Chapitre 9

Dérivabilité et convexité

9.1 Introduction historique

[...] les principes du calcul différentiel, dégagés de toute considération d'infiniment petits ou d'évanouissans, de limites ou de fluxions [...]

Joseph-Louis Lagrange - *Théorie des fonctions analytiques*, Imprimerie de la République, 1797

9.1.1 Des tangentes avant les dérivées

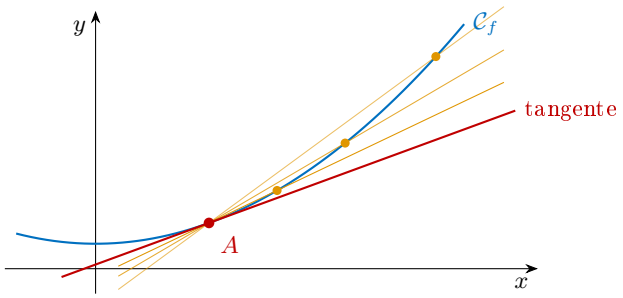
Vers 1636, **Pierre de Fermat** met au point une méthode pour trouver les maximums et les minimums d'une quantité : il compare sa valeur en deux points voisins, divise la différence par l'écart, puis fait disparaître cet écart. Le procédé donne aussi les tangentes. En 1684, **Gottfried Wilhelm Leibniz** publie dans les *Acta Eruditorum* sa *Nova methodus pro maximis et minimis* : ce premier article de calcul différentiel, déjà évoqué au chapitre 6, donne les règles de dérivation d'un produit et d'un quotient, et la condition d'extremum.

9.1.2 Un siècle de doutes

Le calcul fonctionne, mais ses fondements inquiètent. **Michel Rolle**, l'un de ses critiques à l'Académie des sciences, établit en 1691, pour les seuls polynômes, que deux racines consécutives encadrent une racine du polynôme dérivé : c'est l'ancêtre du théorème qui porte aujourd'hui son nom. En 1797, **Joseph-Louis Lagrange** veut fonder la dérivation sur les développements en séries, sans limites, comme l'annonce le titre cité en épigraphe ; il y obtient la future égalité des accroissements finis.

9.1.3 Le choix de Cauchy

C'est pourtant la voie que Lagrange voulait éviter qui l'emporte. En 1823, dans son *Résumé des leçons sur le calcul infinitésimal*, **Augustin-Louis Cauchy** définit la dérivée comme la limite du rapport des accroissements, et en déduit ses propriétés : c'est la démarche de ce chapitre, qui s'appuie sur le chapitre 8. On y ajoute la convexité, étudiée par **Johan Jensen** en 1906. Sur la figure, les sécantes passant par le point A, en orange, se rapprochent de la tangente en rouge lorsque le second point se rapproche de A.



9.2 Dérivabilité en un point

Dans tout le chapitre, I désigne un intervalle de $\mathbb{R}$ non vide et non réduit à un point, et les fonctions sont à valeurs réelles, sauf dans la section 9.5. Le chapitre 5 a utilisé la dérivation comme au lycée, en admettant ses règles ; nous les démontrons maintenant à partir des limites du chapitre 8.

9.2.1 Nombre dérivé

Définition : Soit $f : I \to \mathbb{R}$ et $a \in I$. On dit que f est **dérivable en a** si son **taux d'accroissement** en a,

$$\tau_a : x \mapsto \frac{f(x)-f(a)}{x-a}, \qquad \text{défini sur } I \setminus \{a\},$$

admet une limite finie en a. Cette limite est le **nombre dérivé** de f en a :

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}.$$

On définit de même la **dérivabilité à gauche** et **à droite** en a, à l'aide des limites latérales du taux d'accroissement, notées $f'_g(a)$ et $f'_d(a)$. Enfin, f est **dérivable sur I** si elle l'est en tout point de I, et $f' : x \mapsto f'(x)$ est sa **fonction dérivée**.

Les deux écritures de la limite se déduisent l'une de l'autre par le changement de variable $x = a + h$ (théorème 3 du chapitre 8). D'après la proposition 5 du chapitre 8, f est dérivable en un point a intérieur à I si et seulement si elle y est dérivable à gauche et à droite, avec $f'_g(a) = f'_d(a)$.

Interprétations. Le taux d'accroissement $\tau_a(x)$ est la pente de la *sécante* à la courbe de f passant par les points d'abscisses a et x. Lorsque f est dérivable en a, ces sécantes se rapprochent de la droite

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

la **tangente** à la courbe au point d'abscisse a, comme le montrait la figure de l'introduction. Si $f(t)$ est la position d'un mobile sur un axe à l'instant t, le taux d'accroissement est sa vitesse moyenne entre les instants a et t, et $f'(a)$ sa *vitesse instantanée* à l'instant a.

Le résultat suivant a déjà été établi au chapitre 8 (proposition 10) ; nous en rappelons la démonstration, qui revient directement au taux d'accroissement.

Proposition 1 : Une fonction dérivable en a est continue en a.

∘ Démonstration : Pour $x \in I$, $x \neq a$, on écrit

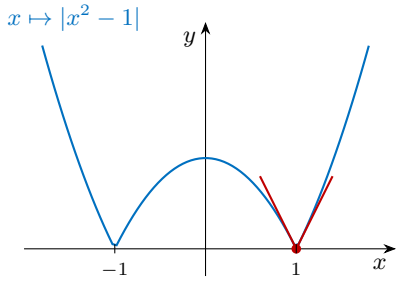
$$f(x) = f(a) + (x-a) \times \frac{f(x)-f(a)}{x-a}.$$

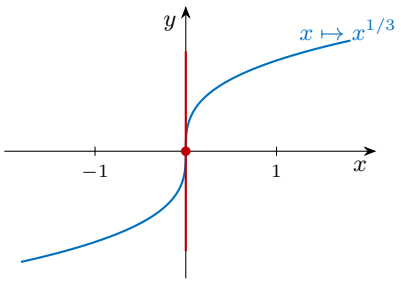
Quand x tend vers a, le produit tend vers $0 \times f'(a) = 0$, par les opérations sur les limites : donc $\lim_{x\to a} f(x) = f(a)$. □

La réciproque est fausse, comme le montrent les exemples suivants : la continuité dit que $f(x) - f(a)$ tend vers 0, la dérivabilité qu'il tend vers 0 « aussi vite que $x - a$ », avec un rapport qui a une limite.

Exemples :

- La valeur absolue est continue en 0, mais son taux d'accroissement $\frac{|h|}{h}$ vaut 1 pour $h > 0$ et -1 pour $h < 0$: elle est dérivable à gauche et à droite en 0, avec $f'_g(0) = -1 \neq 1 = f'_d(0)$, mais pas dérivable. Sa courbe présente en 0 un *point anguleux*, avec deux demi-tangentes.
- La fonction racine cubique, réciproque de $x \mapsto x^3$, est continue en 0, et son taux d'accroissement $\frac{h^{1/3}}{h} = h^{-2/3}$, pour $h > 0$, tend vers $+\infty$; il en est de même pour $h < 0$. Elle n'est pas dérivable en 0, et sa courbe y a une *tangente verticale*.
- La fonction $f : x \mapsto x \sin\frac{1}{x}$, prolongée par $f(0) = 0$, est continue en 0 (chapitre 8), mais son taux d'accroissement $\sin\frac{1}{h}$ n'a pas de limite en 0 : elle n'est pas dérivable en 0, même à droite. En revanche, $g : x \mapsto x^2 \sin\frac{1}{x}$, prolongée par $g(0) = 0$, est dérivable en 0, de nombre dérivé nul, puisque $\left|h \sin\frac{1}{h}\right| \leqslant |h|$.





À gauche, la courbe de $x \mapsto |x^2-1|$ arrive en $(1,0)$ avec deux demi-tangentes de pentes -2 et 2, en rouge : le point est anguleux. À droite, la courbe de la racine cubique traverse sa tangente verticale en l'origine.

Remarque : L'erreur la plus répandue consiste à étudier la dérivabilité en un point en calculant la dérivée *autour* du point et en passant à la limite. Ce procédé, qui sera justifié sous des hypothèses précises par le théorème 8, n'est pas la définition : pour la fonction g ci-dessus, $g'(x) = 2x\sin\frac{1}{x} - \cos\frac{1}{x}$ n'a pas de limite en 0, alors que g est dérivable en 0. Revenir au taux d'accroissement est toujours sûr.

