Chapitre 13 : Dérivation

MPSI Lycée Camille Jullian

22 février 2023

Toute littérature dérive du péché.

Charles Baudelaire

Les constantes et e^x sont dans le métro.
Un opérateur différentiel terroriste monte dans la rame,
menaçant de dériver tout le monde.
Alors que les constantes paniquent, e^x se moque de lui :
« Vas-y, dérive, je crains rien ».
L'opérateur répond alors : « Tremble, misérable exponentielle, je suis $\dfrac{d}{dy}$ » !

Pour ce dernier chapitre d'analyse du premier semestre, après celui sur la continuité, nouveau retour sur une notion fondamentale avec laquelle vous avez beaucoup travaillé au lycée. Le principe sera d'ailleurs le même que dans le chapitre 12, reprendre une notion bien connue et la revoir en profondeur avec des définitions et démonstrations rigoureuses. Rien de très nouveau donc, si ce n'est que la section des théorèmes classiques va s'enrichir notamment de l'inégalité des accroissements finis, fondamentale pour l'étude des suites récurrentes que nous aborderons en fin de chapitre.

Objectifs du chapitre :

- ne plus hésiter une seconde avant de calculer une dérivée classique (notamment à l'aide de la formule de la dérivée d'une composée).
- maîtriser l'application de l'IAF à l'étude des suites récurrentes.
- savoir exploiter la notion de convexité pour démontrer des inégalités.

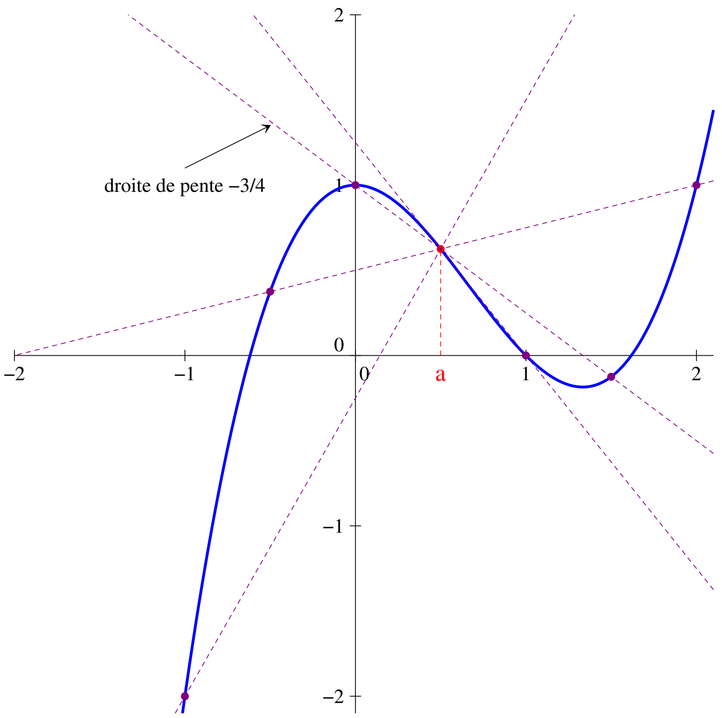
1 Définitions et formulaire.

1.1 Aspect géométrique.

L'idée cachée derrière le calcul de dérivée, que vous utilisez déjà depuis plusieurs années pour étudier les variations de fonctions, est en gros la suivante : les seules fonctions dont le sens de variation est réellement facile à déterminer sont les fonctions affines, pour lesquelles il est simplement donné par le signe du coefficient directeur de la droite représentant la fonction affine. Pour des fonctions plus complexes, on va donc chercher à se ramener au cas d'une droite en cherchant, pour chaque point de la courbe, la droite « la plus proche » de la courbe autour de ce point. C'est ainsi qu'est née la notion de tangente, à laquelle celle de dérivée est intimement liée. Plus précisément :

Définition 1. Soit f une fonction définie sur un intervalle I et $a \in I$, le **taux d'accroissement de f en a** est la fonction définie par $\tau_{a,f}(h) = \dfrac{f(a+h) - f(a)}{h}$.

Remarque 1. Le taux d'accroissement n'est pas défini en 0. Pour $h \neq 0$, $\tau_{a,f}(h)$ représente le coefficient directeur de la droite passant par les points d'abscisse a et $a+h$ de la courbe représentative de f (il s'agit techniquement d'une **corde** de cette courbe représentative).



Sur ce schéma, on a pris $a = \dfrac{1}{2}$ (ce qui correspond au point rouge de la courbe) et tracé plusieurs droites reliant ce point avec d'autres points situés sur la courbe. Ainsi, pour $h = -\dfrac{1}{2}$ ou $h = 1$ (la droite correspondante est la même dans ces deux cas), on aurait $\tau_{a,f}(h) = -\dfrac{3}{4}$

Définition 2. Une fonction f est **dérivable** en a si son taux d'accroissement en a admet une limite finie quand h tend vers 0. On appelle alors **nombre dérivé de f en a** cette limite et on la note $f'(a) = \lim\limits_{h \to 0} \dfrac{f(a+h) - f(a)}{h}$.

Remarque 2. En reprenant l'interprétation géométrique précédente, la droite tracée se rapproche quand h tend vers 0 de la tangente à la courbe représentative de f au point de la courbe d'abscisse a. Le nombre dérivé de f en a est donc le coefficient directeur de cette tangente.

Remarque 3. Pour des raisons pratiques, on aura parfois besoin pour certains calculs d'une définition légèrement différente du nombre dérivé : $f'(a) = \lim\limits_{x \to a} \dfrac{f(x) - f(a)}{x - a}$, qui est équivalente à la précédénte (en posant $h = x - a$, on se ramène en effet à notre première définition).

Exemples :

- Considérons la fonction carré définie par $f(x) = x^2$ et calculons à l'aide de cette définition la dérivée (ou plutôt pour l'instant le nombre dérivé au point d'abscisse a) de f. Le taux d'accroissement de la fonction carré en a vaut $\tau_{a,f}(h) = \dfrac{(a+h)^2 - a^2}{h} = \dfrac{a^2 + 2ha + h^2 - 1}{h} =$

$2a + h$. Ce taux d'accroissement a une limite égale à $2a$ quand h tend vers 0, donc f est dérivable en a et $f'(a) = 2a$ (ce qui correspond bien à la formule que vous connaissez).

- Considérons à présent la fonction racine carrée définie par $g(a) = \sqrt{a}$, le taux d'accroissement de g en a vaut $\tau_{a,g}(h) = \dfrac{\sqrt{a+h} - \sqrt{a}}{h} = \dfrac{(\sqrt{a+h} - \sqrt{a})(\sqrt{a+h} + \sqrt{a})}{h(\sqrt{a+h} + \sqrt{a})} = \dfrac{a+h-a}{h(\sqrt{a+h} + \sqrt{a})} = \dfrac{1}{\sqrt{a+h} + \sqrt{a}}$. Si $a \neq 0$, ce taux d'accroissement a pour limite $\dfrac{1}{2\sqrt{a}}$, ce qui correspond une nouvelle fois à une formule bien connue. Par contre, $\lim\limits_{h \to 0} \tau_0(h) = +\infty$, ce qui prouve que la fonction racine carrée n'est pas dérivable en 0. On a tout de même une interprétation graphique intéressante dans ce cas : la courbe représentative de la fonction racine carrée admet en son point d'abscisse 0 une tangente verticale.

Définition 3. La fonction f est **dérivable à gauche** en a si son taux d'accroissement admet une limite quand h tend vers 0^-. On note alors $f'_g(a) = \lim\limits_{h \to 0^-} \dfrac{f(a+h) - f(a)}{h}$. De même, f est **dérivable à droite** en a si $\tau_a(h)$ admet une limite en 0^+ et on note $f'_d(a) = \lim\limits_{h \to 0^+} \dfrac{f(a+h) - f(a)}{h}$.

Remarque 4. La fonction f est dérivable en a si et seulement si elle y est dérivable à gauche et à droite et que $f'_d(a) = f'_g(a)$.

Définition 4. Dans le cas où $f'_g(a) \neq f'_d(a)$ (ou si une seule des deux limites existe) on dit que la courbe de f admet une (ou deux) **demi-tangente à droite ou à gauche en** a. Si $\tau_a(h)$ admet une limite infinie en 0^+ ou en 0^-, on dit que la courbe de f admet une demi-tangente verticale au point d'abscisse a.

Exemple : Considérons $f(x) = |x|$ et $a = 0$. On a donc $\tau_{0,f}(h) = \dfrac{|h|}{h}$. Si $h > 0$, $\tau_0(h) = \dfrac{h}{h} = 1$, donc $f'_d(0) = 1$; mais si $h < 0$, $\tau_0(h) = \dfrac{-h}{h} = -1$, donc $f'_g(h) = -1$. La fonction valeur absolue n'est donc pas dérivable en 0, mais y admet à gauche une demi-tangente d'équation $y = -x$, et à droite une demi-tangente d'équation $y = x$ (qui sont d'ailleurs confondues avec la courbe).

