

Dérivation, accroissements finis et développements limités

Cours — L1 maths

1 1. Ce que ce chapitre ajoute au précédent

Le chapitre « Fonctions d’une variable réelle : limites et continuité » a laissé une lacune, et elle était volontaire : **les variations d’une fonction s’y démontrent à la main**. On prend deux points $x < y$ d’un intervalle, on calcule $f(y) - f(x)$, on factorise par $y - x$, on discute le signe du second facteur. Cela s’écrit proprement, et cela devient impraticable dès que la fonction cesse d’être un polynôme ou un quotient simple. Même gêne pour compter les solutions d’une équation $f(x) = \lambda$: le théorème de la bijection exigeait une **stricte monotonie** qu’il fallait établir sans autre outil que cette factorisation.

Ce chapitre remplace ce travail artisanal par un unique objet local, la **dérivée**, et il en tire quatre choses hors de portée du chapitre précédent : les **variations** en lisant un signe, sans jamais comparer deux valeurs ; la **tangente**, c’est-à-dire le remplacement d’une fonction compliquée par une fonction affine valable près d’un point ; les **accroissements finis**, qui transportent une information ponctuelle sur f' en une information globale sur f ; et l’**approximation polynomiale locale**, qui pousse la tangente aux ordres supérieurs et calcule des limites qu’aucune technique antérieure n’atteint.

1.1 Le plan du chapitre

Les vingt-cinq sections qui suivent se lisent en six blocs. **Dérivabilité et calcul des dérivées** (sections 2 à 9) : la définition en un point, la tangente, le lien avec la continuité, les dérivées latérales, les opérations, la dérivée d’une réciproque, le tableau des dérivées usuelles. **Extremums, Rolle et accroissements finis** (sections 10 à 14) : le lemme de Fermat, le théorème de Rolle, puis l’égalité des accroissements finis à partir de la section 12, l’inégalité qui en découle, les fonctions lipschitziennes et les suites récurrentes qu’elles contrôlent. **Variations et régularité de la dérivée** (sections 15 à 18) : le signe de f' et la monotonie, les accroissements finis généralisés et la règle de L’Hôpital, le théorème de la limite de la dérivée, le théorème de Darboux, les dérivées successives et les classes $\mathcal{C}^k$.

Viennent ensuite les **formules de Taylor**, celle de Taylor-Young qui est locale et celle de Taylor-Lagrange qui est globale, toutes deux démontrées ici sans aucun calcul d’aire ; les **développements limités**, définition, unicité, développements usuels en 0, opérations, puis les applications aux limites, aux équivalents, à la position d’une courbe et aux asymptotes ; enfin la **convexité** et la **méthode de Newton**, dont la vitesse de convergence se démontre avec exactement les outils du chapitre.

1.2 Ce que la dérivée ne donne pas

Cette puissance produit une illusion qu’il vaut mieux dissiper tout de suite. La dérivabilité **ne se déduit d’aucune régularité plus faible** : aucun argument de « courbe régulière » ne remplace le calcul d’une limite de taux d’accroissement, et une fonction continue n’est pas dérivable pour autant (section 4). Le signe de la dérivée **ne dit rien hors d’un intervalle** : la fonction inverse a une dérivée strictement négative en tout point de $\mathbb{R}^*$ sans être décroissante sur $\mathbb{R}^*$, et le mot *intervalle* est une hypothèse, jamais un décor. Une dérivée nulle **ne signale pas un extremum**, l’implication allant dans l’autre sens et seulement en un point intérieur (section 10). Enfin la dérivée d’une fonction dérivable **n’est pas forcément continue** : le théorème de Darboux dira ce qui survit malgré tout.

Remarque

Ce chapitre et les autres. Il n'existe pas de texte national fixant les contenus d'une licence de mathématiques : chaque université arrête les siens. Ce cours couvre le socle commun aux maquettes consultées et signale par un encadré *Selon les facs* ce qui varie d'un établissement à l'autre, le théorème de la limite de la dérivée, celui de Darboux, la formule de Leibniz ou la méthode de Newton notamment. Ces notions sont traitées, jamais supprimées.

