j = 0$. Or, la somme de droite contient n termes, les $(\lambda_i)_{i \neq j}$ sont positifs de somme 1 et les x_i appartiennent à I : on peut donc appliquer l'hypothèse de récurrence, d'où :

$$f\left(\sum_{i=1}^{n+1} \lambda_i x_i\right) \leqslant \sum_{i \neq j} \lambda_i f(x_i) = \sum_{i=1}^{n+1} \lambda_i f(x_i)$$

et le résultat est démontré. On suppose à présent que tous les λ_i sont non nuls (et donc strictement positifs). On veut appliquer l'hypothèse de récurrence, donc on veut se ramener à n réels. Le problème avec $\lambda_1, \dots, \lambda_n$ est que leur somme n'est pas égale à 1. Pour que leur somme soit égale à 1, il suffit de diviser par leur somme. Posons donc $S_n = \displaystyle\sum_{i=1}^{n} \lambda_i$. Puisque $(\lambda_1, \dots, \lambda_{n+1})$ sont positifs (strictement d'après ce qui précède) de somme 1 alors $S_n = 1 - \lambda_{n+1} \in \,]\,0\,;1\,[$. On pose

$$y = \frac{\lambda_1 x_1 + \cdots + \lambda_n x_n}{S_n}$$

Dès lors,

$$\begin{aligned}
\sum_{i=1}^{n+1} \lambda_i x_i &= S_n \times \frac{\lambda_1 x_1 + \cdots + \lambda_n x_n}{S_n} + \lambda_{n+1} x_{n+1} \\
&= S_n y + (1 - S_n) x_{n+1}
\end{aligned}$$

La fonction f étant convexe,

$$f\left(\sum_{i=1}^{n+1}\lambda_i x_i\right) \leqslant S_n f(y) + (1-S_n)f(x_{n+1}) = S_n f(y) + \lambda_{n+1} f(x_{n+1})$$

Pour tout $i \in [\![1\,;\, n]\!]$, posons $\mu_i = \dfrac{\lambda_i}{S_n}$ si bien que

$$\begin{aligned} y &= \frac{\lambda_1}{S_n}x_i + \cdots + \frac{\lambda_n}{S_n}x_n \\ &= \mu_1 x_1 + \cdots + \mu_n x_n \end{aligned}$$

Les μ_i sont positifs et de somme 1 (et au nombre de n) donc, par hypothèse de récurrence,

$$f(y) = f\left(\sum_{i=1}^{n}\mu_i x_i\right) \leqslant \sum_{i=1}^{n}\mu_i f(x_i) = \sum_{i=1}^{n}\frac{\lambda_i}{S_n} f(x_i)$$

Finalement,

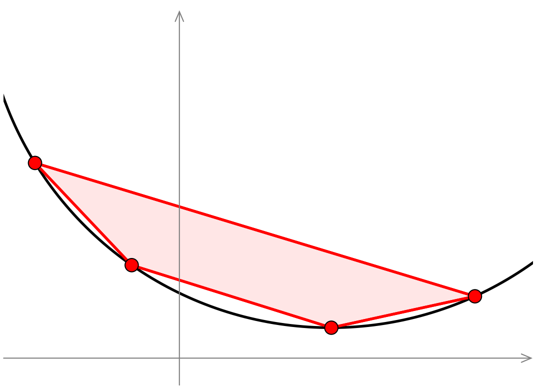
$$\begin{aligned} f\left(\sum_{i=1}^{n+1}\lambda_i x_i\right) &\leqslant S_n \times \sum_{i=1}^{n}\frac{\lambda_i}{S_n}f(x_i) + \lambda_{n+1}f(x_{n+1}) \\ &\leqslant \sum_{i=1}^{n}\lambda_i f(x_i) + \lambda_{n+1}f(x_{n+1}) \\ &\leqslant \sum_{i=1}^{n+1}\lambda_i f(x_i) \end{aligned}$$

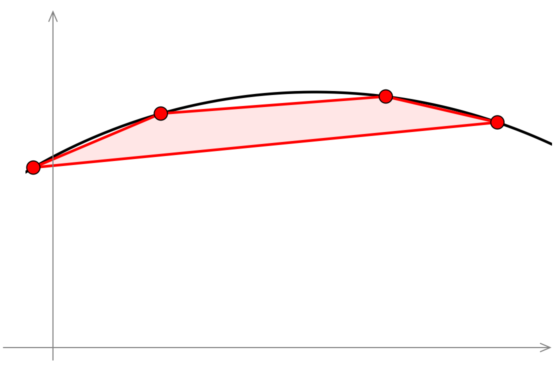
□

c'est-à-dire que H_{n+1} est vraie.

- D'après le principe de récurrence, H_n est vraie pour tout $n \geqslant 2$.

Interprétation géométrique : Si on choisit des points sur le graphe de f, le polygone qu'ils forment est au-dessus du graphe pour une fonction convexe, et en dessous pour une fonction concave.





Exemple : Comparaison des moyennes géométrique et arithmétique. Soient $x_1, \dots, x_n$ des réels strictement positifs. Montrer à l'aide de la concavité de la fonction ln que

$$\frac{x_1 + \cdots + x_n}{n} \geqslant (x_1 \cdots x_n)^{1/n}$$

Vos professeurs sont vraiment très gentils : parmi les deux formules possibles de la moyenne, ils prennent celle qui vous donne la moyenne la plus élevée !