9.2.2 Opérations

Théorème 1 : Opérations sur les fonctions dérivables :
Soit f et g deux fonctions dérivables en $a \in I$, et λ, μ deux réels. Alors $\lambda f + \mu g$ et fg sont dérivables en a, et

$$(\lambda f + \mu g)'(a) = \lambda f'(a) + \mu g'(a) \qquad \text{et} \qquad (fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

Si de plus $g(a) \neq 0$, alors $\frac{f}{g}$ est définie au voisinage de a, dérivable en a, et

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

∘ Démonstration :

▷ *Combinaison linéaire.* Le taux d'accroissement de $\lambda f + \mu g$ est la combinaison correspondante des taux de f et de g, et l'on conclut par les opérations sur les limites (chapitre 8, théorème 2).

▷ *Produit.* Pour $x \neq a$, on fait apparaître un terme intermédiaire :

$$\frac{f(x)g(x) - f(a)g(a)}{x-a} = \frac{f(x)-f(a)}{x-a}\,g(x) + f(a)\,\frac{g(x)-g(a)}{x-a}.$$

Le premier terme tend vers $f'(a)g(a)$, car g est continue en a (proposition 1), et le second vers $f(a)g'(a)$.

▷ *Quotient.* La fonction g est continue en a et $g(a) \neq 0$: elle ne s'annule pas au voisinage de a (chapitre 8, théorème 2). Pour $x \neq a$ dans ce voisinage,

$$\frac{1}{x-a}\left(\frac{1}{g(x)} - \frac{1}{g(a)}\right) = -\frac{g(x)-g(a)}{x-a} \times \frac{1}{g(x)g(a)} \xrightarrow[x\to a]{} -\frac{g'(a)}{g(a)^2},$$

et l'on écrit $\frac{f}{g} = f \times \frac{1}{g}$, pour appliquer la formule du produit. □

Ces formules sont celles de la proposition 3 du chapitre 5, admises jusqu'ici. Comme les fonctions constantes et l'identité sont dérivables, de dérivées 0 et 1, toutes les fonctions polynomiales sont dérivables sur $\mathbb{R}$, et les fractions rationnelles sur leur ensemble de définition.

Théorème 2 : Dérivée d'une composée :
Soit $f : I \to \mathbb{R}$ dérivable en a, à valeurs dans un intervalle J, et $g : J \to \mathbb{R}$ dérivable en $b = f(a)$. Alors $g \circ f$ est dérivable en a, et

$$(g \circ f)'(a) = g'\big(f(a)\big) \times f'(a).$$

∘ Démonstration : On ne peut pas écrire directement $\frac{g(f(x))-g(b)}{x-a} = \frac{g(f(x))-g(b)}{f(x)-b} \times \frac{f(x)-b}{x-a}$, car $f(x)-b$ peut s'annuler pour des $x \neq a$ arbitrairement proches de a. On contourne la difficulté avec la fonction φ définie sur J par

$$\varphi(y) = \frac{g(y)-g(b)}{y-b} \quad \text{si } y \neq b, \qquad \varphi(b) = g'(b).$$

Elle est continue en b, par définition de $g'(b)$, et vérifie $g(y) - g(b) = \varphi(y)(y-b)$ pour tout $y \in J$, y compris $y = b$. Ainsi, pour $x \neq a$,

$$\frac{g(f(x)) - g(f(a))}{x-a} = \varphi\big(f(x)\big) \times \frac{f(x)-f(a)}{x-a}.$$

Par continuité de f en a et de φ en b, $\varphi(f(x))$ tend vers $\varphi(b) = g'(b)$, et le second facteur tend vers $f'(a)$. □

Exemple : La fonction $x \mapsto \sqrt{1+x^2}$ est dérivable sur $\mathbb{R}$, comme composée de $x \mapsto 1+x^2$, à valeurs dans $[1\,,+\infty[$, et de la racine carrée, dérivable sur $]0\,,+\infty[$; sa dérivée est

$$x \mapsto \frac{2x}{2\sqrt{1+x^2}} = \frac{x}{\sqrt{1+x^2}}.$$

En revanche, la composée $x \mapsto \sqrt{x^2} = |x|$ n'est pas dérivable en 0 : le théorème ne s'applique pas, la racine carrée n'étant pas dérivable en $f(0) = 0$.

9.2.3 Dérivée d'une fonction réciproque

Théorème 3 : Dérivée d'une fonction réciproque :
Soit f une fonction continue et strictement monotone sur I, et $f^{-1} : f(I) \to I$ sa réciproque. Si f est dérivable en $a \in I$ et si $f'(a) \neq 0$, alors f^{-1} est dérivable en $b = f(a)$, et

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'\big(f^{-1}(b)\big)}.$$

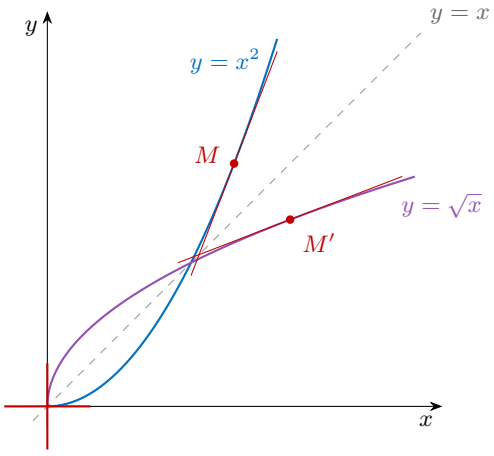
Si $f'(a) = 0$, alors f^{-1} n'est pas dérivable en b, et sa courbe y a une tangente verticale.

∘ Démonstration : D'après le théorème de la bijection continue (chapitre 8, théorème 11), $J = f(I)$ est un intervalle et f^{-1} est continue et strictement monotone sur J. Pour $y \in J$, $y \neq b$, posons $x = f^{-1}(y)$: on a $x \neq a$ par injectivité, et

$$\frac{f^{-1}(y) - f^{-1}(b)}{y-b} = \frac{x-a}{f(x)-f(a)} = \frac{1}{\tau_a(x)}.$$

Quand y tend vers b, $x = f^{-1}(y)$ tend vers a par continuité de f^{-1}, et $\tau_a(x)$ tend vers $f'(a)$ par composition des limites. Si $f'(a) \neq 0$, le quotient tend vers $\frac{1}{f'(a)}$. Si $f'(a) = 0$, le taux $\tau_a(x)$ garde le signe constant de la monotonie de f, et son inverse tend vers $+\infty$ ou $-\infty$. □

Ce théorème achève la démonstration du théorème 2 du chapitre 5, dont le chapitre 8 avait établi la partie continue. Géométriquement, les courbes de f et de f^{-1} sont symétriques par rapport à la première bissectrice, et la symétrie transforme une droite de pente $p \neq 0$ en une droite de pente $\frac{1}{p}$.



La tangente en $M = (1{,}3\,;1{,}69)$ à la parabole, de pente 2,6, a pour symétrique la tangente en M' à la courbe de la racine carrée, de pente $\frac{1}{2{,}6}$. À l'origine, la tangente horizontale de la parabole devient la tangente verticale de la racine carrée, qui n'est pas dérivable en 0.

Exemple : La restriction de tan à $]-\frac{\pi}{2}\,,\frac{\pi}{2}[$ est continue, strictement croissante, de dérivée $1+\tan^2$ qui ne s'annule pas. Pour $y \in \mathbb{R}$, avec $x = \operatorname{Arctan} y$,

$$\operatorname{Arctan}'(y) = \frac{1}{1+\tan^2(\operatorname{Arctan} y)} = \frac{1}{1+y^2},$$

ce qui justifie la formule utilisée au chapitre 5. De même, ln est dérivable de dérivée $\frac{1}{\exp(\ln y)} = \frac{1}{y}$, et Arcsin est dérivable sur $]-1\,,1[$ mais pas en ± 1, où $\sin' = \cos$ s'annule en $\pm\frac{\pi}{2}$.

9.3 Extrema, théorèmes de Rolle et des accroissements finis

9.3.1 Extremum local

Définition : Soit $f : I \to \mathbb{R}$ et $a \in I$. On dit que f admet un **maximum local** en a s'il existe un voisinage V de a tel que $f(x) \leqslant f(a)$ pour tout $x \in V \cap I$, et un **minimum local** si $f(x) \geqslant f(a)$ sur un tel voisinage. Un **extremum local** est un maximum ou un minimum local. On dit que a est un **point critique** de f si f est dérivable en a et $f'(a) = 0$.