Définition 5. Une fonction f est **dérivable sur un intervalle** I si elle est dérivable en tout point de I. On appelle alors **fonction dérivée** de f la fonction $f : x \mapsto f'(x)$.

Proposition 1. Soit f une fonction dérivable en a, alors l'équation de la tangente à la courbe représentative de f en a est $y = f'(a)(x-a) + f(a)$.

Proposition 2. Si une fonction f est dérivable en a, alors f est continue en a.

Remarque 5. La réciproque est fausse ! Par exemple la fonction valeur absolue est continue sur $\mathbb{R}$ mais pas dérivable en 0.

Démonstration. Si f est dérivable en a, on sait que $\lim\limits_{h \to 0} \dfrac{f(a+h) - f(a)}{h} = f'(a)$. Autrement dit, $\dfrac{f(a+h) - f(a)}{h} \underset{0}{=} f'(a) + \varepsilon(h)$, avec $\lim\limits_{h \to 0} \varepsilon(h) = 0$. En multipliant tout par h, on obtient $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$. Comme $\lim\limits_{h \to 0} f(a) + hf'(a) + \varepsilon(h) = f(a)$, on a donc $\lim\limits_{h \to 0} f(a+h) = f(a)$, ce qui prouve que f est continue en a. $\square$

Définition 6. On appelle **développement limité à l'ordre** 1 de f en a l'égalité $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$, avec $\lim_{h\to 0} \varepsilon(h) = 0$.

Remarque 6. Cette égalité signifie simplement que, lorsque h est proche de 0, $f(a+h)$ peut être approché par la fonction affine $h \mapsto f(a) + f'(a)h$ (qui n'est autre que la tangente à la courbe en son point d'abscisse a), et que l'erreur commise quand on effectue cette approximation est en ordre de grandeur plus petite que la valeur approchée calculée (le terme $h\varepsilon(h)$ tend plus vite vers 0 que le terme $f'(a)h$ qui le précède). On verra plus tard que cela prouve que cette approximation est la meilleure possible par une fonction affine. On peut généraliser cette notion en approchant la fonction f par un polynome de degré 2, 3 ou plus (mais il faut alors que f soit deux, trois fois dérivable, etc). On parle alors de développement limité à l'ordre 2, 3 ou n., le principe étant par exemple pour l'ordre 2 d'avoir un développement du type $f(a+h) = f(a) + f'(a)h + kh^2 + h^2\varepsilon(h)$, avec $k \in \mathbb{R}$ et $\lim_{h\to 0} \varepsilon(h) = 0$, de façon à avoir la meilleure approximation possible. Il faut pour cela choisir $k = \dfrac{f''(a)}{2}$, nous reverrons en détail ces calculs une fois que nous posséderons un outil (la formule de Taylor) permettant de mieux les comprendre.

1.2 Opérations.

Proposition 3. Soient f et g deux fonctions dérivables en a. Alors $f+g$ est dérivable en a et $(f+g)'(a) = f'(a) + g'(a)$.

Démonstration. En effet, le taux d'accroissement de $f+g$ en a vaut
$\tau_{a,f+g}(h) = \dfrac{f(a+h)+g(a+h)-f(a)-g(a)}{h} = \dfrac{f(a+h)-f(a)}{h} + \dfrac{g(a+h)-g(a)}{h}$. Autrement dit, c'est la somme des taux d'accroissements de f et de g en a. Sa limite existe donc et est égale à la somme des limites de ces taux d'accroissement, c'est-à-dire que $\lim_{h\to 0} \tau_{a,f+g}(h) = f'(a) + g'(a)$, d'où la formule. □

Proposition 4. Soient f et g deux fonctions dérivables en a, alors fg est dérivable en a et $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.

Démonstration. Calculons le taux d'accroissement de la fonction fg en a :
$\tau_{a,fg}(h) = \dfrac{f(a+h)g(a+h) - f(a)g(a)}{h} = \dfrac{f(a+h)g(a+h) - f(a)g(a+h) + f(a)g(a+h) - f(a)g(a)}{h} = g(a+h)\dfrac{f(a+h)-f(a)}{h} + f(a)\dfrac{g(a+h)-g(a)}{h}$. Le premier terme a pour limite $g(a)f'(a)$ quand h tend vers 0 (la fonction g étant dérivable en a, elle y est continue, donc $g(a+h)$ tend vers $g(a)$ quand h tend vers 0), et le second a pour limite $f(a)g'(a)$ puisqu'on reconnait le taux d'accroissement de g. On obtient donc bien la formule attendue. □

Proposition 5. Soit g une fonction dérivable en a, et ne s'annulant pas en a, alors $\dfrac{1}{g}$ est dérivable en a et $\left(\dfrac{1}{g}\right)'(a) = -\dfrac{g'(a)}{g(a)^2}$. Si f est une autre fonction dérivable en a, alors $\dfrac{f}{g}$ est dérivable en a et $\left(\dfrac{f}{g}\right)'(a) = \dfrac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$.

Démonstration. Le taux d'accroissement de $\dfrac{1}{g}$ en a vaut $\tau_{a,\frac{1}{g}}(a) = \dfrac{\frac{1}{g(a+h)} - \frac{1}{g(a)}}{h}$. Il n'est défini que si $g(a+h) \neq 0$, mais on admettra que, si $g(a) \neq 0$ (c'est une des hypothèses de la proposition) et g est continue, alors g ne s'annule pas au voisinage de a (c'est une conséquence de la définition de la limite, il faut trafiquer un peu avec les ε et les η pour faire les choses tout à fait rigoureusement). On peut alors réduire au même dénominateur : $\tau_{a,\frac{1}{g}}(h) = \dfrac{1}{g(a+h)g(a)} \dfrac{g(a) - g(a+h)}{h}$. On reconnait à droite l'opposé du taux d'accroissement de g, qui tend donc vers $-g'(a)$, et le dénominateur à gauche tend vers $g(a)^2$ car g est dérivable donc continue en a.
La deuxième formule s'obtient en appliquant simplement la formule de dérivation d'un produit à f et $\dfrac{1}{g}$: $\left(\dfrac{f}{g}\right)'(a) = f'(a) \times \dfrac{1}{g(a)} - f(a) \times \dfrac{g'(a)}{g(a)^2} = \dfrac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}$. □

Proposition 6. Soient f et g deux fonctions dérivables respectivement en a et en $f(a)$, alors la composée $g \circ f$ est dérivable en a et $(g \circ f)'(a) = f'(a).(g'(f(a))$.

Démonstration. L'idée est de séparer le taux d'accroissement de $g \circ f$ pour faire apparaitre ceux de g et de f. Notons pour cela $b = f(a)$, et k le réel $f(a+h) - f(a)$, alors $\tau_{a,g\circ f}(h) = \dfrac{g \circ f(a+h) - g \circ f(a)}{h} = \dfrac{g(f(a+h)) - g(f(a))}{k} \times \dfrac{f(a+h) - f(a)}{h}$. Le deuxième quotient est le taux d'accroissement de f en a, il converge donc vers $f'(a)$. Mais le premier quotient est en fait aussi un taux d'accroissement : en effet, par définition, $f(a+h) = f(a) + k$, donc il est égal $\dfrac{g(b+k) - g(b)}{k}$, donc à un taux d'accroissement de la fonction g en $b = f(a)$. De plus, la variable k tend bien vers 0 quand h tend vers 0 par continuité de la fonction f en a. On peut donc conclure que ce quotient a pour limite $g'(b) = g'(f(a))$ quand h tend vers 0, ce qui achève la démonstration de la formule.
Il y a en fait un (gros) problème, c'est que ce premier dénominateur risque très fort de s'annuler (quand $f(a+h) = f(a)$) et (contrairement à ce qui se passait pour l'inverse) cela peut se produire une infinité de fois au voisinage de a. Pour corriger cette imprécision, une autre façon de prouver cette propriété est de passer par les développements limités à l'ordre 1. On sait que $f(a+h) = f(a) + hf'(a) + h\varepsilon(h)$, et que $g(b+k) = g(b) + kg'(b) + k\eta(k)$ (avec $\lim\limits_{k\to 0} \eta(k) = 0$, on a modifié les notations pour ne pas engendrer de confusion). On veut désormais calculer $g \circ f(a+h) = g(f(a) + hf'(a) + h\varepsilon(h))$. En posant $b = f(a)$ et $k = hf'(a) + h\varepsilon(h)$ (qui tend bien vers 0 quand h tend vers 0), on a donc $g \circ f(a+h) = g(f(a)) + (hf'(a) + h\varepsilon(h))g'(f(a)) + \eta(hf'(a) + h\varepsilon(h)) = g \circ f(a) + hf'(a)g' \circ f(a) + h\alpha(h)$, avec $\lim\limits_{h\to 0} \alpha(h) = 0$ (tout les termes restants sont des produits de h par des choses qui tendent vers 0). Comme on sait par ailleurs que $g \circ f(a+h) = g \circ f(a) + h(g \circ f)'(a) + h\alpha(h)$, une simple identification donne $(g \circ f)'(a) = f'(a)g' \circ f(a)$. Pour être totalement honnête, cette dernière identification repose sur l'unicité du développement limité à l'ordre 1, unicité que nous n'avons pas vraiment évoquée dans ce cours, et donc encore moins démontrée. □

Proposition 7. Soit f une fonction dérivable et bijective sur un intervalle I, à valeurs dans J. Alors f^{-1} est dérivable en tout point $b \in J$ tel que $f'(f^{-1}(b)) \neq 0$, et dans ce cas $(f^{-1})'(b) = \dfrac{1}{f'(f^{-1}(b))}$.