Appliquons l'inégalité de Jensen à $x_1, \dots, x_n$ et à la fonction ln qui est concave sur $\mathbb{R}_+^*$. Les $1/n$ étant positifs de somme 1,

$$\ln\left(\frac{x_1 + \cdots + x_n}{n}\right) \geqslant \frac{\ln(x_1) + \cdots + \ln(x_n)}{n} = \ln\left((x_1 \cdots x_n)^{1/n}\right).$$

Nous le prouverons dans le paragraphe II.3.

La fonction exponentielle étant croissante, on en déduit :

$$\frac{x_1 + \cdots + x_n}{n} \geqslant (x_1 \cdots x_n)^{1/n}$$

II Pentes, dérivées, tangentes

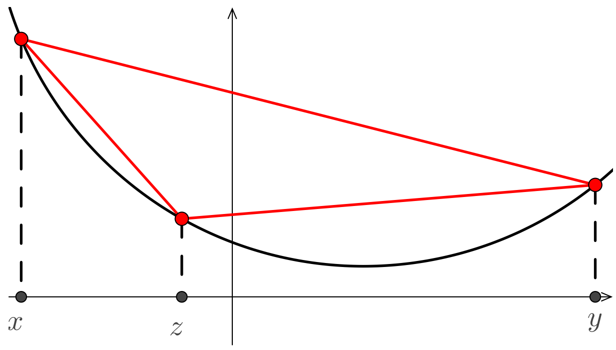
II.1 Pentes

Théorème (Théorème des trois pentes).
Soit $f : I \to \mathbb{R}$ convexe. Soient $x < z < y \in I$. Alors :

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

Le résultat est toujours valable en changeant le sens des inégalités entre les pentes.

On retrouve facilement le résultat à l'aide d'un dessin, sur lequel on voit quelles sont les pentes les plus fortes : quelles sont les pentes que l'on préfère prendre à vélo ?



Démonstration. $z \in \left]x\,;y\right[$ donc il existe $\lambda \in \left]0\,;1\right[$ tel que $z = \lambda x + (1-\lambda)y$. La fonction f est convexe donc

$$f(z) \leqslant \lambda f(x) + (1-\lambda) f(y)$$

d'où $f(z) - f(y) \leqslant \lambda(f(x) - f(y))$. Or, $z = \lambda x + (1-\lambda)y$ donc $z - y = \lambda(x - y) < 0$. Ainsi (et on n'oublie pas de changer le sens de l'inégalité),

$$\frac{f(z) - f(y)}{z - y} \geqslant \frac{\lambda(f(x) - f(y))}{\lambda(x - y)} = \frac{f(y) - f(x)}{y - x}$$ □

De même pour l'autre inégalité.

Proposition (Caractérisation de la convexité par la croissance des cordes). Soit $f : I \to \mathbb{R}$. La fonction f est convexe si et seulement si, pour tout $a \in I$, la fonction

$$\tau_a : \left\{\begin{array}{ccc} I\backslash\{a\} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{f(x) - f(a)}{x - a} \end{array}\right.$$

est croissante.

Le résultat est toujours valable pour une fonction concave en remplaçant « croissante » par « décroissante ».

Démonstration. Supposons f convexe. Soit $a \in I$. Montrons que τ_a est croissante. Soient $x_1 < x_2 \in I \backslash \{a\}$. Supposons que $x_1 < x_2 < a$. Appliquons le théorème des trois pentes à $x = x_1, z = x_2, y = a$:

$$\frac{f(x_2) - f(a)}{x_2 - a} \geqslant \frac{f(x_1) - f(a)}{x_1 - a}$$

c'est-à-dire $\tau_a(x_2) \geqslant \tau_a(x_1)$. De même dans les deux autres cas ($x_1 < a < x_2$ et $a < x_1 < x_2$). En d'autres termes, τ_a est croissante.

Attention, $I \backslash \{a\}$ n'est pas un intervalle ! Même si f était dérivable (ce qui n'est pas le cas), on ne pourrait pas utiliser le signe de ${\tau_a}'$ pour donner la monotonie de τ_a car le critère reliant monotonie et signe de la dérivée n'est valable que sur un intervalle. Par conséquent, pour prouver la croissance de τ_a, un seul moyen : la définition.

Réciproquement : on suppose que pour tout $a \in I, \tau_a$ est croissante sur $I \backslash \{a\}$. Montrons que f est convexe : soient $(x, y) \in I^2$ et $\lambda \in [\,0\,;1\,]$. Si $x = y$ ou si $\lambda = 0$ ou 1 alors $f(\lambda x + (1-\lambda)y) = \lambda f(x) + (1-\lambda)f(y)$ et en particulier

$$f(\lambda x + (1-\lambda)y) \leqslant \lambda f(x) + (1-\lambda)f(y)$$

Supposons donc $x \neq y$ et $\lambda \in \,]\,0\,;1\,[$. Sans perte de généralité, on peut supposer $x < y$ (en effet, le raisonnement est analogue si $x > y$). Soit $z = \lambda x + (1-\lambda)y \in \,]\,x\,;y\,[$. La fonction τ_x est croissante sur $I \backslash \{x\}$ donc $\tau_x(z) \leqslant \tau_x(y)$ c'est-à-dire

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

Or, $z - x > 0$ donc (on ne change pas le sens de l'inégalité en multipliant par $z - x$)

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

Rappelons qu'on ne peut pas faire de produit en croix dans les inégalités. Ce qu'on peut faire, en revanche, c'est multiplier par $z - x$ et par $y - x$ en étudiant le signe à chaque fois pour voir si les inégalités changent de sens.