Un extremum sur I tout entier, que l'on dit aussi *global*, est en particulier local. La réciproque est fausse : la fonction $x \mapsto x^3 - 3x$ admet en -1 un maximum local, de valeur 2, alors qu'elle n'est pas majorée sur $\mathbb{R}$.

Proposition 2 : *(Condition nécessaire d'extremum.)* Soit $f : I \to \mathbb{R}$ et a un point *intérieur* à I. Si f admet un extremum local en a et si f est dérivable en a, alors $f'(a) = 0$.

∘ Démonstration : Supposons que f admette un maximum local en a : puisque a est intérieur, il existe $\delta > 0$ tel que $[a-\delta\,,a+\delta]$ soit contenu dans I et que $f(x) \leqslant f(a)$ pour $|x-a| \leqslant \delta$. Le taux d'accroissement $\tau_a(x)$ est alors négatif pour $a < x \leqslant a+\delta$ et positif pour $a-\delta \leqslant x < a$. En passant à la limite dans ces inégalités (chapitre 8, théorème 4),

$$f'(a) = f'_d(a) \leqslant 0 \qquad \text{et} \qquad f'(a) = f'_g(a) \geqslant 0,$$

donc $f'(a) = 0$. On raisonne de même pour un minimum, ou l'on applique ce qui précède à $-f$. □

Remarques :

- La condition n'est pas suffisante : $x \mapsto x^3$ a un point critique en 0, sans y avoir d'extremum, puisqu'elle change de signe. Un point critique est seulement un *candidat*.
- L'hypothèse que a est intérieur est indispensable : la fonction $x \mapsto x$ sur $[0\,,1]$ atteint son minimum en 0 et son maximum en 1, où sa dérivée vaut 1. Pour chercher les extrema d'une fonction dérivable sur un segment, on compare donc ses valeurs aux points critiques *et* aux extrémités ; le théorème des bornes atteintes du chapitre 8 garantit que le maximum est l'une de ces valeurs.

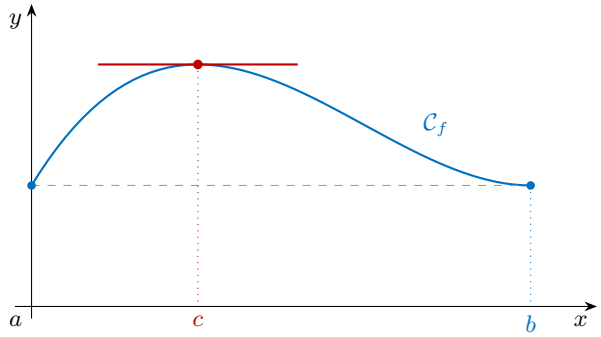
9.3.2 Le théorème de Rolle

Théorème 4 : Théorème de Rolle :
Soit $a < b$ et f une fonction continue sur $[a\,,b]$, dérivable sur $]a\,,b[$, telle que $f(a) = f(b)$. Alors il existe $c \in]a\,,b[$ tel que $f'(c) = 0$.

∘ Démonstration : D'après le théorème des bornes atteintes (chapitre 8, théorème 10), f atteint sur le segment $[a\,,b]$ son minimum m et son maximum M.

▷ *Si* $m = M$, la fonction f est constante sur $[a\,,b]$, et tout $c \in]a\,,b[$ convient.

▷ *Sinon,* l'une des deux valeurs m ou M est différente de $f(a) = f(b)$, disons M. Le maximum est alors atteint en un point c qui n'est ni a ni b, donc intérieur à $[a\,,b]$, et la proposition 2 donne $f'(c) = 0$. □



La courbe part et arrive à la même hauteur, sur la droite orange : entre les deux, elle atteint un sommet c, où la tangente, en rouge, est horizontale.

Remarques :

- Les trois hypothèses sont nécessaires. La fonction $x \mapsto |x|$ sur $[-1\,,1]$ vérifie $f(-1) = f(1)$, mais sa dérivée ne s'annule pas : elle n'est pas dérivable en 0. La fonction qui vaut x sur $[0\,,1[$ et 0 en 1 est dérivable sur $]0\,,1[$ de dérivée 1 : elle n'est pas continue en 1.
- Le théorème est faux pour une fonction à valeurs complexes : $t \mapsto \mathrm{e}^{\mathrm{i}t}$ prend la même valeur en 0 et en 2π, alors que sa dérivée $\mathrm{i}\mathrm{e}^{\mathrm{i}t}$ est de module 1 (section 9.5).

9.3.3 Les accroissements finis

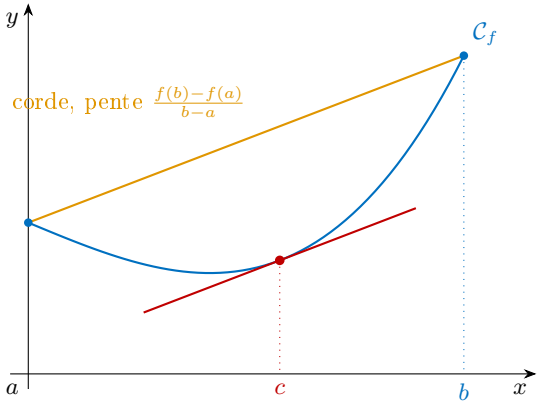
Théorème 5 : Égalité des accroissements finis :
Soit $a < b$ et f une fonction continue sur $[a\,,b]$ et dérivable sur $]a\,,b[$. Alors il existe $c \in]a\,,b[$ tel que

$$f(b) - f(a) = f'(c)\,(b-a).$$

∘ Démonstration : On retranche à f la fonction affine dont la courbe est la corde joignant les points d'abscisses a et b. Posons

$$g(x) = f(x) - f(a) - p\,(x-a), \qquad \text{avec} \qquad p = \frac{f(b)-f(a)}{b-a}.$$

La fonction g est continue sur $[a\,,b]$, dérivable sur $]a\,,b[$, et $g(a) = g(b) = 0$. Le théorème de Rolle fournit $c \in]a\,,b[$ tel que $g'(c) = f'(c) - p = 0$. □



La tangente en rouge au point d'abscisse $c \approx 1{,}386$ est parallèle à la corde orange : c'est la figure de Rolle, inclinée. Cinématiquement, un mobile qui a parcouru une distance $f(b) - f(a)$ pendant la durée $b - a$ a eu, à un certain instant c, une vitesse instantanée égale à sa vitesse moyenne.

Remarque : Le point c n'est en général pas unique, et l'on ne sait pas où il se trouve dans $]a\,,b[$. La force du théorème vient de ce qu'il transforme une information locale, la valeur de f' en chaque point, en une information globale sur l'accroissement $f(b)-f(a)$. Il ne s'utilise presque jamais pour calculer c, mais pour *majorer* ou *minorer* un accroissement à l'aide d'une borne sur f'.

Théorème 6 : Inégalité des accroissements finis :
Soit f une fonction dérivable sur un intervalle I.

(i) Si $m \leqslant f' \leqslant M$ sur I, alors $m(y-x) \leqslant f(y)-f(x) \leqslant M(y-x)$ pour tous $x \leqslant y$ dans I.

(ii) Si $|f'| \leqslant K$ sur I, alors f est K-lipschitzienne sur I : $|f(y)-f(x)| \leqslant K|y-x|$.

∘ Démonstration : Pour $x<y$ dans I, f est continue sur $[x\,,y]$ et dérivable sur $]x\,,y[$, puisque $[x\,,y] \subset I$. Le théorème 5 donne $c \in]x\,,y[$ tel que $f(y)-f(x)=f'(c)(y-x)$, et il suffit d'encadrer $f'(c)$; le cas $x=y$ est trivial. Le point (ii) est le point (i) avec $m=-K$ et $M=K$. □

C'est le critère annoncé au chapitre 8 : une dérivée bornée donne une fonction lipschitzienne, donc uniformément continue. Ainsi sin est 1-lipschitzienne, ce que l'on retrouve sans formule de trigonométrie, et Arctan l'est aussi, puisque $\left|\frac{1}{1+x^2}\right| \leqslant 1$.