Sont **employés** ici sans être refondés : les limites, la continuité, le théorème des valeurs intermédiaires, celui des bornes atteintes, le théorème de la bijection et la construction d'Arcsin, Arccos, Arctan, tous acquis du chapitre « Fonctions d'une variable réelle : limites et continuité » ; les suites, du chapitre « Suites réelles » ; les quantificateurs et les modes de raisonnement, du chapitre « Logique, raisonnement, ensembles et applications ». Tout ce qui touche aux primitives et au calcul d'aires relève du chapitre « Calcul intégral élémentaire » : **aucune démonstration de ce chapitre n'en a besoin**, y compris celle de Taylor-Lagrange, qui s'obtient par une fonction auxiliaire et le théorème de Rolle.

Un dernier mot, qui est celui du chapitre précédent. Ce qui coûte des points en première année n'est presque jamais le calcul : c'est l'hypothèse qu'on n'a pas vérifiée avant d'appliquer un théorème, l'intervalle qu'on a oublié de nommer, le contre-exemple qu'on annonce sans l'exhiber. Les encadrés **Comment on rédige** montrent la phrase exacte qu'un correcteur cherche des yeux, et ils sont à lire comme du cours.

2 2. Dérivabilité en un point

2.1 Le taux d'accroissement

Dans toute cette section, I désigne un intervalle de $\mathbb{R}$ non réduit à un point, f une fonction définie sur I et à valeurs réelles, et a un point de I . Cette hypothèse garantit l'essentiel : I contient des points distincts de a et aussi proches de a qu'on veut, donc la limite qui va suivre a un sens.

Définition

Le **taux d'accroissement** de f entre a et x est le quotient

$$\tau_a(x) = \frac{f(x) - f(a)}{x - a}, \quad x \in I \setminus \{a\}.$$

Géométriquement, $\tau_a(x)$ est la **pente de la sécante** qui joint les deux points de la courbe d'abscisses a et x . C'est un objet parfaitement défini pour tout $x \neq a$, et la seule question du chapitre est de savoir ce qu'il devient quand x se rapproche de a .

Définition

On dit que f est **dérivable en a** lorsque le taux d'accroissement τ_a admet en a une limite **finie**. Ce réel s'appelle alors le **nombre dérivé** de f en a , et se note $f'(a)$:

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

Le mot « finie » fait partie de la définition, ce n'est pas un ornement. Une limite peut exister dans $\mathbb{R}$ sans être un réel, et la fonction n'est alors pas dérivable. C'est le cas de la racine carrée en 0, dont le taux vaut $\tau_0(x) = \frac{\sqrt{x}}{x} = \frac{1}{\sqrt{x}}$ pour $x > 0$ et tend vers $+\infty$: la limite existe, elle est infinie, et $x \mapsto \sqrt{x}$ n'est **pas** dérivable en 0. La section 5 dira ce que cela signifie sur le dessin.

Lorsque a est une extrémité de I , la limite ne peut être prise que d'un seul côté : la dérivée obtenue est une **dérivée latérale**, et c'est encore la section 5 qui en traite.

2.2 Les trois écritures

Propriété

Trois formulations équivalentes. Avec les données ci-dessus, les trois énoncés suivants sont équivalents, et le réel α du troisième vaut alors $f'(a)$.

1. Le taux τ_a admet en a une limite finie, notée $f'(a)$.
2. Le quotient $\frac{f(a+h) - f(a)}{h}$ admet une limite finie quand h tend vers 0.
3. Il existe un réel α et une fonction ε , définie au voisinage de 0, de limite nulle en 0, telle que

$$f(a+h) = f(a) + \alpha h + h \varepsilon(h)$$

pour tout h tel que $a+h$ appartienne à I .

Démonstration. L'équivalence entre 1 et 2 est le changement de variable $x = a + h$: la condition $x \in I \setminus \{a\}$ devient $a + h \in I$ avec $h \neq 0$, et $x \rightarrow a$ devient $h \rightarrow 0$.