De plus, $z = \lambda x + (1-\lambda)y$, donc $z - x = (1-\lambda)(y - x)$, si bien que

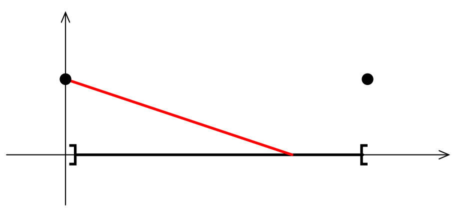
$$f(z) \leqslant f(x) + (1-\lambda)(f(y) - f(x)) = \lambda f(x) + (1-\lambda)f(y) \qquad \square$$

La fonction f est donc convexe.

II.2 Régularité des fonctions convexes.

Tous les résultats de ce paragraphe sont aussi valables pour les fonctions concaves.

Remarque : Nous avons vu plus haut qu'une fonction convexe n'est pas forcément dérivable (par exemple, la valeur absolue n'est pas dérivable alors qu'elle est convexe). Une fonction convexe n'est pas non plus forcément continue. Ci-dessous le graphe d'une fonction convexe non continue sur $[\,0\,;4\,]$: plus précisément, c'est la fonction nulle sur $]\,0\,;4\,[$ qui vaut 1 en 0 et en 4. Elle est discontinue en 0 et en 4, mais est tout de même convexe sur $[\,0\,;4\,]$: le graphe est bien en-dessous de ses cordes !



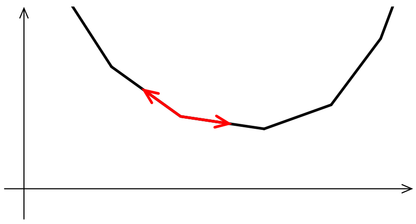
Cependant, on peut prouver qu'une fonction convexe est tout de même assez régulière.

C'est d'ailleurs fou qu'une fonction ayant vérifiant une propriété géométrique aussi simple ait de telles propriétés de régularité !

Proposition. Une fonction convexe est dérivable à gauche et à droite en tout point intérieur à son domaine de définition.

Remarque : Comme on peut le voir sur le dessin ci-contre, si f est convexe, même en les points en lesquels f n'est pas dérivable, f est dérivable à gauche et à droite.

Attention, une fonction dérivable à droite et à gauche n'est pas forcément dérivable (mais ça marche avec les fonctions continues) ! Une fonction est dérivable en a...



*... ssi f est dérivable à droite et à gauche **et** si ${f_d}'(a) = {f_g}'(a)$.*

Démonstration. Soit $f : I \to \mathbb{R}$ convexe. Soit $a \in \overset{o}{I}$. D'après le paragraphe précédent, la fonction τ_a est croissante sur $I\backslash\{a\}$. En particulier, elle est croissante sur $I \cap\,]-\infty\,;a\,[$ (qui est non vide car a est un point intérieur). De plus, a étant un point intérieur, il existe $b \in I$ strictement supérieur à a. La fonction τ_a étant croissante, $\tau_a(x) \leqslant \tau_a(b)$ pour tout $x \in I, x < a$. En conclusion, τ_a est croissante majorée sur $I \cap\,]-\infty\,;a\,[$ donc admet une limite finie en a^- : f est dérivable à gauche en a. De même, f est dérivable à droite en a.

Plus fort : si $x < a < y$ alors, par croissance de τ_a, $\tau_a(x) \leqslant \tau_a(y)$. Si on fait tendre x vers a^- et y vers a^+, l'inégalité large passant à la limite, ${f_g}'(a) \leqslant {f_d}'(a)$. On peut en déduire que f est dérivable sauf en un nombre dénombrable de points, cf. deuxième année.

Corollaire. Une fonction convexe est continue sur l'intérieur de son domaine de définition. En particulier, une fonction convexe sur un intervalle ouvert est continue.

Cas particulier important : Une fonction convexe définie sur $\mathbb{R}$ est continue.

Démonstration. Si f est convexe, alors elle est dérivable à droite et à gauche en tout point intérieur de son domaine de définition, donc est continue à droite et à gauche, donc continue, en tout point intérieur.

Comme on l'a vu plus haut, une fonction convexe sur un intervalle I n'est pas forcément continue sur I. Cependant, comme dit ci-contre, elle est continue sur $\overset{o}{I}$. On peut même montrer qu'elle est localement lipschitzienne sur l'intérieur de son domaine de définition, c'est-à-dire que pour tout $x \in \overset{o}{I}$, f est lipschitzienne sur un voisinage de x.

II.3 Dérivées

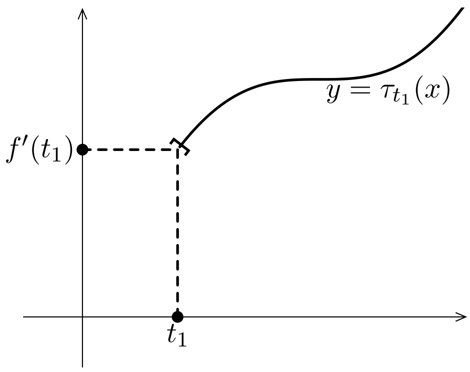
Proposition. Soit $f : I \longrightarrow \mathbb{R}$ dérivable.

- f est convexe si et seulement si f' est croissante.
- f est concave si et seulement si f' est décroissante.

Démonstration. Supposons f convexe. Soient $t_1 \leqslant t_2 \in I$. Si $t_1 = t_2$, on a évidemment $f'(t_1) \leqslant f'(t_2)$. Supposons à présent $t_1 < t_2$. Soit $x \in\,]\,t_1\,;t_2\,[$.