Exemple : Pour tous réels $0<x<y$, l'inégalité des accroissements finis appliquée à ln sur $[x\,,y]$, où $\frac{1}{y} \leqslant \ln' \leqslant \frac{1}{x}$, donne

$$\frac{y-x}{y} \leqslant \ln y - \ln x \leqslant \frac{y-x}{x}.$$

Avec $x=n$ et $y=n+1$, on retrouve l'encadrement $\frac{1}{n+1} \leqslant \ln(n+1)-\ln n \leqslant \frac{1}{n}$, qui compare les sommes $1+\frac{1}{2}+\cdots+\frac{1}{n}$ au logarithme, comme le chapitre 20 le précisera. L'inégalité est obtenue sans connaître la position du point c.

9.3.4 Dérivée et monotonie

Théorème 7 : Dérivée et monotonie :
Soit f une fonction dérivable sur un *intervalle* I.

(i) f est constante sur I si et seulement si $f'=0$ sur I.

(ii) f est croissante sur I si et seulement si $f' \geqslant 0$ sur I.

(iii) f est strictement croissante sur I si et seulement si $f' \geqslant 0$ sur I et f' n'est identiquement nulle sur aucun intervalle $[x\,,y]$ avec $x<y$.

On a les énoncés analogues pour la décroissance, en remplaçant $f' \geqslant 0$ par $f' \leqslant 0$.

∘ Démonstration :

▷ *(ii)* Si f est croissante, ses taux d'accroissement sont positifs, et leur limite $f'(x)$ aussi. Réciproquement, si $f' \geqslant 0$, l'inégalité des accroissements finis avec $m=0$ donne $f(y)-f(x) \geqslant 0$ pour $x \leqslant y$.

▷ *(i)* On applique (ii) à f et à $-f$: f est constante si et seulement si elle est à la fois croissante et décroissante.

▷ *(iii)* Si f est strictement croissante, $f' \geqslant 0$ par (ii), et si f' était nulle sur un intervalle $[x\,,y]$ avec $x<y$, f y serait constante par (i), ce qui est exclu. Réciproquement, supposons $f' \geqslant 0$ et f' nulle sur aucun tel intervalle : f est croissante par (ii). Si l'on avait $f(x)=f(y)$ avec $x<y$, la croissance donnerait $f(x) \leqslant f(t) \leqslant f(y)=f(x)$ pour tout $t \in [x\,,y]$: f serait constante sur $[x\,,y]$, et f' y serait nulle, ce qui est exclu.

▷ *Décroissance.* La fonction f est décroissante, ou strictement décroissante, si et seulement si $-f$ est croissante, ou strictement croissante. Il suffit donc d'appliquer (ii) et (iii) à $-f$, dont la dérivée est $-f'$. □

Ce théorème démontre le théorème 1 du chapitre 5, dont le point (iii) est un cas particulier : une dérivée positive qui ne s'annule qu'en un nombre fini de points n'est nulle sur aucun intervalle non trivial. L'hypothèse d'intervalle est indispensable, comme l'avait montré la fonction $x \mapsto \frac{1}{x}$ sur $\mathbb{R}^*$.

9.3.5 Application aux suites récurrentes

Le chapitre 7 a étudié les suites définies par $u_{n+1} = f(u_n)$ à l'aide de la monotonie. L'inégalité des accroissements finis donne une méthode plus puissante, qui fournit en outre une vitesse de convergence.

Proposition 3 : *(Point fixe attractif.)* Soit f une fonction dérivable sur un segment I, telle que $f(I) \subset I$, et $k \in [0\,,1[$ tel que $|f'| \leqslant k$ sur I. Alors f admet un unique point fixe $\ell \in I$, et toute suite définie par $u_0 \in I$ et $u_{n+1} = f(u_n)$ vérifie

$$|u_n - \ell| \leqslant k^n\,|u_0 - \ell|, \qquad \text{donc} \qquad \lim_{n\to+\infty} u_n = \ell.$$

∘ Démonstration :

▷ *Existence et unicité du point fixe.* La fonction f est continue sur le segment I et à valeurs dans I : elle admet un point fixe ℓ (chapitre 8, proposition 14). Si ℓ' en est un autre, l'inégalité des accroissements finis donne $|\ell - \ell'| = |f(\ell) - f(\ell')| \leqslant k|\ell - \ell'|$, d'où $|\ell - \ell'| = 0$, puisque $k < 1$.

▷ *Convergence.* La suite est bien définie et reste dans I, qui est stable par f. Pour tout n, l'inégalité des accroissements finis donne

$$|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leqslant k|u_n - \ell|,$$

et une récurrence immédiate conclut, puisque $\lim_{n\to+\infty} k^n = 0$. □

Exemple : Soit $u_0 \in [0\,,1]$ et $u_{n+1} = \cos u_n$. Le segment $[0\,,1]$ est stable par cos, puisque $0 < \cos 1 \leqslant \cos x \leqslant 1$, et $|\cos'| = |\sin| \leqslant \sin 1 \approx 0{,}841$ sur $[0\,,1]$. La suite converge donc vers l'unique solution $\ell \approx 0{,}739085$ de l'équation $\cos \ell = \ell$, avec $|u_n - \ell| \leqslant (\sin 1)^n$. Cette majoration garantit six décimales exactes dès que $(\sin 1)^n \leqslant 10^{-6}$, soit $n \geqslant 81$. En pratique, la convergence est plus rapide, car la constante pertinente est $|\sin \ell| \approx 0{,}674$: la majoration est une garantie, pas une estimation fine.

9.3.6 Le théorème de la limite de la dérivée

Théorème 8 : Limite de la dérivée :
Soit f une fonction continue sur I, dérivable sur $I \setminus \{a\}$, et telle que f' admette une limite $\ell \in \overline{\mathbb{R}}$ en a. Alors

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = \ell.$$

En particulier, si ℓ est réel, f est dérivable en a, $f'(a) = \ell$, et f' est continue en a. Si $\ell = \pm\infty$, f n'est pas dérivable en a, et sa courbe y a une tangente verticale.

∘ Démonstration : Soit $x \in I$, $x \neq a$. La fonction f est continue sur le segment d'extrémités a et x, et dérivable à l'intérieur : le théorème 5 fournit c_x strictement compris entre a et x tel que

$$\frac{f(x) - f(a)}{x - a} = f'(c_x).$$

Quand x tend vers a, on a $|c_x - a| < |x - a|$, donc c_x tend vers a, et $f'(c_x)$ tend vers ℓ. Pour le justifier proprement, soit V un voisinage de ℓ : il existe $\delta > 0$ tel que $f'(t) \in V$ pour $0 < |t - a| \leqslant \delta$, et alors $\tau_a(x) = f'(c_x) \in V$ pour $0 < |x - a| \leqslant \delta$. □

Exemple : La fonction $f : x \mapsto \operatorname{Arcsin}\left(\sqrt{x}\right)$ sur $[0\,,1]$ est continue, et dérivable sur $]0\,,1[$, avec

$$f'(x) = \frac{1}{2\sqrt{x}}\,\frac{1}{\sqrt{1-x}}.$$

Cette dérivée tend vers $+\infty$ en 0 et en 1 : la courbe de f a des tangentes verticales aux deux extrémités.

Remarque : La réciproque est fausse, et c'est le piège le plus courant : une fonction peut être dérivable en a sans que f' y ait de limite. C'est le cas de $g : x \mapsto x^2 \sin \frac{1}{x}$, prolongée par $g(0) = 0$, vue à la section 9.2. Le théorème 8 donne une condition suffisante de dérivabilité, commode en pratique, mais quand f' n'a pas de limite on ne peut rien conclure, et il faut revenir au taux d'accroissement.

9.4 Fonctions de classe $\mathcal{C}^k$

9.4.1 Définition

Définition : Soit $f : I \to \mathbb{R}$. On pose $f^{(0)} = f$ et, si $f^{(n)}$ est dérivable sur I, $f^{(n+1)} = \left(f^{(n)}\right)'$. Pour $k \in \mathbb{N}$, on dit que f est **de classe $\mathcal{C}^k$** sur I si f est k fois dérivable sur I et si $f^{(k)}$ est continue ; on note $\mathcal{C}^k(I, \mathbb{R})$ leur ensemble. La fonction f est **de classe $\mathcal{C}^\infty$** si elle est de classe $\mathcal{C}^k$ pour tout $k \in \mathbb{N}$.