Remarque 7. Les images des valeurs où la dérivée de f s'annule, qui sont donc les points où la fonction réciproque n'est pas dérivable, correspondent en fait à des endroits où la courbe de f^{-1} admet des tangentes verticales (ce qui se comprend graphiquement puisqu'une tangente horizontale pour f devient après symétrie par rapport à la droite d'équation $y = x$ une tangente verticale pour f^{-1}).

Démonstration. Soient donc $b \in J$ et $a = f^{-1}(b)$. La taux d'accroissement de f^{-1} en b est défini par $\tau_{b,f^{-1}}(h) = \dfrac{f^{-1}(b+h) - f^{-1}(b)}{h} = \dfrac{f^{-1}(b+h) - a}{h}$. La fonction f étant bijective de I sur J, $b + h$ admet un unique antécédent, qu'on va noter c, dans l'intervalle I. On a donc $f(c) = b + h$ et par ailleurs $f(a) = b$, donc $h = f(c) - b = f(c) - f(a)$, ce qui permet d'écrire $\tau_{b,f^{-1}}(h) = \dfrac{c-a}{f(c) - f(a)}$. Pour mettre ce quotient sous une forme plus habituelle (on devrait déjà reconnaitre ici un inverse de taux d'accroissement), on pose $k = c - a$, et on a $\tau_{b,f^{-1}}(h) = \dfrac{k}{f(a+k) - f(a)}$, avec k qui tend vers 0 quand h tend vers 0 par continuité de la fonction f^{-1} (quand h tend vers 0, $c = f^{-1}(b+h)$ tend vers $f^{-1}(b)$, donc vers a). On reconnait donc la limite quand h tend vers 0 de l'inverse du taux d'accroissement de f en a. Si $f'(a) \neq 0$, on a donc $\lim\limits_{h \to 0} \tau_{b,f^{-1}}(h) = \dfrac{1}{f'(a)} = \dfrac{1}{f'(f^{-1}(b))}$. Si $f'(a) = 0$, la limite de $\tau_{b,f^{-1}}(h)$ est infinie, on a donc une tangente verticale. □

Terminons ce paragraphe en rappelant les (rares) cas de fonctions usuelles qui ne sont pas dérivables sur tout leur ensemble de définition (on ne rappellera pas les formules pour les dérivées de fonctions usuelles, qui ont déjà été revues en début d'année), et qui, à l'exception du cas de la valeur absolue, découlent tous du dernier résultat énoncé sur la dérivabilité d'une réciproque :

- la fonction valeur absolue n'est pas dérivable en 0.
- la fonction racine carrée n'est pas dérivable en 0 (réciproque de la fonction carré dont la dérivée s'annule en 0).
- les fonctions arccos et arcsin ne sont pas dérivables en -1 et en 1 (réciproques des fonctions cos et sin dont les dérivées s'annulent respectivement en 0 et en π pour le cosinus, et en $\pm\dfrac{\pi}{2}$ pour le sinus).

1.3 Dérivées d'ordre supérieur.

Définition 7. Soit f une fonction dérivable sur un intervalle I, et telle que f' est elle-même dérivable sur I, alors la dérivée de f' est appelée **dérivée seconde** de la fonction f, et notée f''. On note de même f''' la dérivée tierce de f (sous réserve d'existence), puis plus généralement $f^{(n)}$ la dérivée n-ième de la fonction f.

Définition 8. Une fonction définie sur un intervalle I y est :

- **de classe** $\mathcal{D}^n$ si elle est n fois dérivable sur I
- **de classe** $\mathcal{C}^n$ si de plus sa dérivée n-ème $f^{(n)}$ est continue sur I
- **de classe** $\mathcal{C}^\infty$ si elle dérivable n fois sur I pour tout entier naturel n

Théorème 1. Toutes les fonctions usuelles sont de classe $\mathcal{C}^\infty$ sur tous les intervalles où elles sont dérivables.

Théorème 2. Le caractère $\mathcal{C}^n$ (ou $\mathcal{C}^\infty$) est stable par toutes les opérations usuelles : une somme, produit, quotient (si le dénominateur ne s'annule pas), composée de fonctions $\mathcal{C}^n$ (ou $\mathcal{C}^\infty$) sera également $\mathcal{C}^n$ (ou $\mathcal{C}^\infty$).

Proposition 8. Formule de Leibniz.
Si f et g sont deux fonctions de classe $\mathcal{D}^n$ sur I, alors $(fg)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} f^{(k)} g^{(n-k)}$.

Démonstration. Ce résultat nous rappelle étrangement la formule du binôme de Newton. Il se démontre exactement de la même façon (on ne le fera donc pas). □

Exemple : Appliquée pour $n = 4$, la formule donne par exemple $(fg)^{(4)} = f^{(4)}g + 4f'''g' + 6f''g'' + 4f'g''' + fg^{(4)}$. On fait bien attention au fait que la notation $f^{(0)}$ désigne la fonction f elle-même et pas une fonction constante égale à 1.

2 Théorème des accroissements finis et applications.

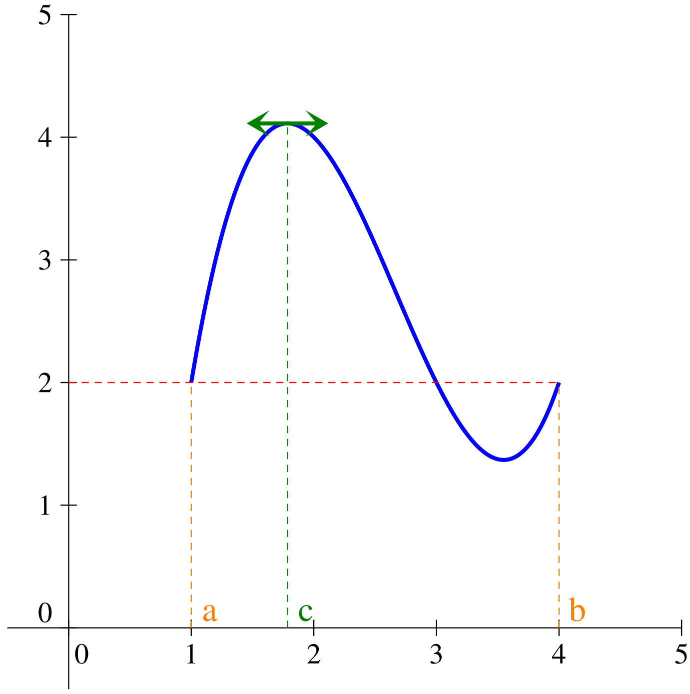
2.1 Énoncés.

Le théorème des accroissements finis, même s'il énonce un résultat qui n'a rien de spectaculaire (et même rien de très utile en tant que tel), est un outil absolument fondamental en analyse puisque c'est notamment grâce à lui qu'on démontre le lien entre signe de la dérivée et sens de variation d'une fonction qui est à la base de tous les tableaux de variations que vous avez dressés depuis la moyenne section de maternelle. Avant de l'énoncer, on a besoin de quelques résultats préliminaires, dont le premier est lui-même un cas particulier d'information habituellement présente dans un tableau de variations.

Proposition 9. Soit f une fonction dérivable sur un intervalle borné $]a, b[$ et $x \in]a, b[$. Si x est un point en lequel f atteint un extremum local, alors $f'(x) = 0$.

Démonstration. Supposons par exemple qu'il s'agisse d'un maximum (l'autre cas est très similaire). Le taux d'accroissement de f en x est défini par $\tau_{x,f}(h) = \dfrac{f(x+h) - f(x)}{h}$. Or, au voisinage de x, on aura $f(x+h) \leqslant f(x)$ puisque $f(x)$ est un maximum local. On en déduit que $\forall h < 0$ (et tel que $x + h$ appartienne au voisinage en question), $\tau_{x,f}(h) \geqslant 0$, donc $f'(x) = \lim\limits_{h \to 0^-} \tau_{x,f}(h) \geqslant 0$ (le point x étant situé dans un intervalle ouvert, on peut toujours calculer cette limite à gauche, tout comme la limite à droite qui va suivre). Mais de même $\forall h > 0$, $\tau_{x,f}(h) \leqslant 0$, donc $f'(x) = \lim\limits_{h \to 0^+} \tau_{x,f}(h) \leqslant 0$. Finalement, on a nécessairement $f'(x) = 0$. □

Théorème 3. Théorème de Rolle.
Soit f une fonction continue sur un segment $[a, b]$ et dérivable sur $]a, b[$, telle que $f(a) = f(b)$, alors $\exists c \in]a, b[$, $f'(c) = 0$.