L'application τ_{t_1} est croissante sur $]\,t_1\,;t_2\,[$ (car f est convexe, cf. paragraphe II.1) et

$$\tau_{t_1}(x) = \frac{f(x) - f(t_1)}{x - t_1} \xrightarrow[x \to {t_1}^+]{} f'(t_1)$$



Or, une fonction croissante sur $]\,t_1\,;t_2\,[$ est supérieure à sa limite en t_1 (voir le dessin ci-contre) donc

$$f'(t_1) \leqslant \frac{f(x) - f(t_1)}{x - t_1}$$

De même, $\dfrac{f(t_2) - f(x)}{t_2 - x} \leqslant f'(t_2)$. Or, τ_x est croissante et $t_1 < t_2$ donc

$$\tau_x(t_1) = \frac{f(t_1) - f(x)}{t_1 - x} \leqslant \tau_x(t_2) = \frac{f(t_2) - f(x)}{t_2 - x}$$

Finalement, à l'aide des trois inégalités ci-dessus, il vient : $f'(t_1) \leqslant f'(t_2)$: f' est croissante.

Réciproquement, supposons f' croissante. Soient $(x, y) \in I^2$ et $\lambda \in [\,0\,;1\,]$. Comme précédemment, on peut supposer $\lambda \in\,]\,0\,;1\,[$ et $x < y$. Soit $z = \lambda x + (1-\lambda)y \in\,]\,x\,;y\,[$. La fonction

f est continue sur $[x;z]$ et dérivable sur $]x;z[$: d'après l'égalité des accroissements finis, il existe $c_1 \in]x;z[$ tel que

$$f'(c_1) = \frac{f(z)-f(x)}{z-x}$$

De même, il existe $c_2 \in]z;y[$ tel que

$$f'(c_2) = \frac{f(y)-f(z)}{y-z}$$

En particulier, $c_1 \leqslant c_2$: f' étant croissante, $f'(c_1) \leqslant f'(c_2)$, donc

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

Or, $z-x>0$ et $y-z>0$ donc on a successivement

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

puis

$$(y-z)(f(z)-f(x)) \leqslant (z-x)(f(y)-f(x))$$

c'est-à-dire $f(z)(y-x) \leqslant f(y)(z-x)+f(x)(y-z)$. Or, $z=\lambda x+(1-\lambda)y$ donc $z-x = (1-\lambda)(y-x)$ et $y-z=\lambda(y-x)$ donc

$$f(z)(y-x) \leqslant f(y)\times(1-\lambda)(y-x)+f(x)\times\lambda(y-x) \qquad \square$$

Il suffit de simplifier par $(y-x)>0$ (et donc on ne change pas le sens de l'inégalité) pour conclure : f est convexe. D'où la première équivalence.

Pour finir, démontrons la deuxième équivalence : f est concave $\iff$ $-f$ est convexe $\iff$ $-f'$ est croissante $\iff$ f' est décroissante.

Corollaire. Soit $f : I \longrightarrow \mathbb{R}$ dérivable deux fois.

- f est convexe si et seulement si f'' est positive.
- f est concave si et seulement si f'' est négative.

Remarques :

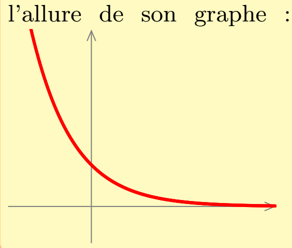
- Ici, le mot« positif » est à prendre au sens large, c'est-à-dire « positif ou nul », « supérieur ou égal à 0 ». De même pour « négatif ».
- Attention, comme on l'a déjà vu, il existe des fonctions convexes non dérivables et donc non dérivables deux fois. Pour appliquer ce corollaire, il faut donc préciser que f est dérivable deux fois !
- Ne pas confondre f'' et f', en d'autres termes, ne pas confondre une fonction convexe (qui vérifie $f'' \geqslant 0$) et une fonction croissante (qui vérifie $f' \geqslant 0$). Il n'y a aucun lien entre la convexité et la monotonie d'une fonction !
- Ce corollaire nous fournit un moyen très simple de montrer qu'une fonction est convexe : si elle est dérivable deux fois, il suffit d'étudier le signe de f''. Par conséquent, quand on se demandera si une fonction f est convexe, la première chose à faire est de se demander si elle est dérivable deux fois. Si c'est le cas, il suffit d'étudier le signe de f'', sinon : pas le choix, on revient à la définition. Attention, les fonctions convexes non dérivables tombent dans beaucoup d'exercices, il faut impérativement connaître la définition d'une fonction convexe ! Ce corollaire est très utile, mais il ne fait pas tout !

Exemples : Donnons des exemples de fonctions convexes ou concaves : on peut l'affirmer directement, mais on le montre très facilement en donnant le signe de la dérivée seconde. Remarquons que, pour certaines fonctions, cela dépend de l'intervalle sur lequel on se place (quand rien n'est précisé, il est sous-entendu qu'on se place sur le domaine de définition de la fonction). On donnera les graphes de ces fonctions au fil du chapitre, dans la marge.

Fonctions convexes :

- exp.
- Fonction carré.
- Fonction cube sur $\mathbb{R}_+$.
- Arctan sur $\mathbb{R}_-$.
- $x \mapsto x^{2p}$ pour tout $p \in \mathbb{N}$.
- tan sur $\left[0\,;\frac{\pi}{2}\right[$.
- Fonction inverse sur $\mathbb{R}_+^*$.
- $x \longmapsto e^{-x}$.