Ces définitions sont celles du chapitre 5. La classe $\mathcal{C}^0$ est celle des fonctions continues du chapitre 8, et d'après la proposition 1, une fonction k fois dérivable est de classe $\mathcal{C}^{k-1}$, toutes ses dérivées d'ordre inférieur à k étant dérivables, donc continues. On a ainsi les inclusions

$$\mathcal{C}^\infty(I, \mathbb{R}) \subset \cdots \subset \mathcal{C}^{k+1}(I, \mathbb{R}) \subset \mathcal{C}^k(I, \mathbb{R}) \subset \cdots \subset \mathcal{C}^0(I, \mathbb{R}).$$

Exemples :

- Les fonctions polynomiales, exp, sin, cos sont de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, ainsi que ln sur $]0, +\infty[$, puisque chacune de leurs dérivées est encore dérivable.
- Les inclusions sont strictes. La fonction $x \mapsto x|x|$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}$, de dérivée $x \mapsto 2|x|$, mais pas deux fois dérivable (chapitre 5). La fonction $g : x \mapsto x^2 \sin\frac{1}{x}$, prolongée par $g(0) = 0$, est dérivable sur $\mathbb{R}$, mais sa dérivée $g'(x) = 2x \sin\frac{1}{x} - \cos\frac{1}{x}$ pour $x \neq 0$, avec $g'(0) = 0$, n'est pas continue en 0 : g est dérivable sans être de classe $\mathcal{C}^1$.

9.4.2 Formule de Leibniz

Théorème 9 : Formule de Leibniz :
Soit $n \in \mathbb{N}$ et f, g deux fonctions n fois dérivables sur I. Alors fg est n fois dérivable sur I, et

$$(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}\, g^{(n-k)}.$$

Si f et g sont de classe $\mathcal{C}^n$, il en est de même de fg.

∘ Démonstration : On raisonne par récurrence sur n. Pour $n = 0$, la formule se réduit à $fg = fg$. Supposons-la vraie au rang n, et soit f, g deux fonctions $n+1$ fois dérivables. Chaque terme $f^{(k)}g^{(n-k)}$ de la somme est un produit de deux fonctions dérivables, donc dérivable par le théorème 1, de dérivée $f^{(k+1)}g^{(n-k)} + f^{(k)}g^{(n+1-k)}$. Ainsi $(fg)^{(n)}$ est dérivable, et

$$(fg)^{(n+1)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k+1)} g^{(n-k)} + \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n+1-k)}.$$

Dans la première somme, on pose $j = k + 1$; en regroupant les termes de même indice, la formule de Pascal $\binom{n}{j-1} + \binom{n}{j} = \binom{n+1}{j}$ du chapitre 3 donne

$$(fg)^{(n+1)} = \sum_{j=0}^{n+1} \binom{n+1}{j} f^{(j)} g^{(n+1-j)}.$$

Si f et g sont de classe $\mathcal{C}^n$, chaque terme de la formule est continu, comme produit de fonctions continues, et $(fg)^{(n)}$ l'est aussi. □

La démonstration est exactement celle de la formule du binôme, au chapitre 3 : la dérivation d'un produit joue le rôle de la distributivité, et la formule de Pascal fait le reste.

Exemple : Calculons les dérivées successives de $x \mapsto x^2 \mathrm{e}^x$. Avec $f(x) = x^2$, dont les dérivées sont nulles à partir de l'ordre 3, et $g = \exp$, seuls trois termes subsistent :

$$\left(x^2 \mathrm{e}^x\right)^{(n)} = x^2 \mathrm{e}^x + n \times 2x\, \mathrm{e}^x + \frac{n(n-1)}{2} \times 2\, \mathrm{e}^x = \left(x^2 + 2nx + n(n-1)\right) \mathrm{e}^x.$$

La formule est d'autant plus efficace que l'un des facteurs a des dérivées qui s'annulent : on choisit pour f le polynôme.

9.4.3 Opérations sur les fonctions de classe $\mathcal{C}^k$

Proposition 4 : Soit $k \in \mathbb{N} \cup \{\infty\}$, et f, g deux fonctions de classe $\mathcal{C}^k$ sur I.

(i) Les fonctions $\lambda f + \mu g$ et fg sont de classe $\mathcal{C}^k$, et $\frac{f}{g}$ aussi si g ne s'annule pas sur I.

(ii) Si f est à valeurs dans un intervalle J, et si $h : J \to \mathbb{R}$ est de classe $\mathcal{C}^k$, alors $h \circ f$ est de classe $\mathcal{C}^k$.

(iii) Si $k \geqslant 1$ et si f' ne s'annule pas sur I, alors f réalise une bijection de I sur l'intervalle $f(I)$, et f^{-1} est de classe $\mathcal{C}^k$ sur $f(I)$.

∘ Démonstration : Pour $k = 0$, ce sont les résultats du chapitre 8. On raisonne par récurrence sur $k \geqslant 1$, le cas $k = \infty$ s'en déduisant. Pour la combinaison linéaire, $(\lambda f + \mu g)^{(k)} = \lambda f^{(k)} + \mu g^{(k)}$ est continue, et le produit relève du théorème 9.

▷ *Quotient.* Il suffit de traiter $\frac{1}{g}$, puis d'utiliser le produit. Si g est de classe $\mathcal{C}^k$, alors $\frac{1}{g}$ est dérivable, de dérivée $-\frac{g'}{g^2}$; or g' est de classe $\mathcal{C}^{k-1}$, ainsi que $\frac{1}{g^2} = \frac{1}{g} \times \frac{1}{g}$ par hypothèse de récurrence, de sorte que $\left(\frac{1}{g}\right)'$ est de classe $\mathcal{C}^{k-1}$, et $\frac{1}{g}$ de classe $\mathcal{C}^k$.

▷ *Composée.* La fonction $h \circ f$ est dérivable, de dérivée $(h' \circ f) \times f'$. Par hypothèse de récurrence appliquée à h' et à f, qui sont de classe $\mathcal{C}^{k-1}$, la fonction $h' \circ f$ est de classe $\mathcal{C}^{k-1}$, et le produit par f' aussi.

▷ *Réciproque.* La dérivée f' est continue et ne s'annule pas sur l'intervalle I : elle garde un signe constant (chapitre 8, proposition 15), et f est strictement monotone par le théorème 7. Par le théorème 3, f^{-1} est dérivable sur $f(I)$, et $(f^{-1})' = \frac{1}{f' \circ f^{-1}}$. Si f^{-1} est de classe $\mathcal{C}^j$ avec $j \leqslant k-1$, alors $f' \circ f^{-1}$ l'est aussi, par le point (ii), puis son inverse, et f^{-1} est de classe $\mathcal{C}^{j+1}$. En partant de $j = 0$, on atteint $j = k$. □

Les démonstrations relatives à la composée et à la réciproque ne sont pas exigibles ; elles montrent néanmoins le schéma commun, qui consiste à exprimer la dérivée à l'aide de fonctions de classe $\mathcal{C}^{k-1}$. Par exemple, Arctan est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, comme réciproque de la restriction de tan, de dérivée $1 + \tan^2$ qui ne s'annule pas.

9.5 Fonctions à valeurs complexes

Le chapitre 5 a défini la dérivabilité d'une fonction $f : I \to \mathbb{C}$ par celle de ses parties réelle et imaginaire, en posant $f' = (\mathrm{Re}\, f)' + \mathrm{i}\,(\mathrm{Im}\, f)'$. Cette définition coïncide avec celle par le taux d'accroissement.

Proposition 5 : Soit $f : I \to \mathbb{C}$ et $a \in I$. La fonction f est dérivable en a si et seulement si son taux d'accroissement $\frac{f(x)-f(a)}{x-a}$ admet une limite finie dans $\mathbb{C}$ quand x tend vers a, et cette limite est alors $f'(a)$. Les théorèmes 1, 2 et 9 restent valables, ainsi que l'équivalence de la constance de f avec l'annulation de f' sur un intervalle.

∘ Démonstration : Les parties réelle et imaginaire du taux d'accroissement de f sont les taux d'accroissement de $\mathrm{Re}\, f$ et de $\mathrm{Im}\, f$. D'après la proposition 20 du chapitre 8, le taux de f a une limite si et seulement si ceux de $\mathrm{Re}\, f$ et $\mathrm{Im}\, f$ en ont une, et la limite est alors $(\mathrm{Re}\, f)'(a) + \mathrm{i}\,(\mathrm{Im}\, f)'(a)$. Les démonstrations des théorèmes 1 et 9 n'utilisent que les opérations sur les limites, valables dans $\mathbb{C}$. La caractérisation des fonctions constantes s'obtient en appliquant le théorème 7 aux parties réelle et imaginaire. □

En revanche, tous les résultats qui utilisent l'ordre de $\mathbb{R}$ disparaissent : il n'y a ni extremum d'une fonction complexe, ni théorème de Rolle, ni égalité des accroissements finis. La fonction $t \mapsto \mathrm{e}^{\mathrm{i}t}$ vérifie $\mathrm{e}^{\mathrm{i}\cdot 2\pi} - \mathrm{e}^{\mathrm{i}\cdot 0} = 0$, alors que sa dérivée ne s'annule jamais ; il n'existe donc aucun c tel que $\mathrm{e}^{2\mathrm{i}\pi} - 1 = 2\pi\,\mathrm{i}\,\mathrm{e}^{\mathrm{i}c}$. L'inégalité, elle, subsiste.