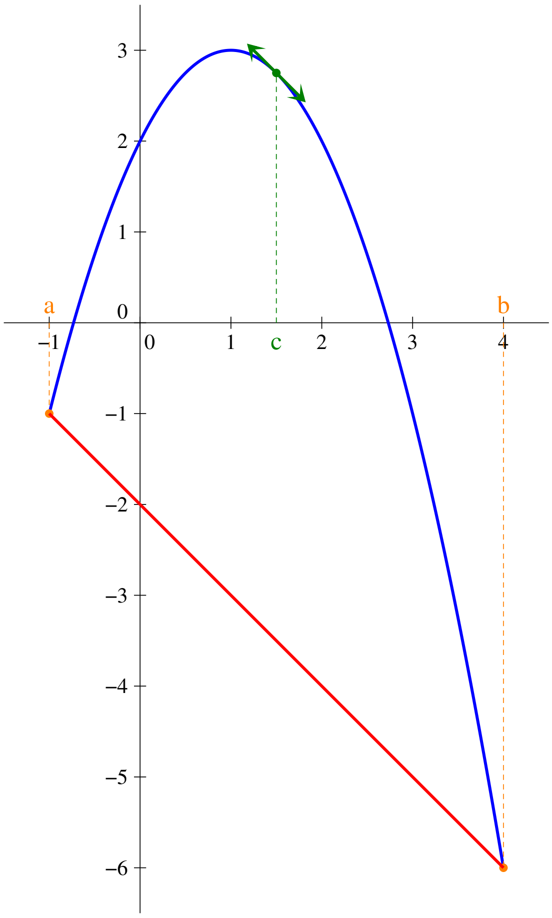


Démonstration. Commençons par éliminer le cas où la fonction f est constante sur $[a, b]$ puisque dans ce cas la dérivée de f est nulle, donc le théorème est manifestement vérifié.
La fonction f étant dérivable, elle est continue sur $[a, b]$, donc y atteint un maximum M et un minimum m d'après le théorème du maximum. Si on suppose f non constante, l'un des deux, par exemple M (dans l'autre cas, la démonstration est similaire), est distinct de $f(a)$ (et de $f(b)$ qui lui est égal), donc atteint en un réel $c \in]a, b[$. D'après la propriété précédente, $f'(c) = 0$. □

Remarque 8. On peut proposer une interprétation cinématique de ce résultat : si un objet se déplace sur un axe (déplacement à une dimension) et qu'il revient au bout d'un certain temps à son point de départ, il y aura forcément eu un instant lors de son déplacement où sa vitesse instantanée aura été nulle (ce qui est effectivement indispensable pour qu'il puisse faire demi-tour).

Théorème 4. Théorème des accroissements finis.
Soit f une fonction continue sur un segment $[a, b]$ et dérivable sur $]a, b[$, alors $\exists c \in]a, b[$, $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.

Remarque 9. Autrement dit, il existe un point où la tangente est parallèle à la corde passant par les points de la courbe de coordonnées $(a, f(a))$ et $(b, f(b))$. Là encore, on peut fournir une interprétation cinématique de notre résultat : toujours en supposant un mouvement unidimensionnel, mais en supprimant l'hypothèse de retour au point de départ, il y aura un instant pendant le déplacement où la vitesse instantanée du mobile aura coïncidé avec sa vitesse moyenne sur l'ensemble du déplacement. Le théorème fonctionne d'ailleurs aussi si le mouvement n'est pas unidimensionnel, quitte à prendre en compte la norme de la vitesse : si on parcourt 24 kilomètres en deux heures lors d'un jogging, on aura eu à un moment donné une vitesse instantanée de 12 km.h^{-1}.



Démonstration. Le principe est de se ramener au théorème précédent. Définissons une deuxième fonction g par $g(x) = \dfrac{f(b)-f(a)}{b-a}x - f(x)$ (ce qui correspond, à une constante près, à l'écart entre la courbe représentative de f et la droite passant par les points $(a, f(a))$ et $(b, f(b))$). Cette fonction est dérivable sur $]a, b[$ puisque f l'est et vérifie $g(b) - g(a) = \dfrac{f(b)-f(a)}{b-a}b - f(b) - \dfrac{f(b)-f(a)}{b-a}a + f(a) = \dfrac{f(b)-f(a)}{b-a}(b-a) - f(b) + f(a) = 0$, c'est-à-dire que $g(b) = g(a)$. On peut donc lui appliquer le théorème de Rolle : $\exists c \in]a, b[$, $g'(c) = 0$. Or, $g'(x) = \dfrac{f(b)-f(a)}{b-a} - f'(x)$, donc on a $f'(c) = \dfrac{f(b)-f(a)}{b-a}$, ce qu'on cherchait à prouver. □

Théorème 5. Soit f une fonction dérivable sur un intervalle I, alors f est croissante sur I si et seulement si f' est positive sur I. De même, f est décroissante sur I si et seulement si f' est négative sur I.

Démonstration. Supposons f croissante sur I, et soit $a \in I$, considérons le taux d'accroissement de f en a : $\tau_{a,f}(h) = \dfrac{f(a+h)-f(a)}{h}$. Ce taux d'accroissement est toujours positif, puisque numérateur et dénominateur sont de même signe (par croissance de f, $f(a+h) - f(a) < 0$ si $h < 0$, et $f(a+h) - f(a) > 0$ si $h > 0$). En passant à la limite, on en déduit $f'(a) \geqslant 0$. Comme cette inégalité est vraie pour tout a, la dérivée f' est bien positive sur l'intervalle I. Le sens réciproque est bien entendu le plus important : si $f'(x) \geqslant 0$ sur tout l'intervalle I, en choisissant un élément $y > x$, le théorème des accroissements finis assure l'existence d'un réel x tel que $\dfrac{f(y) - f(x)}{y - x} = f'(c)$. Comme $f'(c) \geqslant 0$ et $y - x \geqslant 0$, on a donc $f(y) - f(x) \geqslant 0$, ce qui prouve que f est croissante sur I. La preuve dans le cas de la décroissance est identique aux signes près. □

Remarque 10. Ces théorèmes seront bien entendus utilisés sans être cités lors de l'étude des variations de fonctions, comme vous en avez déjà l'habitude. Mais vous avez désormais une preuve complète de ces résultats très classiques.

Théorème 6. Théorème du prolongement de la dérivée.

Soit f une fonction continue sur un segment $[a, b]$ et dérivable sur $]a, b]$. Si la dérivée f' de la fonction f admet une limite finie l quand x tend vers a, alors f est dérivable en a et $f'(a) = l$.

Démonstration. Considérons le taux d'accroissement de f en a : $\tau_{a,f}(h) = \dfrac{f(a+h) - f(a)}{h}$ (ici, h sera nécessairement positif pusique f n'est définie qu'à droite de a). D'après le théorème des accroissements finis, on peut écrire $\tau_{a,f}(h) = f'(c_h)$, où c_h est une constante (dépendant de h) appartenant à l'intervalle $]a, a+h[$. Si on fait tendre h vers 0, d'après le théorème des gendarmes, c_h aura pour limite a. Alors, les hypothèses du théorème nous permettent d'affirmer que $\lim\limits_{h \to 0} f'(c_h) = l$, ce qui prouve bien que la fonction f est dérivable en a, puisque son taux d'accroissement y tend vers l. □

Exemple : Ce théorème sera souvent appliqué dans le cas où on prolonge une fonction par continuité, pour déterminer si le prolongement effectué est dérivable ou non. Il évite de revenir au calcul du taux d'accroissement (qui est toutefois rarement plus complexe, les deux options sont en pratique équivalentes). Considérons la fonction $f : x \mapsto x^2 \ln(x)$. Cette fonction est définie et de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^{+*}$ et peut se prolonger par continuité en 0 en une fonction g vérifiant $g(0) = 0$ (par croissance comparée). Par ailleurs, $f'(x) = 2x \ln(x) + x$ a certainement aussi une limite nulle en 0 (toujours de la croissance comparée ici). Le théorème de prolongement de la dérivée permet alors d'affirmer que la fonction prolongée g est dérivable en 0, et que $g'(0) = 0$. Cette information est essentielle pour tracer une allure précise de la courbe au voisinage de 0.

Remarque 11. On pourra également utiliser la variante suivante du théorème de prolongement de la dérivée : sous les mêmes hypothèses, si la dérivée f' admet en a une limite infinie, alors f n'est pas dérivable en a mais sa courbe y admet une (demi)-tangente verticale.

Proposition 10. Inégalité des accroissements finis (IAF).

Soit $f$une fonction dérivable sur un intervalle I :

- s'il existe deux réels m et M tels que, $\forall x \in I$, $m \leqslant f'(x) \leqslant M$, alors $\forall (x, y) \in I^2$ tels que $x < y$, on aura $m(y - x) \leqslant f(y) - f(x) \leqslant M(y - x)$.
- s'il existe un réel M tel que $\forall x \in I$, $|f'(x)| \leqslant M$, alors $\forall (x, y) \in I^2$, on aura $|f(y) - f(x)| \leqslant M|y - x|$.

Remarque 12. On peut énoncer autrement l'IAF dans sa seconde version : si $|f'|$ est majorée par K sur un intervalle I, alors f est K-Lipschitzienne sur I.