Supposons 2, et notons α la limite. Posons, pour $h \neq 0$ tel que $a + h \in I$,

$$\varepsilon(h) = \frac{f(a+h) - f(a)}{h} - \alpha, \quad \text{et} \quad \varepsilon(0) = 0.$$

Par construction $\varepsilon(h)$ tend vers 0 quand h tend vers 0, et en multipliant l'égalité par h on obtient $f(a+h) - f(a) = \alpha h + h \varepsilon(h)$, ce qui est 3. L'égalité vaut aussi pour $h = 0$, où les deux membres valent $f(a)$.

Réciproquement, supposons 3. Pour $h \neq 0$, on divise par h :

$$\frac{f(a+h) - f(a)}{h} = \alpha + \varepsilon(h).$$

Le second membre tend vers α quand h tend vers 0, donc le premier aussi : la limite existe, elle est finie, elle vaut α . L'unicité de la limite impose alors $\alpha = f'(a)$. $\square$

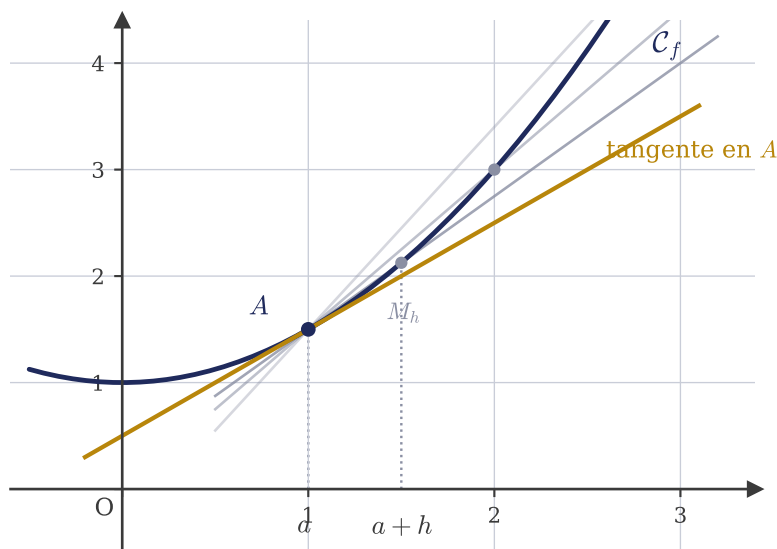
La troisième écriture est celle qui fera tout le travail du chapitre, et il faut voir pourquoi : elle **remplace un quotient par une somme**. Là où les deux premières exigent de diviser, donc de vérifier qu'on ne divise pas par zéro, la troisième est une égalité entre deux expressions toujours définies. La démonstration de la dérivée d'une composée, en section 7, échoue avec les deux premières et passe sans effort avec la troisième.

Remarque

Deux avertissements de notation. La fonction ε de l'écriture 3 est une **fonction de la variable h** , sans aucun rapport avec le réel $\varepsilon > 0$ des définitions de limite du chapitre précédent. La collision de symboles est traditionnelle et fâcheuse ; quand les deux apparaissent dans une même démonstration, il faut renommer.

Ensuite, l'écriture 3 est **déjà un développement limité à l'ordre 1** : elle dit que f coïncide, près de a , avec la fonction affine $h \mapsto f(a) + f'(a)h$, à un terme près négligeable devant h . Le bloc consacré aux développements limités notera ce terme $o(h)$ et poussera l'idée aux ordres supérieurs : tout y est déjà dans cette ligne.

2.3 Ce que la figure montre



La figure représente la courbe de $f : x \mapsto \frac{x^2}{2} + 1$, le point A d'abscisse $a = 1$, trois sécantes (AM_h) dont les points M_h d'abscisse $a + h$ se rapprochent de A , et la tangente en A : les pentes des sécantes décroissent et se stabilisent sur celle de la tangente.