Proposition 6 : *(Inégalité des accroissements finis.)* Soit $f : I \to \mathbb{C}$ dérivable, et $K \geqslant 0$ tel que $|f'| \leqslant K$ sur I. Alors $|f(y) - f(x)| \leqslant K|y - x|$ pour tous x, $y \in I$.

∘ Démonstration : Soit x, $y \in I$ tels que $f(y) \neq f(x)$, le cas d'égalité étant évident. On écrit l'accroissement sous forme trigonométrique, $f(y) - f(x) = r\mathrm{e}^{\mathrm{i}\theta}$ avec $r > 0$, et l'on pose $g = \mathrm{Re}\left(\mathrm{e}^{-\mathrm{i}\theta} f\right)$. La fonction réelle g est dérivable sur I, de dérivée $g' = \mathrm{Re}\left(\mathrm{e}^{-\mathrm{i}\theta} f'\right)$, et donc $|g'| \leqslant |\mathrm{e}^{-\mathrm{i}\theta} f'| = |f'| \leqslant K$. Par le théorème 6 appliqué à g,

$$|f(y) - f(x)| = r = \mathrm{Re}\left(\mathrm{e}^{-\mathrm{i}\theta}(f(y) - f(x))\right) = g(y) - g(x) \leqslant K|y - x|.$$

□

L'astuce consiste à faire tourner le plan pour amener l'accroissement $f(y)-f(x)$ sur l'axe réel, où l'on peut appliquer le résultat réel. Le chapitre 18 en donnera une autre démonstration, pour les fonctions de classe $\mathcal{C}^1$, en écrivant $f(y)-f(x)$ comme l'intégrale de f' et en majorant cette intégrale.

9.6 Fonctions convexes

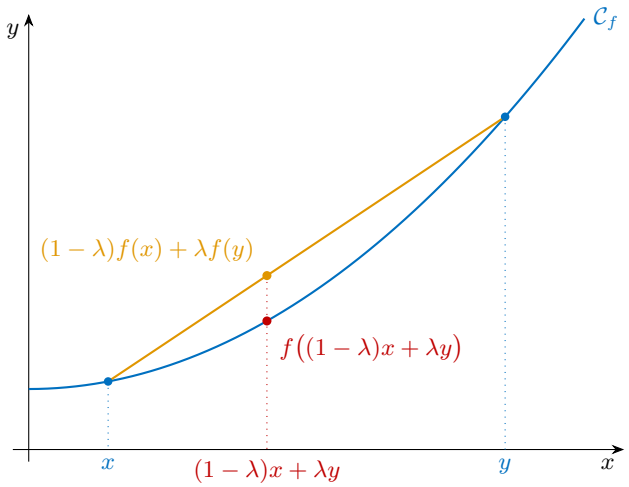
9.6.1 Définition et inégalité de Jensen

Définition : Une fonction $f : I \to \mathbb{R}$ est **convexe** sur I si

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

Elle est **concave** si $-f$ est convexe, c'est-à-dire si l'inégalité inverse est vérifiée.

Lorsque λ parcourt $[0\,,1]$, le réel $(1-\lambda)x+\lambda y$ parcourt le segment d'extrémités x et y, et le point de coordonnées $\big((1-\lambda)x+\lambda y, (1-\lambda)f(x)+\lambda f(y)\big)$ parcourt la *corde* joignant les points $(x,f(x))$ et $(y,f(y))$ de la courbe. La convexité signifie donc que, sur tout segment, *la courbe est au-dessous de ses cordes.*



Au point d'abscisse $(1-\lambda)x+\lambda y$, ici avec $\lambda = 0{,}4$, la valeur de f, en rouge, est inférieure à l'ordonnée du point de la corde, en orange : l'écart entre les deux est l'écart de convexité.

Exemples :

- Les fonctions affines sont à la fois convexes et concaves, avec égalité dans la définition ; ce sont les seules, comme on peut le déduire de la proposition 9 ci-dessous.
- La fonction carré est convexe sur $\mathbb{R}$: un calcul direct donne

$$(1-\lambda)x^2+\lambda y^2-\big((1-\lambda)x+\lambda y\big)^2 = \lambda(1-\lambda)(x-y)^2 \geqslant 0.$$

- La valeur absolue est convexe sur $\mathbb{R}$, par l'inégalité triangulaire et la positivité de $1-\lambda$ et λ.

Proposition 7 : *(Inégalité de Jensen.)* Soit f convexe sur I, $n \geqslant 1$, $x_1,\dots,x_n$ des éléments de I et $\lambda_1,\dots,\lambda_n$ des réels positifs de somme 1. Alors $\sum_{i=1}^{n}\lambda_i x_i \in I$, et

$$f\Big(\sum_{i=1}^{n}\lambda_i x_i\Big) \leqslant \sum_{i=1}^{n}\lambda_i f(x_i).$$

∘ Démonstration : On raisonne par récurrence sur n. Pour $n=1$, on a $\lambda_1=1$ et il y a égalité. Supposons le résultat vrai au rang n, et soit $\lambda_1,\dots,\lambda_{n+1}$ positifs de somme 1. Si $\lambda_{n+1}=1$, tous les autres sont nuls, et il y a égalité. Sinon, $s=\lambda_1+\cdots+\lambda_n = 1-\lambda_{n+1} > 0$, et l'on pose

$$z=\sum_{i=1}^{n}\frac{\lambda_i}{s}x_i,\qquad \text{de sorte que}\qquad \sum_{i=1}^{n+1}\lambda_i x_i = s\,z+\lambda_{n+1}x_{n+1}.$$

Les coefficients $\frac{\lambda_i}{s}$ sont positifs de somme 1 : par hypothèse de récurrence, $z \in I$ et $f(z) \leqslant \sum_{i=1}^{n} \frac{\lambda_i}{s} f(x_i)$. Comme $s + \lambda_{n+1} = 1$, le point $s\,z + \lambda_{n+1} x_{n+1}$ est entre z et x_{n+1}, donc dans l'intervalle I, et la définition de la convexité donne

$$f\Big(\sum_{i=1}^{n+1} \lambda_i x_i\Big) \leqslant s\,f(z) + \lambda_{n+1} f(x_{n+1}) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) + \lambda_{n+1} f(x_{n+1}).$$

□

Pour une fonction concave, l'inégalité s'inverse. Avec $\lambda_i = \frac{1}{n}$, l'inégalité de Jensen compare l'image de la moyenne et la moyenne des images : $f\big(\frac{x_1+\cdots+x_n}{n}\big) \leqslant \frac{f(x_1)+\cdots+f(x_n)}{n}$.

9.6.2 Croissance des pentes

Proposition 8 : *(Lemme des trois pentes.)* Une fonction $f : I \to \mathbb{R}$ est convexe si et seulement si, pour tous $x < y < z$ dans I,

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

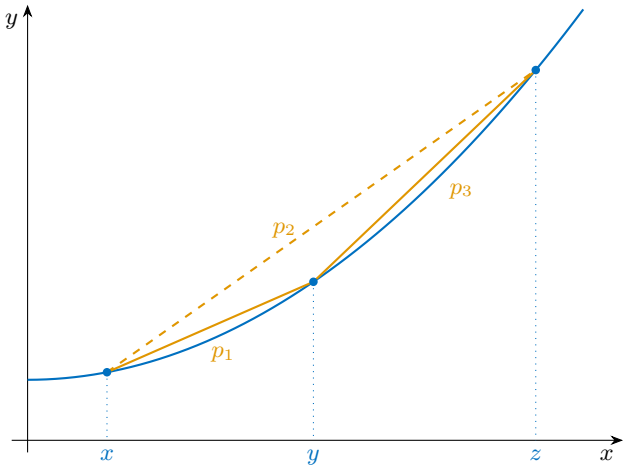
Il est équivalent de demander que, pour tout $a \in I$, le taux d'accroissement τ_a soit croissant sur $I \setminus \{a\}$, ou encore que l'inégalité entre les deux pentes extrêmes soit vérifiée.