La première version de l'IAF est plus précise que la deuxième, mais en pratique c'est surtout la deuxième qu'on utilisera, notamment quand on l'appliquera à l'étude des suites récurrentes.

Démonstration. C'est vraiment une application directe du théorème des accroissements finis. Démontrons la première version : il existe un réel $c \in]x, y[$ tel que $\dfrac{f(y) - f(x)}{y - x} = f'(c)$. Les hypothèses imposent que $m \leqslant f'(c) \leqslant M$, il suffit donc de multiplier l'encadrement par $y - x$ (qui est par hypothèse positif) pour obtenir le résultat souhaité. □

Remarque 13. Ces inégalités ont une interprétation cinématique assez évidente : si on reprend l'exemple du coureur qui fait son jogging, s'il a couru deux heures avec une vitesse instantanée maximale de 15 kilomètres par heure, il aura parcouru au maximum 30 kilomètres lors de ses deux heures de jogging.

2.2 Application à l'étude de suites récurrentes.

Définition 9. Une **suite récurrente** est une suite définie par une relation de récurrence du type $u_{n+1} = f(u_n)$, où la fonction f est une fonction continue.

Comme pour l'étude des suites implicites dans le chapitre précédent, le but de ce paragraphe est de présenter les méthodes principales d'étude des suites récurrentes sur un exemple simple, sans réellement énoncer de résultats à retenir. Il faut par contre vraiment connaitre les étapes principales de la méthode, notamment en ce qui concerne l'application de l'IAF : les récurrences sont toujours les mêmes et apparaissent au même endroit, on doit donc savoir à quel moment y recourir. Tout de même, un seul résultat théorique :

Théorème 7. Si une suite (u_n) vérifiant la relation de récurrence $u_{n+1} = f(u_n)$ converge, alors sa limite l est un point fixe de la fonction f, c'est-à-dire une solution de l'équation $f(x) = x$.

Remarque 14. Il est bien entendu tout à fait possible qu'une telle suite ne converge pas, les points fixes de f représentent donc simplement les limites potentielles d'une suite récurrente faisant intervenir la fonction f.

Démonstration. Il s'agit simplement de passer à la limite dans la relation de récurrence. Comme u_{n+1} tend vers l et $f(u_n)$ vers $f(l)$, l'égalité $f(l) = l$ en découle immédiatement. Notons l'importance de la continuité de la fonction f pour ce calcul. □

Ajoutons un tout petit peu de vocabulaire avant d'étudier notre exemple :

Définition 10. Un intervalle I est **stable** par la fonction f si $f(I) \subset I$.

Ainsi, une suite récurrente u_n pour laquelle $u_0 \in I$ aura **tous** ses termes dans l'intervalle I si l'intervalle est stable (c'est une récurrence triviale, mais qu'il faut citer à chaque fois sur une copie).

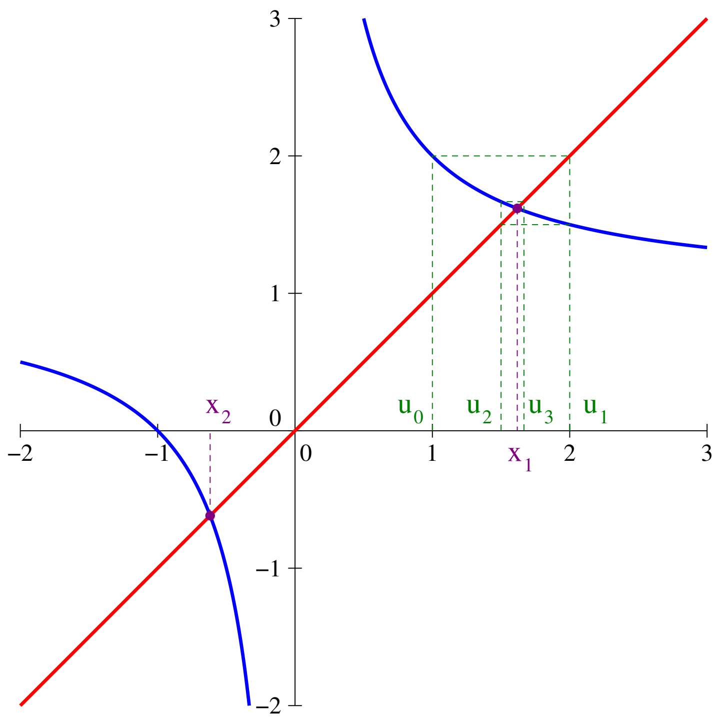
Exemple : Considérons la suite définie par $u_0 = 1$ et $\forall n \in \mathbb{N}$, $u_{n+1} = 1 + \dfrac{1}{u_n}$. On posera donc pour notre étude $f(x) = 1 + \dfrac{1}{x}$.

Étude de la fonction f et détermination des points fixes.

On commence toujours une étude de suite récurrente par l'étude de la fonction f, en lui adjoignant celle du signe de $f(x)-x$ (qui donnera en passant la valeur des points fixes éventuels de la fonction). Ici, f est dérivable sur $\mathbb{R}^*$, de dérivée $f'(x) = -\dfrac{1}{x^2}$. Elle est donc décroissante sur $]-\infty,0[$ et sur $]0,+\infty[$, avec des limites égales à 1 en $\pm\infty$ et infinies à gauche et à droite de 0 (on ne détaille pas, tout ça est très facile). Par ailleurs, $f(x)-x = 1+\dfrac{1}{x}-x = \dfrac{x+1-x^2}{x}$. Le numérateur a pour discriminant $\Delta = 5$, et s'annule en $x_1 = \dfrac{-1-\sqrt{5}}{-2} = \dfrac{1+\sqrt{5}}{2}$, et en $x_2 = \dfrac{1-\sqrt{5}}{2}$, qui sont donc les deux points fixes de f. On va ajouter dans le tableau de variations le signe de $f(x)-x$:

x	$-\infty$		x_2		0		x_1		$+\infty$
$f(x)$	1	↘	x_2	↘	$-\infty$ ‖ $+\infty$	↘	x_1	↘	1
$f(x)-x$		$+$	0	$-$	‖	$+$	0	$-$	

Tant qu'on y est, une représentation graphique ne peut pas faire de mal, et permet de visualiser aisément les principales propriétés de la suite, à condition d'indiquer sur le graphique la droite d'équation $y = x$. On peut même visualiser les premiers termes de la suite grâce à la construction géométrique simple suivante : on part de l'abscisse u_0 puis on « monte » jusqu'au point de la courbe d'abscisse u_0 (et donc d'ordonnée u_1 puisque par définition $f(u_0) = u_1$). On trace ensuite à partir de ce point une horizontale jusqu'à atteindre la droite d'équation $y = x$ (qu'on coupera donc au point de coordonnées (u_1, u_1)) et on « redescend » jusqu'à l'axe des abscisses pour se retrouver à l'abscisse u_1. Il ne reste plus qu'à itérer pour les termes suivants.



Ici, on constate que la suite ne semble pas monotone, mais semble par contre converger vers x_1. Si on veut être plus précis, la sous-suite (u_{2n}) des termes d'indices pairs semble croissante, et la sous-suite (u_{2n+1}) des termes d'indices impairs semble décroissante, avec adjacence des deux suites vers leur limite commune x_1. On pourrait en fait prouver beaucoup plus généralement que la suite récurrente est monotone lorsque la fonction f est croissante (sur l'intervalle stable où vont se situer tous les termes de la suite), et qu'on aura toujours la situation de notre exemple (convergence « en escargot ») si f est décroissante. Le tout bien sûr à condition que la suite converge.

Détermination d'un intervalle stable.

L'observation du tableau de variations, éventuellement assortie du calcul de quelques valeurs, permet de trouver facilement des intervalles stables. Ici, l'intervalle le plus naturel serait l'intervalle $[1, 2]$ (qui est effectivement stable), mais on va plutôt prendre $I = \left[\frac{3}{2}, 2\right]$ (pour des raisons qu'on expliquera ensuite). Bien sûr, u_0 n'appartient pas à cet intervalle mais ce n'est pas très gênant. Vérifions donc que I est stable par f : $f\left(\frac{3}{2}\right) = \frac{5}{3} < 2$, $f(2) = \frac{3}{2}$ et f est décroissante sur I, donc $f\left(\left[\frac{3}{2}, 2\right]\right) = \left[\frac{3}{2}, \frac{5}{3}\right]$ et notre intervalle est bien stable. On démontre alors facilement par récurrence que, $\forall n \geqslant 1$, $u_n \in I$: c'est vrai pour $u_1 = 2$, et si on le suppose vrai pour u_n, alors $u_{n+1} = f(u_n) \in I$ par stabilité de l'intervalle.

Utilisation de l'IAF.