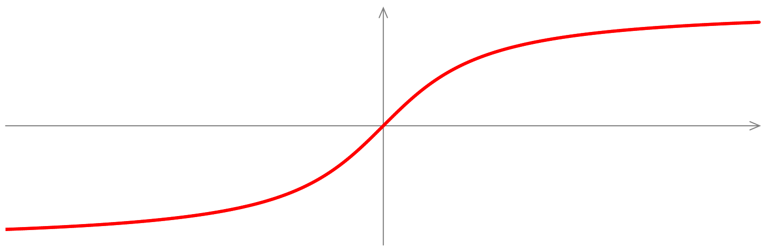
La fonction $x \longmapsto e^{-x}$ est un exemple de fonction convexe strictement décroissante : il n'y a aucun lien entre convexité et monotonie ! Ci-dessous l'allure de son graphe :



Fonctions concaves :

- ln.
- sin sur $[0\,;\pi]$.
- La racine carrée sur $\mathbb{R}_+$.
- cos sur $\left[-\frac{\pi}{2}\,;\frac{\pi}{2}\right]$.
- Fonction cube sur $\mathbb{R}_-$.
- Arctan sur $\mathbb{R}_+$.
- tan sur $\left]-\frac{\pi}{2}\,;0\right]$.
- Fonction inverse sur $\mathbb{R}_-^*$.

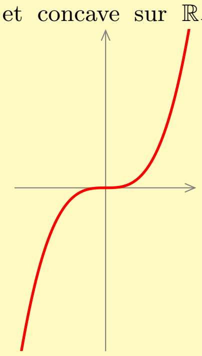
Ci-dessous le graphe de la fonction Arctan, convexe sur $\mathbb{R}_-$ et concave sur $\mathbb{R}_+$:



Remarques :

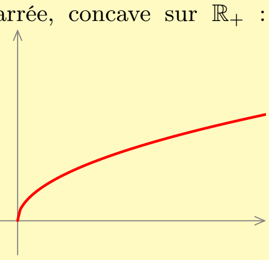
- Les fonctions $x \mapsto x$ et $x \mapsto x^2$ sont convexes sur $\mathbb{R}^-$ mais $x \mapsto x^3$ ne l'est pas : un produit de fonctions convexes n'est pas forcément convexe.
- Les fonctions $x \mapsto x^2$ et $x \mapsto e^{-x}$ sont convexes sur $\mathbb{R}$ mais $x \mapsto e^{-x^2}$ n'est pas convexe sur $\left[-\frac{1}{\sqrt{2}}\,;\frac{1}{\sqrt{2}}\right]$ (cf. exercice 17 du chapitre 2) : une composée de fonctions convexes n'est pas forcément convexe.
- D'ailleurs, pendant qu'on en parle, la fonction $x \mapsto e^{-x}$ est un exemple de fonction convexe strictement décroissante : puisqu'on vous dit qu'il n'y a aucun lien entre convexité et monotonie !

Ci-dessous le graphe de la fonction cube, convexe sur $\mathbb{R}_+$ et concave sur $\mathbb{R}_-$:



Remarque : La racine carrée est dérivable deux fois (car $\mathscr{C}^\infty$) sur $\mathbb{R}_+^*$ de dérivée seconde négative donc est concave sur $\mathbb{R}_+^*$ donc sur $\mathbb{R}_+$ car elle est continue.

Ci-dessous le graphe de la fonction racine carrée, concave sur $\mathbb{R}_+$:



Montrons plus généralement que si $f : I \longrightarrow \mathbb{R}$ est continue sur $\overset{o}{I}$ (l'ensemble I privé de ses éventuelles extrémités finies) et convexe (respectivement concave) sur $\overset{o}{I}$ alors f est convexe (respectivement concave) sur I. Cela se voit assez bien : la convexité est une notion géométrique. Si f est convexe sur l'intérieur d'un intervalle et si f est continue, ce n'est pas en rajoutant un point (en lequel f est continue) que cela changera la position relative de f par rapport aux cordes !

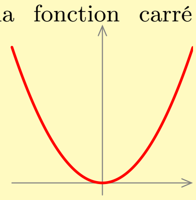
Bref, montrons ce résultat. Soient $(x,y) \in I^2$ et $\lambda \in [0\,;1]$. Alors il existe une suite $(x_n)_{n\in\mathbb{N}}$ d'éléments de $\overset{o}{I}$ qui converge vers x (si $x \in \overset{o}{I}$ alors il n'y a rien à prouver, il suffit de prendre $(x_n)_{n\in\mathbb{N}}$ constante égale à x, sinon $x = \sup I$ ou $x = \inf I$ donc il existe une suite d'éléments de $\overset{o}{I}$ qui converge vers x). De même, il existe une suite (y_n) d'éléments de $\overset{o}{I}$ qui converge vers y. La fonction f étant convexe sur $\overset{o}{I}$, pour tout $n \in \mathbb{N}$, $f(\lambda x_n + (1-\lambda)y_n) \leqslant \lambda f(x_n) + (1-\lambda) f(y_n)$. La fonction f étant continue, le membre de gauche converge, quand

n tend vers $+\infty$, vers $f(\lambda x+(1-\lambda)y)$ et le membre de droite vers $\lambda f(x)+(1-\lambda)f(y)$. L'inégalité large passe à la limite donc $f(\lambda x+(1-\lambda)y) \leqslant \lambda f(x)+(1-\lambda)f(y)$. Ainsi f est convexe sur I. De même pour les fonctions concaves.

Remarque : La fonction cube est convexe sur $\mathbb{R}_+$ et concave sur $\mathbb{R}_-$. Elle change de convexité en 0. De même pour Arctan et tan en 0 et $x \longmapsto e^{-x^2}$ en $1/\sqrt{2}$ et en $-1/\sqrt{2}$ (voir l'exercice 17 du chapitre 2). Cela justifie la définition suivante.