∘ Démonstration : Soit $x < y < z$ dans I. On écrit $y = (1-\lambda)x + \lambda z$, avec $\lambda = \frac{y-x}{z-x} \in]0\,,1[$, et l'on note p_1, p_2, p_3 les trois pentes, dans l'ordre de l'énoncé. L'inégalité de convexité en y s'écrit $f(y) \leqslant (1-\lambda)f(x) + \lambda f(z)$, soit, en retranchant $f(x)$,

$$f(y) - f(x) \leqslant \lambda\big(f(z) - f(x)\big), \qquad \text{c'est-à-dire} \qquad p_1 \leqslant p_2,$$

après division par $y - x = \lambda(z - x) > 0$. En retranchant $f(z)$ au lieu de $f(x)$, on obtient de même qu'elle équivaut à $p_2 \leqslant p_3$. Enfin, p_2 est une moyenne pondérée de p_1 et p_3, à poids positifs, puisque $f(z) - f(x) = \big(f(y) - f(x)\big) + \big(f(z) - f(y)\big)$; il est donc compris entre eux, et $p_1 \leqslant p_3$ équivaut aux deux autres inégalités.

Comme tout point d'un segment $[x\,,z]$ s'écrit $(1-\lambda)x + \lambda z$, la convexité équivaut à l'inégalité de convexité en chaque point y strictement compris entre deux autres, donc à la condition de l'énoncé. La croissance des taux τ_a en est une reformulation : elle donne $p_1 \leqslant p_2$ avec $a = x$, et $p_2 \leqslant p_3$ avec $a = z$; réciproquement, elle se lit sur les trois positions possibles de a par rapport aux deux autres points. □



Les trois cordes joignant les points d'abscisses $x < y < z$ ont des pentes croissantes : $p_1 = 0{,}575$, $p_2 = 0{,}925$ pour la corde en pointillés, et $p_3 = 1{,}25$. Plus on se déplace vers la droite, plus la courbe monte vite.

Proposition 9 : *(Position par rapport aux sécantes.)* Soit f convexe sur I, et $x < y$ dans I. La courbe de f est au-dessous de la droite (xy), sécante passant par les points d'abscisses x et y, sur $[x\,,y]$, et au-dessus de cette droite en dehors de $[x\,,y]$.

∘ Démonstration : La première affirmation est la définition. Soit $t \in I$, $t > y$. Le lemme des trois pentes appliqué à $x < y < t$ donne $\frac{f(t)-f(x)}{t-x} \geqslant \frac{f(y)-f(x)}{y-x}$, soit, en multipliant par $t - x > 0$,

$$f(t) \geqslant f(x) + \frac{f(y) - f(x)}{y - x}(t - x) :$$

le point de la courbe d'abscisse t est au-dessus de la sécante. Le cas $t < x$ se traite avec $t < x < y$. □

9.6.3 Fonctions convexes dérivables

Théorème 10 : Convexité et dérivée :
Soit f une fonction dérivable sur I. Alors f est convexe sur I si et seulement si f' est croissante sur I. Dans ce cas, la courbe de f est au-dessus de chacune de ses tangentes :

$$\forall (a, x) \in I^2, \qquad f(x) \geqslant f(a) + f'(a)(x - a).$$

∘ Démonstration :

▷ *Sens direct.* Soit $x < y$ dans I, et t tel que $x < t < y$. Le lemme des trois pentes donne

$$\tau_x(t) \leqslant \frac{f(y) - f(x)}{y - x} \leqslant \tau_y(t).$$

En faisant tendre t vers x dans la première inégalité, puis vers y dans la seconde, on obtient

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

et f' est croissante.

▷ *Réciproque.* Soit $x < y < z$ dans I. Le théorème des accroissements finis fournit $c_1 \in]x, y[$ et $c_2 \in]y, z[$ tels que

$$\frac{f(y) - f(x)}{y - x} = f'(c_1) \leqslant f'(c_2) = \frac{f(z) - f(y)}{z - y},$$

puisque $c_1 < c_2$ et que f' est croissante. C'est l'inégalité entre les pentes extrêmes, qui suffit d'après la proposition 8.

▷ *Tangentes.* Soit $a \in I$ et $x > a$. Par le théorème des accroissements finis, $f(x) - f(a) = f'(c)(x - a)$ avec $c \in]a, x[$, et $f'(c) \geqslant f'(a)$, d'où l'inégalité. Pour $x < a$, on a $f'(c) \leqslant f'(a)$ avec $c \in]x, a[$, et $x - a < 0$, ce qui donne encore $f'(c)(x - a) \geqslant f'(a)(x - a)$. □

Théorème 11 : Convexité et dérivée seconde :
Soit f une fonction deux fois dérivable sur I. Alors f est convexe sur I si et seulement si $f'' \geqslant 0$ sur I.

En effet, f' est croissante si et seulement si $f'' \geqslant 0$, par le théorème 7, et l'on conclut par le théorème 10. C'est le critère le plus utilisé en pratique : exp est convexe sur $\mathbb{R}$, ln concave sur $]0, +\infty[$, $x \mapsto x^\alpha$ convexe sur $]0, +\infty[$ pour $\alpha \geqslant 1$ ou $\alpha \leqslant 0$ et concave pour $0 \leqslant \alpha \leqslant 1$, puisque sa dérivée seconde vaut $\alpha(\alpha - 1)x^{\alpha-2}$.

♡ **Méthode : Pour établir une inégalité entre fonctions**, on peut reconnaître la position d'une courbe par rapport à une tangente ou à une corde : une fonction convexe est au-dessus de ses tangentes et au-dessous de ses cordes, et c'est l'inverse pour une fonction concave. Il suffit alors de calculer une dérivée seconde, au lieu d'étudier les variations d'une différence.

Exemples :

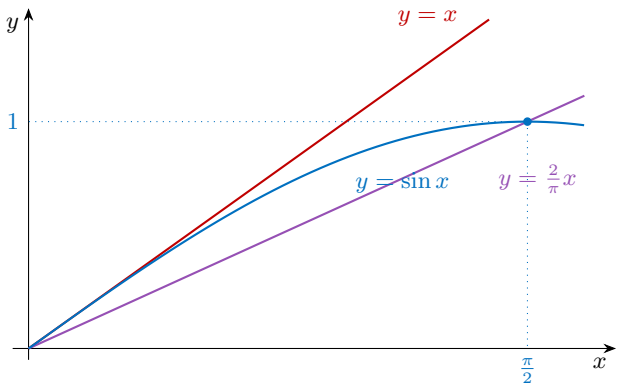
- La fonction exp est convexe : sa courbe est au-dessus de sa tangente en 0, d'où $\mathrm{e}^x \geqslant 1+x$ pour tout réel x. La fonction ln est concave : sa courbe est au-dessous de sa tangente en 1, d'où $\ln x \leqslant x-1$ pour tout $x>0$. Ces deux inégalités, établies au chapitre 5 par l'étude d'une différence, sont symétriques l'une de l'autre par rapport à la première bissectrice.
- *Inégalité arithmético-géométrique.* Soit $x_1,\dots,x_n$ des réels strictement positifs. L'inégalité de Jensen pour la fonction concave ln, avec les coefficients $\frac{1}{n}$, donne

$$\frac{\ln x_1+\cdots+\ln x_n}{n} \leqslant \ln\Big(\frac{x_1+\cdots+x_n}{n}\Big), \qquad \text{soit} \qquad \sqrt[n]{x_1\cdots x_n} \leqslant \frac{x_1+\cdots+x_n}{n},$$

par croissance de exp. La moyenne géométrique est inférieure à la moyenne arithmétique.

- *Inégalité de Jordan.* La fonction sin est concave sur $\big[0\,,\frac{\pi}{2}\big]$, où $\sin''=-\sin\leqslant 0$. Sa courbe est au-dessous de sa tangente $y=x$ en 0, et au-dessus de sa corde joignant $(0,0)$ à $\big(\frac{\pi}{2},1\big)$:

$$\frac{2}{\pi}x \leqslant \sin x \leqslant x \qquad \text{pour tout } x\in\Big[0\,,\frac{\pi}{2}\Big].$$



Sur $\big[0\,,\frac{\pi}{2}\big]$, la courbe du sinus, en bleu, est prise entre sa tangente en 0, en rouge, et la corde violette qui joint ses extrémités : c'est l'image de la concavité.