Il est important de bien vérifier toutes les hypothèses de l'IAF avant de l'appliquer. Sur notre intervalle I, on peut majorer la valeur absolue de la dérivée $|f'(x)| = \frac{1}{x^2}$ par $\frac{4}{9}$. On peut alors appliquer l'IAF en prenant $y = u_n$ et $x = x_1$ (on choisira toujours le terme général de la suite et le point fixe qui sera la limite pour appliquer l'IAF dans ce genre de cas), qui sont bien tous les deux dans l'intervalle I. On obtient, puisque $|f'|$ est majorée par $\frac{4}{9}$, $|f(u_n) - f(x_1)| \leqslant \frac{4}{9}|u_n - x_1|$, soit $|u_{n+1} - x_1| \leqslant \frac{4}{9}|u_n - x_1|$. Ce résultat permet déjà de comprendre intuitivement pourquoi la suite va forcément converger vers x_1 : à chaque nouvelle étape, la distance entre la suite et la limite va être diminuée (multipliée par un coefficient inférieur à $\frac{4}{9}$), ce qui revient à dire qu'on se rapproche en permanence de x_1. On remarque que, pour que ce raisonnement puisse fonctionner, il est nécessaire que le majorant obtenu pour $|f'|$ soit strictement inférieur à 1, ce qui n'aurait pas été le cas sur l'intervalle $[1, 2]$. En fait, on peut plus généralement constater que, si l est un point fixe de la fonction f vérifiant $|f'(l)| < 1$, la suite va souvent converger vers l (on parle de point fixe **attractif** dans ce cas), alors que si $|f'(l)| > 1$, la suite ne pourra presque jamais converger vers l, sauf si elle est stationnaire (on parle dans ce cas de point fixe **répulsif**). Ce vocabulaire n'est pas à connaitre.

Une fois appliquée l'IAF, les dernières étapes sont toujours les mêmes : on prouve par récurrence une inégalité majorant la distance entre u_n et sa limite directement en fonction de n. Ici on va prouver que $\forall n \in \mathbb{N}^*$, $|u_n - x_1| \leqslant \left(\frac{4}{9}\right)^{n-1}$. En effet, au rang 1, $|u_1 - x_1| \leqslant 1$ puisque $x_1 \in [1, 2]$. En supposant ensuite la propriété vraie au rang n, on peut écrire $|u_{n+1} - x_1| \leqslant \frac{4}{9}|u_n - x_1| \leqslant \frac{4}{9} \times \left(\frac{4}{9}\right)^{n-1} = \left(\frac{4}{9}\right)^n$ en appliquant successivement l'IAF puis l'hypothèse de récurrence. On conclut enfin à l'aide du théorème des gendarmes : comme $\lim_{n \to +\infty} \left(\frac{4}{9}\right)^n = 0$, et comme $0 \leqslant |u_n - x_1| \leqslant \left(\frac{4}{9}\right)^{n-1}$, on aura $\lim_{n \to +\infty} |u_n - x_1| = 0$, c'est-à-dire que $\lim_{n \to +\infty} u_n = x_1 = \frac{1 + \sqrt{5}}{2}$. Si on le souhaite, on peut même exploiter la majoration obtenue pour étudier la vitesse de convergence de la suite. Sur notre exemple, si on veut déterminer une valeur pour laquelle $|u_n - x_1| \leqslant 10^{-2}$ (autrement dit, u_n sera une valeur approchée de sa limite avec une précision de deux décimales), une condition suffisante (mais pas du tout nécessaire, ne mettez surtout pas d'équivalence dans ce cas!) est $\left(\frac{4}{9}\right)^{n-1} \leqslant 10^{-2}$, soit $(n-1)\ln\left(\frac{4}{9}\right) \leqslant -2\ln(10)$, donc $n \geqslant 1 + \frac{\ln(10)}{\ln(3) - \ln(2)}$ (on fait bien sûr attention au changement de sens de l'inégalité quand on divise par $\ln(4) - \ln(9)$ qui est négatif). Plus concrètement (on effectue l'application numérique à la calculatrice) on peut choisir $n = 7$ (en pratique, la valeur approchée est sûrement meilleure que prévue car la majoration utilisée est loin d'être optimale).

3 Généralisation aux suites et fonctions complexes.

3.1 Limites de suites complexes.

Définition 11. Une suite complexe (z_n) converge vers une limite $l \in \mathbb{C}$ si

$$\forall \varepsilon > 0,\ \exists n_0 \in \mathbb{N},\ \forall n \geqslant n_0,\ |z_n - l| \leqslant \varepsilon.$$

Remarque 15. Il s'agit rigoureusement de la même définition que pour les suites réelles, mais bien sur, la distance est ici mesurée par un module et plus par une valeur absolue. Si on représente les termes de la suite dans le plan complexe, cela revient à imposer qu'ils vont tous se retrouver dans un disque de rayon ε autour de l, quitte à attendre suffisamment longtemps.
Notons enfin que la notion de limite infinie (et surtout égale à $+\infty$ ou à $-\infty$) n'a pas grand sens pour une suite complexe. On se contentera donc d'étudier une éventuellement limite infinie de $|z_n|$, ce qui ramène évidemment le calcul à une suite réelle.

Proposition 11. La suite (z_n) converge si et seulement si les deux suites réelles $(\mathrm{Re}\ (z_n))$ et $(\mathrm{Im}\ (z_n))$ convergent. Dans ce cas, $\lim\limits_{n\to+\infty} z_n = \lim\limits_{n\to+\infty} \mathrm{Re}\ (z_n) + i \times \lim\limits_{n\to+\infty} \mathrm{Im}\ (z_n)$.

Autrement dit, on ramènera toujours l'étude de la convergence des suites complexes à celle de (deux) suites réelles. Il arrivera plus rarement qu'on étudie les suites réelles formées par le module et l'argument de z_n pour aboutir au même genre de conclusion.

Parmi les résultats classiques vus dans le chapitre d'étude des suites réelles, tous ceux qui font intervenir des inégalités (convergence monotone, théorème des gendarmes, suites adjacentes) sont bien entendu inapplicables aux suites complexes (mais on peut bien sûr les appliquer à la partie réelle ou à la partie imaginaire de la suite), ce qui laisse en pratique très peu d'outils d'étude directe de ces suites.

3.2 Fonctions complexes.

Il n'est question ici que de fonctions $f : I \to \mathbb{C}$, où $I \subset \mathbb{R}$. Les fonctions dont la variable est elle-même complexe sont l'objet de tout un pan de l'analyse, appelé analyse complexe, utilisant des méthodes très différentes de celles étudiées dans ce chapitre (et pas du tout à votre programme).

Définition 12. La fonction f admet pour limite $l \in \mathbb{C}$ quand x tend vers $a \in I$ si $\forall \varepsilon > 0$, $\exists \eta > 0$, $|x - a| < \eta \Rightarrow |f(x) - l| < \varepsilon$.

Remarque 16. C'est encore une fois la même définition que pour une fonction réelle. On aurait les mêmes difficultés à définir des limites infinies que pour les suites réelles.

Proposition 12. La fonction f admet une limite l en a si et seulement si $\mathrm{Re}\ (f)$ et $\mathrm{Im}\ (f)$ admettent des limites respectives l_1 et l_2 quand x tend vers a, et on a alors $l = l_1 + il_2$.

Définition 13. La fonction f est **continue** en a si $\lim_{x \to a} f(x) = f(a)$.

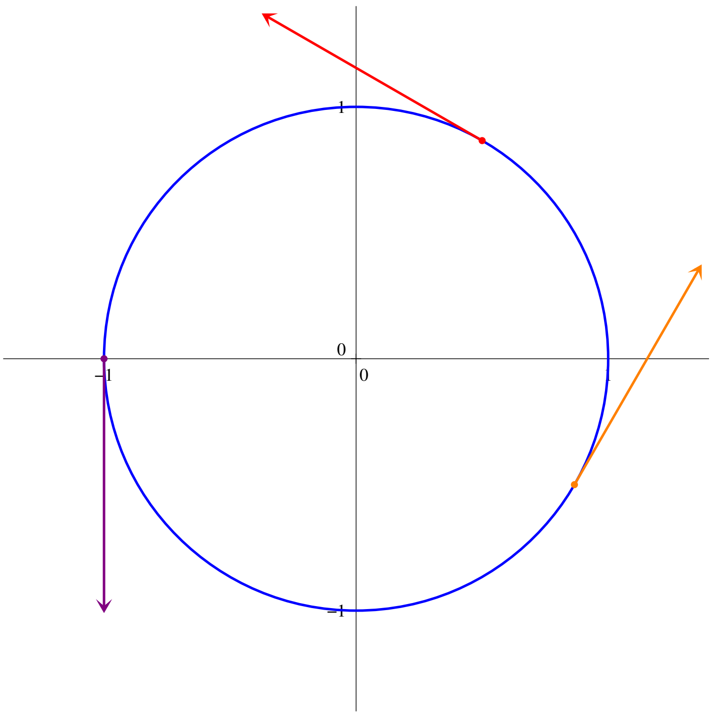
La fonction f est dérivable en a si son taux d'accroissement $\tau_{a,f}(h) = \dfrac{f(a+h) - f(a)}{h}$ admet une limite finie l lorsque x tend vers a. On note alors $f'(a) = l$.

Les autres notions (continuité sur un intervalle, prolongement par continuité) sont identiques au cas des fonctions réelles.