Définition. Soit $f : I \longrightarrow \mathbb{R}$, soit $x_0 \in \overset{o}{I}$. On dit que f admet un point d'inflexion en x_0 si f change de convexité en x_0, c'est-à-dire s'il existe $\varepsilon > 0$ tel que f soit convexe sur $[\,x_0-\varepsilon\,;x_0\,]$ et concave sur $[\,x_0\,;x_0+\varepsilon\,]$, ou concave sur $[\,x_0-\varepsilon\,;x_0\,]$ et convexe sur $[\,x_0\,;x_0+\varepsilon\,]$.

Ci-dessous le graphe de la fonction carré :



Remarque : Ainsi, si f est et dérivable deux fois, f admet un point d'inflexion en x_0 si et seulement si f'' change de signe en x_0 c'est-à-dire s'il existe $\varepsilon > 0$ tel que f'' soit positive sur $[\,x_0-\varepsilon\,;x_0\,]$ et négative sur $[\,x_0\,;x_0+\varepsilon\,]$ (ou le contraire). De plus, on a $f''(x_0) \geqslant 0$ et $f''(x_0) \leqslant 0$ (ou le contraire bien sûr) donc $f''(x_0)=0$. En conclusion :

Théorème. Soit $f : I \longrightarrow \mathbb{R}$ dérivable deux fois, soit $x_0 \in \overset{o}{I}$. f admet un point d'inflexion en x_0 si et seulement si f'' s'annule en changeant de signe en x_0.

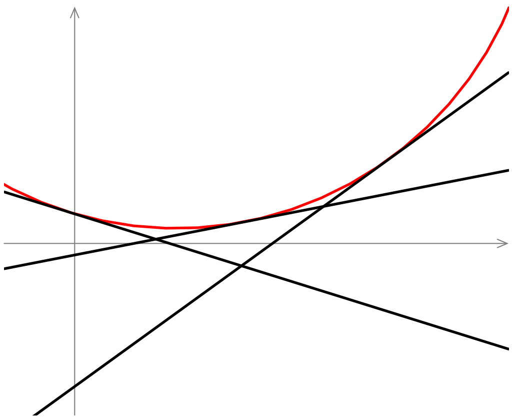
Il ne faut pas oublier de montrer que f'' change de signe : le fait que f'' s'annule ne suffit pas !

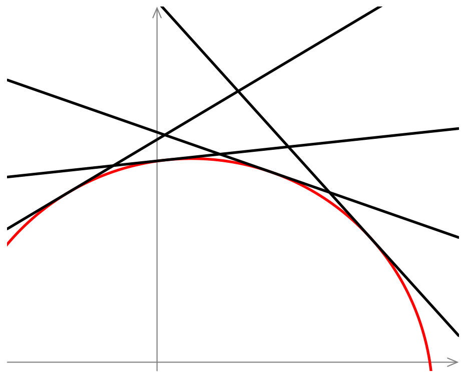
Remarque : Attention, il existe des fonctions convexes qui ne sont pas dérivable deux fois. Dans ce cas, il faut revenir à la définition d'un point d'inflexion, cf exercice 1.

II.4 Tangentes

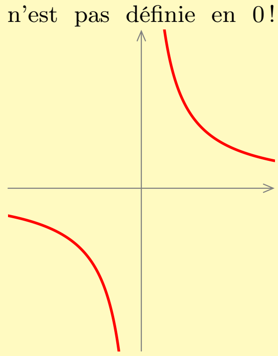
Théorème. Soit $f : I \longrightarrow \mathbb{R}$ dérivable.

- f est convexe si et seulement si son graphe est au-dessus de ses tangentes.
- f est concave si et seulement si son graphe est en dessous de ses tangentes.





Ci-dessous le graphe de la fonction inverse : elle est convexe sur $\mathbb{R}_+^*$ et concave sur $\mathbb{R}_-^*$. Attention, elle n'admet pas de point d'inflexion en 0 car n'est pas définie en 0 !



Rappelons qu'une fonction convexe ou concave est définie sur un **intervalle** ! En effet, la définition d'une fonction convexe impose que, pour tout $(x,y) \in {D_f}^2$, $[\,x\,;y\,] \subset D_f$ (sinon $\lambda x+(1-\lambda)y$ n'appartient pas forcément à D_f), c'est-à-dire (cf. chapitre 12) que D_f doit être un intervalle : parler de fonction convexe définie sur autre chose qu'un intervalle n'a aucun sens !

Démonstration. Rappelons que si $a \in I$, la tangente en a est la droite d'équation $y=f'(a)(x-a)+f(a)$. Supposons f convexe. Soit $a \in I$ et soit $x<a$. On a déjà montré plus haut (cf proposition reliant convexité et croissance de la dérivée) que $\dfrac{f(x)-f(a)}{x-a} \leqslant f'(a)$. Or, $x-a<0$ donc $f(x) \geqslant f(a)+f'(a)\times(x-a)$. De même, si $a<x$, $f'(a) \leqslant \dfrac{f(x)-f(a)}{x-a}$ donc $f(x) \geqslant f(a)+f'(a)\times(x-a)$ et le résultat est aussi vérifié si $x=a$: le graphe de f est donc au-dessus de ses tangentes.