Remarque : L'erreur la plus répandue est de confondre les deux positions : une fonction convexe est au-dessous de ses cordes, mais *au-dessus* de ses tangentes. Un dessin rapide de la parabole $y=x^2$ suffit à ne pas se tromper. Le chapitre 10 précisera l'écart entre la courbe et sa tangente à l'aide de la dérivée seconde.

9.6.4 Points d'inflexion

Définition : Soit f une fonction dérivable sur I, et a un point intérieur à I. On dit que le point $A=(a,f(a))$ est un **point d'inflexion** de la courbe de f s'il existe $\eta>0$ tel que $[a-\eta\,,a+\eta]\subset I$ et que

$$f \text{ soit convexe sur } [a-\eta\,,a] \text{ et concave sur } [a\,,a+\eta],$$

ou l'inverse. La fonction change de convexité en a.

Proposition 10 : *(Inflexion et tangente.)* Soit f dérivable sur I, et A un point d'inflexion de sa courbe, d'abscisse a. Alors la courbe traverse sa tangente en A : elle est d'un côté de cette tangente sur $[a-\eta\,,a]$, et de l'autre sur $[a\,,a+\eta]$. Si de plus f est deux fois dérivable sur I, alors $f''(a)=0$.

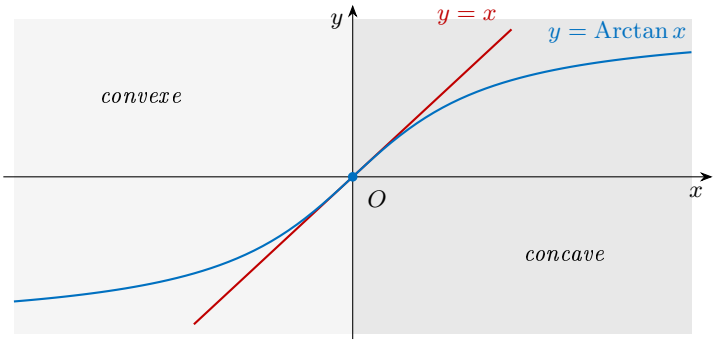
∘ Démonstration :

▷ *Position par rapport à la tangente.* Supposons f convexe sur $[a-\eta\,,a]$ et concave sur $[a\,,a+\eta]$, l'autre cas s'en déduisant en remplaçant f par $-f$. Le théorème 10, appliqué sur l'intervalle $[a-\eta\,,a]$ qui contient a, donne $f(x)\geqslant f(a)+f'(a)(x-a)$ pour $x\in[a-\eta\,,a]$. Appliqué à la fonction convexe $-f$ sur $[a\,,a+\eta]$, il donne $f(x)\leqslant f(a)+f'(a)(x-a)$ pour $x\in[a\,,a+\eta]$.

▷ *Annulation de la dérivée seconde.* Toujours d'après le théorème 10, f' est croissante sur $[a-\eta\,,a]$ et décroissante sur $[a\,,a+\eta]$: elle admet donc en a un maximum local. Comme a est intérieur à I et que f' est dérivable en a, la proposition 2 donne $f''(a)=0$. □

Réciproquement, si f est deux fois dérivable et si f'' s'annule en a *en changeant de signe*, le théorème 11, appliqué sur $[a-\eta\,,a]$ et sur $[a\,,a+\eta]$, montre que A est un point d'inflexion. C'est ainsi qu'on les cherche en pratique : on étudie le signe de f''.

Exemple : Pour $f=\text{Arctan}$, on a $f'(x)=\frac{1}{1+x^2}$ et $f''(x)=-\frac{2x}{(1+x^2)^2}$, qui est positive sur $]-\infty\,,0]$ et négative sur $[0\,,+\infty[$. L'origine est donc un point d'inflexion, où la tangente est la droite $y=x$, et la proposition 10 donne, sans aucune étude de fonction, $\text{Arctan}\,x\leqslant x$ pour $x\geqslant 0$ et $\text{Arctan}\,x\geqslant x$ pour $x\leqslant 0$.

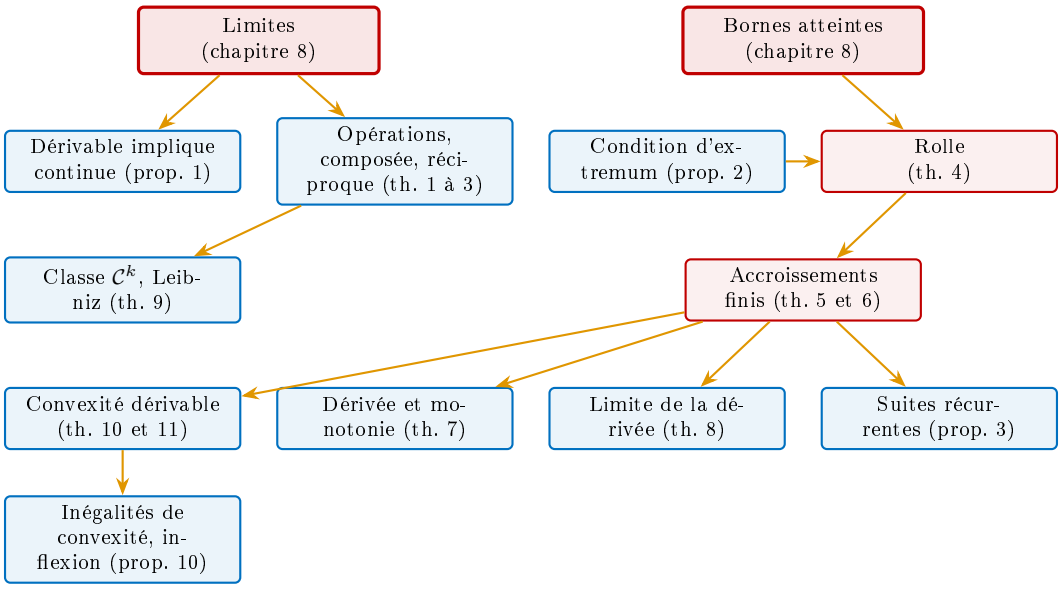


Sur la figure, Arctan est convexe sur fond clair et concave sur fond foncé. Sa courbe, en bleu, est au-dessus de sa tangente en O, en rouge, du côté où elle est convexe, et au-dessous du côté où elle est concave : elle traverse sa tangente au point d'inflexion.

Remarque : L'erreur la plus répandue est de croire que l'annulation de f'' suffit. La fonction $x\mapsto x^4$ vérifie $f''(0)=0$, mais elle est convexe sur $\mathbb{R}$ et sa courbe reste au-dessus de sa tangente en 0 : l'origine n'est pas un point d'inflexion. C'est le changement de signe de f'' qui compte, comme c'était le changement de signe de f' qui comptait pour un extremum.

9.7 Bilan

Tout le chapitre repose sur deux piliers : la définition de la dérivée comme limite, d'où viennent les règles de calcul, et le théorème des bornes atteintes du chapitre 8, d'où vient le théorème de Rolle. Le schéma suivant montre comment les résultats s'enchaînent ; une flèche signifie « sert à démontrer ».



Les deux théorèmes encadrés de rouge sont le cœur du chapitre : tous les résultats globaux, de la monotonie à la convexité, passent par l'égalité des accroissements finis, qui elle-même repose sur Rolle.

Le tableau suivant rassemble les outils du chapitre selon la question posée.

Question	Outil	Référence
étudier la dérivabilité en un point	taux d'accroissement, ou limite de la dérivée	définition, théorème 8
dériver une expression	opérations, composée, réciproque	théorèmes 1, 2 et 3
calculer une dérivée n-ième	formule de Leibniz	théorème 9
chercher un extremum	points critiques et extrémités	proposition 2
montrer qu'une dérivée s'annule	théorème de Rolle	théorème 4
majorer un accroissement	inégalité des accroissements finis	théorème 6, proposition 6
étudier $u_{n+1} = f(u_n)$	point fixe attractif	proposition 3
établir une inégalité	tangente ou corde d'une fonction convexe	théorèmes 10 et 11, proposition 7
chercher un point d'inflexion	changement de signe de f''	proposition 10

Le chapitre 10 prolonge ces idées : la tangente est la meilleure approximation affine d'une fonction dérivable près d'un point, et les développements limités y approcheront une fonction par un polynôme à un ordre quelconque.