On peut définir des notions de dérivée à gauche ou à droite comme dans le cas réel. Par contre, l'interprétation géométrique de la dérivée en termes de tangente est plus compliquée (cf l'exemple qui suit la propriété). Pire, la notion de variations pour une fonction complexe n'existant pas, le calcul même de la dérivée perd une grande partie de son intérêt !

Proposition 13. La fonction f est dérivable en a si et seulement si Re (f) et Im (f) sont dérivables en a, et on a alors $f'(a) = \text{Re } (f)'(a) + i\text{Im } (f)'(a)$.

Exemple : Si on pose $f(t) = e^{it}$ (fonction définie sur $\mathbb{R}$), on peut écrire $f(t) = \cos(t) + i\sin(t)$, donc, en dérivant séparément les parties réelle et imaginaire, $f'(t) = -\sin(t) + i\cos(t) = e^{i(t+\frac{\pi}{2})} = ie^{it}$. On remarque que la dérivée de cette exponentielle complexe se calcule comme celles des exponentielles réelles. Si on souhaite comprendre géométriquement la notion de dérivée pour cette fonction, on peut représenter la trajectoire correspondant à cette fonction dans le plan complexe sous forme de courbe paramétrée (on place tous les points d'affixe $f(t)$ lorsque t varie, sans faire varier la variable t sur un axe du repère). Ici, on obtient une trajectoire circulaire (on parcourt le cercle trigonométrique). La valeur (complexe) de $f'(t)$ correspond en fait à l'affixe du vecteur tangent à cette trajectoire au point de paramètre t (ici, cette affixe a toujours un module 1, ce qui prouve que la trajectoire est parcourue à vitesse uniforme). Un petit schéma ci-dessous, avec les points et les vecteurs tangents correspondant à $t_1 = \dfrac{\pi}{3}$ (en rouge), $t_2 = \pi$ (en violet) et $t_3 = \dfrac{11\pi}{6}$ (en orange) :



Tout le formulaire de calcul de dérivées (y compris la formule de Leibniz) reste valable pour des fonctions complexes. Il est toutefois moins utile que pour les fonctions réelles, car il existe beaucoup moins de fonction usuelles sur $\mathbb{C}$.

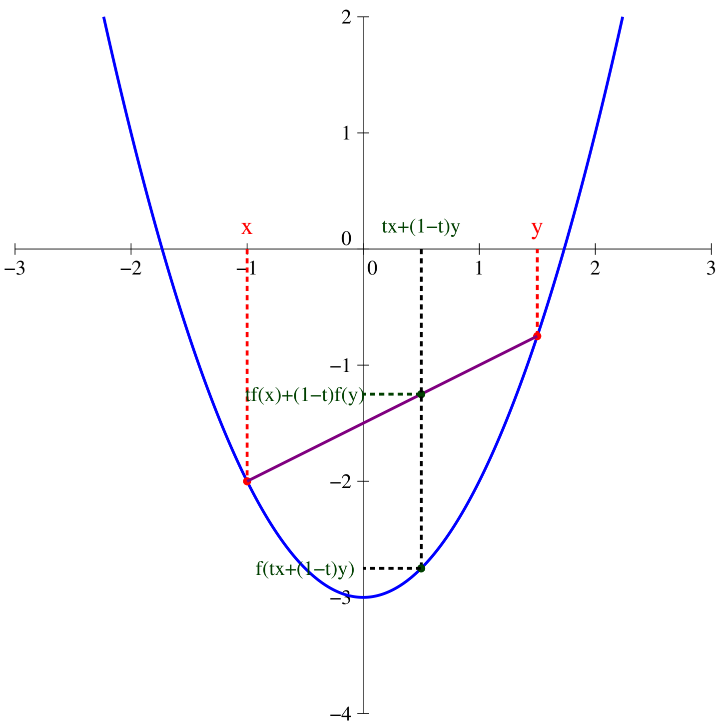
Parmi les théorèmes énoncés plus haut dans ce chapitre, par contre, presque plus rien d'intéressant pour le cas complexe : pas de théorème des accroissements finis, pas de théorème de Rolle. Par exemple, en reprenant $f(t) = e^{it}$, la dérivée $f'(t) = ie^{it}$ ne s'annule jamais, alors que $f(0) = f(2\pi) = 1$ (c'est logique dans la mesure où, lors d'un déplacement en deux dimensions, on peut très bien revenir à son point de départ sans jamais avoir eu une vitesse nulle). Par contre, assez curieusement, l'IAF, du moins sous sa forme « valeur absolue » (qui sera ici remplacée par un module) reste vraie.

4 Convexité.

La notion de convexité/concavité d'une courbe, que vous avez déjà abordée l'an dernier, permet de compléter celle de croissance/décroissance pour visualiser l'allure générale de la courbe représentative d'une fonction. Elle permet également de démontrer énormément d'inégalités liées à la position des courbes représentatives de certaines fonctions par rapport à des droites remarquables (cordes, tangentes). C'est surtout cet aspect qui sera développé dans cette partie de cours, même si le lien avec le signe de f'' quand la fonction est de classe $\mathcal{C}^2$ sera bien sûr démontré.

Définition 14. Soit f une fonction définie sur un intervalle I, alors f est **convexe** sur I si, $\forall (x, y) \in I^2$, $\forall t \in [0, 1]$, $f(tx + (1-t)y) \leqslant tf(x) + (1-t)f(y)$. Symétriquement, f est **concave** sur I si, avec les mêmes quantifications, $f(tx + (1-t)y) \geqslant tf(x) + (1-t)f(y)$.

Cette définition signifie simplement la chose suivante : la courbe représentative d'une fonction convexe (respectivement concave) est située en-dessous (respectivement au-dessus) de chacune de ses cordes. En effet, avec les quantifications données dans notre définition, $tx + (1-t)y$ parcourt toutes les valeurs de l'intervalle $[x, y]$ (en supposant par exemple $x \leqslant y$), et $tf(x) + (1-t)f(y)$ correspond à l'ordonnée du point d'abscisse $tx + (1-t)y$ sur la corde reliant les deux points de coordonnées $(x, f(x))$ et $(y, f(y))$. Un petit dessin sera certainement plus clair :



Plus généralement, une courbe de fonction convexe est située sous ses **sécantes** (droite coupant la courbe en deux points, le segment reliant les deux points est alors une corde de la courbe) entre les deux points d'intersection, et au-dessus à gauche du premier point d'intersection et à droite du second. C'est le contraire pour une fonction concave. En effet, notons x et y les abscisses des deux points d'intersection, et supposons $x < y \leqslant z$. Dans le cas où f est convexe, on veut prouver que $f(z) \geqslant \frac{f(y)-f(x)}{y-x}(z-y)+f(y)$ (le membre de droite de cette inégalité est une équation de la sécante : en effet il s'agit d'une équation de droite en la variable z, dont le coefficient directeur coïncide avec celui de la sécante, et qui prend la valeur $f(y)$ lorsque $z = y$, donc qui passe par le point de coordonnées $(y, f(y))$), soit en mettant tout au même dénominateur $(y-x)f(z)+(z-y)f(x) \geqslant (z-x)f(x)$, ou encore $f(y) \leqslant \left(1-\frac{y-x}{z-x}\right)f(x)+\frac{y-x}{z-x}f(z)$. Cette dernière inégalité n'est rien d'autre qu'une inégalité de convexité, avec $\frac{y-x}{z-x} \in [0,1]$ et $\left(1-\frac{y-x}{z-x}\right)x+\frac{y-x}{z-x}z = \frac{(z-y)x+(y-x)z}{z-x} = y$. On traiterait le cas $z \leqslant x < y$ de la même façon.

Théorème 8. Inégalité de Jensen.

Soit f une fonction convexe sur un intervalle I et $(x_1, x_2, \dots, x_n) \in I^n$, alors pour tout n-uplet de réels $(t_1, t_2, \dots, t_n) \in [0,1]^n$ tels que $\sum_{k=1}^{n} t_k = 1$, on aura

$$f\left(\sum_{k=1}^{n} t_k x_k\right) \leqslant \sum_{k=1}^{n} t_k f(x_k)$$

Démonstration. On va procéder par récurrence sur n. Si $n = 1$, il n'y a rien à prouver. Si $n = 2$, il s'agit tout simplement de la définition de la convexité (puisque $t_1 + t_2 = 1$, on a $t_2 = 1 - t_1$) ! Supposons maintenant la formule vérifiée pour n réels et ajoutons-en un $n+1$-ème x_{n+1}, ainsi que son coefficient t_{n+1}, toujours bien entendu en respectant la condition sur la somme. Si $t_{n+1} = 0$, il n'y rien à prouver puisqu'on est ramenés à l'hypothèse de récurrence. Sinon, on va poser $t'_n = t_n + t_{n+1}$ et $x'_n = \frac{t_n}{t'_n}x_n + \frac{t_{n+1}}{t'_n}x_{n+1}$. On peut appliquer l'hypothèse de récurrence aux réels $(x_1, \dots, x_{n-1}, x'_n)$ et aux réels $(t_1, \dots, t_{n-1}, t'_n)$ (dont la somme est bien égale à 1 par définition de $t'n$) pour obtenir $f\left(\sum_{k=1}^{n+1} t_k x_k\right) = f(t_1x_1 + \dots + t_{n-1}x_{n-1} + t'_n x'_n) \leqslant t_1 f(x_1) + \dots + t_{n-1}f(x_{n-1}) + t'_n f(x'_n)$. Or, $t'_n f(x'_n) = t'_n f\left(\frac{t_n}{t'_n}x_n + \frac{t_{n+1}}{t'_n}x_{n+1}\right) \leqslant t_n f(x_n) + t_{n+1}f(x_{n+1})$ (par inégalité « simple » de convexité), il n'y a plus qu'à remplacer dans l'inégalité précédente pour obtenir l'hérédité voulue. □

Théorème 9. Caractérisation de la convexité par croissance des pentes des sécantes.