Réciproquement, supposons que le graphe de f est situé au-dessus de toutes ses tangentes. Soient $a<b \in I$. Le graphe de f est situé au-dessus de sa tangente en a donc :

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

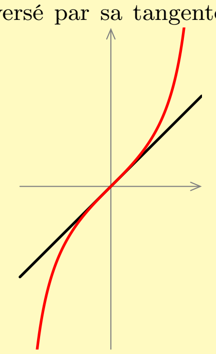
En particulier si $x = b, f(b) \geqslant f'(a)(b-a) + f(a)$. De même, pour tout $x \in I, f(x) \geqslant f'(b)(x-b) + f(b)$ et ce résultat est en particulier vrai pour $x = a$ donc $f(b) \geqslant f'(b)(a-b) + f(b)$. Par somme :

$$f(b) + f(a) \geqslant f'(a)(b-a) + f'(b)(a-b) + f(a) + f(b) \qquad \square$$

donc $0 \geqslant (f'(a) - f'(b)) \times (b-a)$. Or, $b - a > 0$ donc $0 \geqslant f'(a) - f'(b)$ ie $f'(a) \leqslant f'(b)$: f' est croissante donc f est convexe.

Remarque : Par conséquent, si f est dérivable et si f admet un point d'inflexion en x_0, le graphe de f « traverse sa tangente » en x_0.

Ci-dessous le graphe de la tangente sur $\left]-\frac{\pi}{2}\,;\frac{\pi}{2}\right[$, traversé par sa tangente :

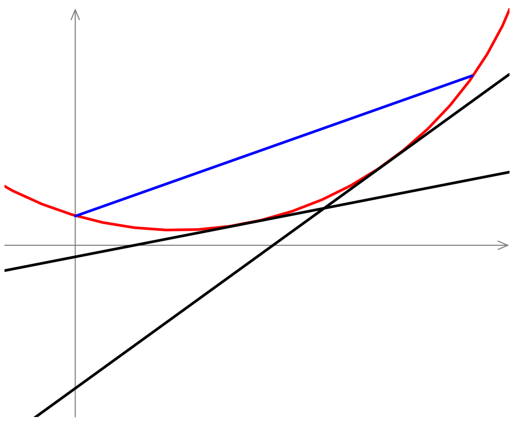


Remarque : En pratique, on se contente de la tangente en 0 dont l'équation est $y = f(0) + f'(0) \times x$. Or, $f(0) + f'(0) \times x$ est la partie principale (c'est-à-dire sans le o) du DL de f à l'ordre 1 (cf. chapitre 24). Par conséquent, si on connaît le DL de f, l'équation de sa tangente en 0 ne pose aucune difficulté.

Exemples :

- La tangente en 0 du sinus, de la tangente, de l'Arctan, de $x \longmapsto \ln(1+x)$ est la droite d'équation $y = x$.
- La tangente en 0 de l'exponentielle est la droite d'équation $y = 1 + x$.
- La tangente en 0 de $x \longmapsto (1+x)^\alpha$ (où α est un réel FIXE) est la droite d'équation $y = 1 + \alpha x$.

II.5 Bilan



f est convexe si et seulement si :

- Le graphe de f est en-dessous de ses cordes.
- Le graphe de f est au-dessus de ses tangentes (si f est dérivable).
- f' est croissante (si f est dérivable).
- f'' est positive (si f est dérivable deux fois).

Ci-contre (à gauche) le graphe d'une fonction convexe (en rouge) avec une corde (en bleu) et deux tangentes (en noir).

Il suffit de tout inverser pour les fonctions concaves.

III Questions types

Montrer les inégalités suivantes :

- $\forall x \in \left[0\,;\frac{\pi}{2}\right], \; \frac{2x}{\pi} \leqslant \sin(x) \leqslant x$.

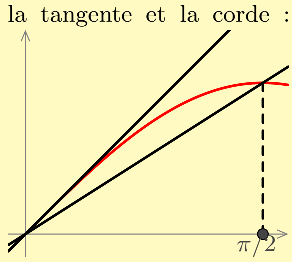
Nous avons déjà prouvé certaines de ces inégalités dans le chapitre 2.

Réponse : La fonction sin est concave sur $\left[0\,;\frac{\pi}{2}\right]$ donc est au-dessus de ses cordes et en dessous de ses tangentes. Or, la droite d'équation $y=x$ est sa tangente en 0 et la corde reliant les points d'abscisses 0 et $\frac{\pi}{2}$ est la droite d'équation

$$y=\frac{\sin\left(\frac{\pi}{2}\right)-\sin(0)}{\frac{\pi}{2}-0}\times(x-0)+\sin(0)=\frac{2x}{\pi}.$$

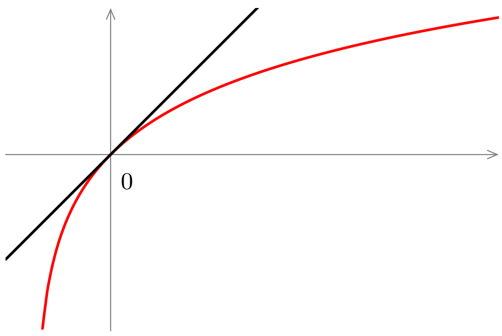
D'où le résultat.

On peut affirmer directement que le sinus est concave sur $\left[0\,;\frac{\pi}{2}\right]$, mais si on veut le montrer, il suffit de dire que sin est dérivable deux fois et $\sin''=-\sin\leqslant 0$ sur cet intervalle. Ci-dessous le graphe, la tangente et la corde :



- $\forall x>-1$, $\ln(1+x)\leqslant x$.

  Réponse : La fonction $x\longmapsto\ln(1+x)$ est concave donc est en dessous de ses tangentes, et la droite d'équation $y=x$ est sa tangente en 0. Voir le dessin ci-dessous.
  Remarque : Idem, on peut l'affirmer directement, mais si on veut le montrer, il suffit de dire que la fonction est dérivable deux fois, de dérivée seconde $x\longmapsto\dfrac{-1}{(1+x)^2}<0$.
  Ci-dessous le graphe et la tangente en 0 :



- $\forall x\in\left[0\,;\frac{\pi}{2}\right[$, $\tan(x)\geqslant x$.