Une fonction f est convexe sur l'intervalle I si et seulement si, pour tout réel $a \in I$, le taux d'accroissement $\tau : x \mapsto \frac{f(x)-f(a)}{x-a}$ est une fonction croissante sur $I \backslash \{a\}$.

Démonstration. Supposons donc f convexe sur I, et fixons $a \in I$. Soient x et y deux valeurs appartenant à I et distinctes de a telles que $x < y$. Si $y > a$, alors y est situé à l'extérieur du segment $[a,x]$ (ou $[x,a]$ selon la position de x), donc $f(y) \geqslant \dfrac{f(x)-f(a)}{x-a}(y-a)+f(a)$ (la courbe est au-dessus de sa sécante, cf remarque effectuée plus haut). Cela revient exactement à dire que $\dfrac{f(y)-f(a)}{y-a} \geqslant \dfrac{f(x)-f(a)}{x-a}$ (on divise par un nombre qui est positif par hypothèse). Si $y < a$, y est situé à l'intérieur du segment $[a,x]$, donc l'inégalité est dans l'autre sens (courbe en-dessous de ses cordes), mais revient dans le même sens après division par $y-a$ qui est négatif, pour la même conclusion. On a bien prouvé la croissance du taux d'accroissement en a pour une fonction convexe.

Supposons réciproquement les taux d'accroissements croissants. Si $(x,y) \in I^2$ avec $x < y$ et $t \in]0,1[$, on pose $a = tx+(1-t)y$, ce qui implique $x < a < y$. L'hypothèse faite permet d'affirmer que $\dfrac{f(x)-f(a)}{x-a} \leqslant \dfrac{f(y)-f(a)}{y-a}$, donc $(y-a)(f(x)-f(a)) \geqslant (x-a)(f(y)-f(a))$ (on multiplie par un nombre positif et par un nombre négatif, d'où le changement de sens de l'inégalité). Après développement et division par $y-x$, on a donc $f(a) \leqslant \dfrac{y-a}{y-x}f(x)+\dfrac{a-x}{y-x}f(y)$, c'est-à-dire exactement $f(tx+(1-t)y) \leqslant tf(x)+(1-t)f(y)$, ce qui est la définition même de la convexité ! □

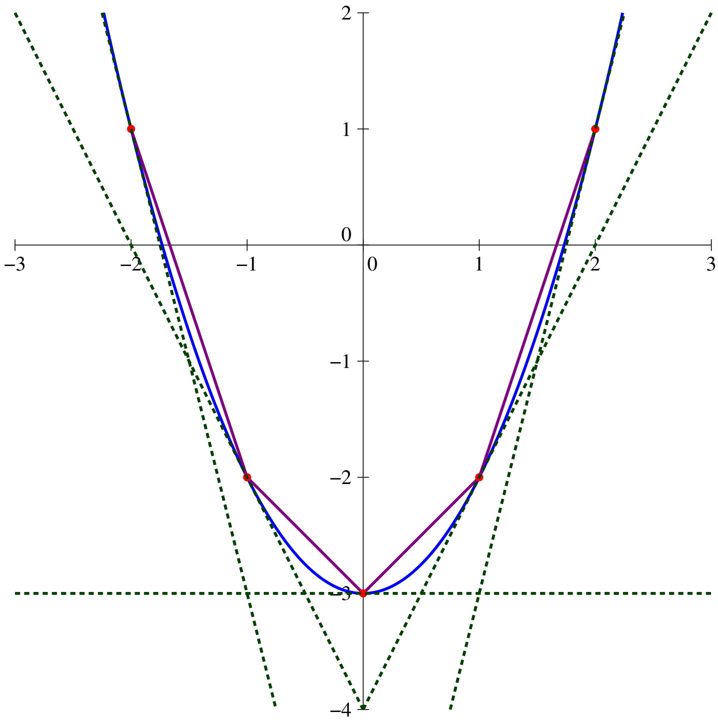
Théorème 10. Caractérisation de la convexité des fonctions dérivables.

Soit f une fonction dérivable sur un intervalle I, alors f est convexe sur I si et seulement si f' est croissante sur I (ou, dans le cas où f est deux fois dérivable, si et seulement si f'' est positive sur I). Dans ce cas, la courbe représentative de f sur l'intervalle I est située au-dessus de toutes ses tangentes.

Démonstration. Si f' est convexe, la croissance de f' découle de la caractérisation précédente sur la croissance des taux d'accroissement (il suffit de faire un passage à la limite). De plus, la tangente à $\mathcal{C}_f$ en son point d'abscisse a a pour équation $y = f'(a)(x-a)+f(a)$. Posons donc $g(x) = f(x)-f'(a)(x-a)-f(a)$, la fonction g est dérivable sur I et $g'(x) = f'(x)-f'(a)$. Par croissance de f', la fonction g est décroissante à gauche de a et croissante à droite de a. Comme $g(a) = 0$, cela suffit à prouver que g est toujours positive, exactement ce qu'on souhaitait démontrer.

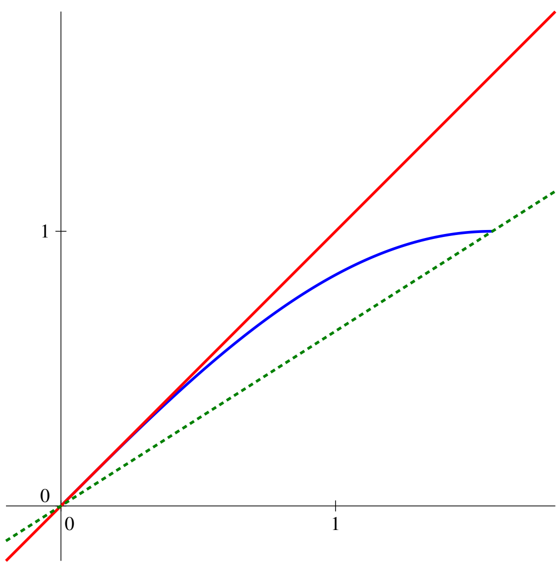
Supposons réciproquement la courbe située au-dessus de ses tangentes, et prouvons que f est convexe : on a par hypothèse $f(x) \geqslant f'(a)(x-a)+f(a)$, donc $tf(x)+(1-t)f(y) \geqslant tf'(a)(x-a)+tf(a)+(1-t)f'(a)(y-a)+(1-t)f(a) = f'(a)(tx+(1-t)y-a)+f(a)$. Il suffit de prendre $a = tx+(1-t)y$ pour retrouver la définition de la convexité. □

Ci-dessous, une illustration de la position relative d'une courbe de fonction convexe (en bleu) et de ses cordes d'une part (segments violets) et tangentes d'autres part (droites vertes en pointillés) :



Exemples : en pratique, on privilégiera bien sûr ces caractérisations faciles à démontrer pour prouver la convexité d'une fonction. Mais les autres caractérisations et définitions seront particulièrement intéressantes à exploiter pour obtenir des inégalités classiques, souvent connues sous le nom d'inégalités de convexité. Quelques exemples à connaître avec les fonctions usuelles :

- $\forall x \in \mathbb{R}$, $e^x \geqslant x + 1$ (la fonction exponentielle est convexe, donc au-dessus de sa tangente en 0).
- $\forall x \in]0, +\infty[$, $\ln(x) \leqslant x - 1$ (la fonction ln est concave, donc en-dessous de sa tangente en 1).
- $\forall x \in \left[0, \frac{\pi}{2}\right]$, $\frac{2x}{\pi} \leqslant \sin(x) \leqslant x$ (la fonction sin est concave sur cet intervalle, donc en-dessous de sa tangente en 0 et au-dessus de la corde reliant les deux points aux extrémités du segment). Une illustration de cet encadrement ci-dessous :



Définition 15. Soit f une fonction définie sur un intervalle I et $a \in I$, a est l'abscisse d'un **point**

d'inflexion de la courbe représentative de f si f change de concavité en a (convexe à gauche de a et concave à droite, ou le contraire).

Remarque 17. Si f est deux fois dérivable, on aura un point d'inflexion en a si f'' change de signe en a (et donc $f''(a) = 0$). La tangente à la courbe représentative de f en un point d'inflexion traverse localement cette courbe (la position relative de la courbe et de la tangente change au passage en a).