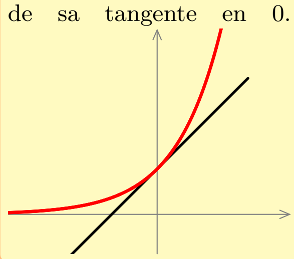
  Réponse : La fonction tan est convexe sur $\left[0\,;\frac{\pi}{2}\right[$ donc est au dessus de ses tangentes, et la droite d'équation $y=x$ est sa tangente en 0. Voir le dessin en II.4.

Remarque : Le seul problème de cette méthode est qu'elle ne permet pas de donner les cas d'égalité car les fonctions strictement convexes ne sont pas au programme. Si on demande le cas d'égalité, pas le choix, on se débrouille autrement ! Mais la bonne nouvelle est que cela n'arrive pas souvent.

Autres exemples : Soit $n\geqslant 1$. Montrer les inégalités suivantes :

1. $\forall x\in\mathbb{R}$, $e^x\geqslant 1+x$.
2. $\forall x\in[\,0\,;n\,]$, $\left(1-\dfrac{x}{n}\right)^n\leqslant e^{-x}$.
3. $\forall x\in[\,0\,;n\,]$, $\left(1-\dfrac{x^2}{n^2}\right)^n\geqslant 1-\dfrac{x^2}{n}$.
4. $\forall x\in[\,0\,;n\,]\,,0\leqslant e^{-x}-\left(1-\dfrac{x}{n}\right)^n\leqslant\dfrac{x^2}{n}e^{-x}$.

Ci-dessous le graphe de la fonction exponentielle, convexe sur $\mathbb{R}$: il est bien au-dessus de sa tangente en 0.



Réponse :

1. La fonction exponentielle est convexe donc est au-dessus de ses tangentes et la droite d'équation $y=1+x$ est la tangente en 0.
2. D'après la question précédente, pour tout $u\in\mathbb{R}$, $e^u\geqslant 1+u$. Soit $x\in[\,0\,;n\,]$. En appliquant ce résultat à $u=-\dfrac{x}{n}$, il vient $e^{-x/n}\geqslant 1-\dfrac{x}{n}$. Or, la fonction $u\longmapsto u^n$ est croissante sur $\mathbb{R}_+$ (attention, elle n'est pas croissante sur $\mathbb{R}$ si n est pair) et $x\leqslant n$ donc $\dfrac{x}{n}\leqslant 1$ donc $1-\dfrac{x}{n}\geqslant 0$. Par conséquent, on ne change pas le sens des inégalités en mettant à la puissance n donc $e^{-x}\geqslant\left(1-\dfrac{x}{n}\right)^n$.

3. Soit $f : u \longmapsto (1-u)^n$ (ce qui est la fonction de gauche, qu'on applique à x^2/n^2). Sa convexité ne saute pas aux yeux : on ne peut pas l'affirmer directement. Cependant, elle est dérivable deux fois et pour tout $u \in [\,0\,;1\,]$ (car on va l'appliquer à $u = x^2/n^2 \in [\,0\,;1\,]$), $f'(u) = -n(1-u)^{n-1}$ et $f''(u) = n(n-1)(1-u)^{n-2} \geqslant 0$: f est convexe donc est au-dessus de ses tangentes, et la droite d'équation $y = 1 - nx$ est sa tangente en 0. Par conséquent, $f(u) \geqslant 1 - nu$ et en appliquant cela à $u = x^2/n^2 \in [\,0\,;1\,]$, il vient

$$\left(1 - \frac{x^2}{n^2}\right)^n \geqslant 1 - n \times \frac{x^2}{n^2} = 1 - \frac{x^2}{n}.$$

4. L'inégalité de gauche découle de la question 2. Montrons à présent celle de droite. Soit $x \in [\,0\,;n\,]$. D'après la question précédente, et en reconnaissant une identité remarquable,

$$\left(1 - \frac{x}{n}\right)^n \times \left(1 + \frac{x}{n}\right)^n \geqslant 1 - \frac{x^2}{n}$$

Or, d'après la question 3, $\left(1 + \dfrac{x}{n}\right)^n \leqslant e^x$ donc (on multiplie par un terme positif),

$$\left(1 - \frac{x}{n}\right)^n \times e^x \geqslant \left(1 - \frac{x}{n}\right)^n \times \left(1 + \frac{x}{n}\right)^n \geqslant 1 - \frac{x^2}{n}$$

En particulier, $\left(1 - \dfrac{x}{n}\right)^n \times e^x \geqslant 1 - \dfrac{x^2}{n}$ donc, en multipliant par $e^{-x} > 0$:

$$\left(1 - \frac{x}{n}\right)^n \geqslant e^{-x} - \frac{x^2}{n} e^{-x}$$

ce qui permet de conclure.

Remarque : Attention à ne pas sortir l'arme de la convexité quand cela n'a rien à voir avec la question. Par exemple, si on demande de montrer que $x - \dfrac{x^2}{2} \leqslant \ln(1+x)$ pour tout $x \geqslant 0$, alors on peut faire une étude de fonction (cf. exercice 40 du chapitre 2), ou appliquer une formule de Taylor (cf. chapitre 23), mais il est hors de question d'utiliser la convexité et de parler de corde ou de tangente ici : en effet, $y = x - \dfrac{x^2}{2}$ n'est pas une équation de droite !