Ce pivotement se calcule exactement, et c'est le meilleur exemple d'entrée dans le chapitre.

Exemple

Le taux d'accroissement de $x \mapsto \frac{x^2}{2} + 1$ en 1. On a $f(1) = \frac{3}{2}$. Pour $h \neq 0$,

$$f(1+h) - f(1) = \frac{(1+h)^2}{2} + 1 - \frac{3}{2} = \frac{1+2h+h^2-1}{2} = h + \frac{h^2}{2},$$

donc, en divisant par h ,

$$\frac{f(1+h) - f(1)}{h} = 1 + \frac{h}{2}.$$

Ce taux tend vers 1 quand h tend vers 0 : la fonction f est dérivable en 1 et $f'(1) = 1$. Les trois sécantes de la figure correspondent à $h = 1,8$, $h = 1$ et $h = 0,5$, de pentes respectives 1,9, 1,5 et 1,25 : elles décroissent bien vers 1.

L'écriture 3 se lit sur la même ligne : $f(1+h) = \frac{3}{2} + h + \frac{h^2}{2}$, c'est-à-dire $f(1+h) = f(1) + 1 \times h + h\varepsilon(h)$ avec $\varepsilon(h) = \frac{h}{2}$, qui tend bien vers 0. Ici la fonction ε est explicite ; c'est exceptionnel, et cela ne doit pas faire croire qu'on doit toujours la calculer.

Méthode

Comment on rédige un calcul de dérivée par le taux d'accroissement. Quatre étapes : annoncer la variable, « soit $h \neq 0$ tel que $a + h \in I$ », faute de quoi la suite manipule une expression qui n'est pas définie partout ; calculer le numérateur $f(a + h) - f(a)$ et le **factoriser par h** , ou le transformer par quantité conjuguée s'il contient une racine ; simplifier par h , ce qui est licite puisque $h \neq 0$ a été annoncé ; conclure par une phrase complète, « le taux tend vers ℓ quand h tend vers 0 ; cette limite est finie, donc f est dérivable en a et $f'(a) = \ell$ ».

Le geste à ne jamais omettre est la mention du caractère fini de la limite : c'est lui qui distingue la dérivabilité d'une tangente verticale, et un correcteur le cherche.

L'exercice 01 inverse le sens de la tâche : il donne six taux d'accroissement bruts, sans dire de quelle fonction ni en quel point, et c'est cela qu'il faut retrouver avant de conclure. Sa seconde question porte sur deux fonctions dont l'écriture laisse croire à un défaut de dérivabilité qui n'existe pas. L'exercice 02 travaille la lecture croisée de f , de f' et de f'' sur un graphique.

3 3. Tangente et approximation affine

Définition

Soit f dérivable en a . La **tangente** à la courbe $\mathcal{C}_f$ au point A de coordonnées $(a, f(a))$ est la droite d'équation

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

C'est la droite qui passe par A et dont la pente est $f'(a)$: elle est la position limite des sécantes (AM) quand M se rapproche de A sur la courbe, ce que la figure de la section précédente met en scène. Notons tout de suite que **cette définition suppose la dérivabilité** : sans elle, la courbe peut avoir une tangente, ou deux demi-tangentes, ou rien du tout, et la section 5 fait le tri.

3.1 Ce que la tangente approche

Notons T la fonction affine dont la tangente est le graphe, c'est-à-dire $T(x) = f(a) + f'(a)(x - a)$. L'écriture 3 de la section 2, appliquée avec $x = a + h$, donne exactement $f(x) - T(x) = (x - a)\varepsilon(x - a)$, avec ε de limite nulle. L'erreur commise en remplaçant f par T est donc non seulement petite, mais **petite devant $x - a$** : c'est bien plus fort que la continuité, qui dirait seulement que l'erreur tend vers 0.

Propriété

La tangente est la seule droite qui approche f à cet ordre. Soit m un réel. La quantité $f(x) - (f(a) + m(x - a))$ est négligeable devant $x - a$ au voisinage de a si et seulement si $m = f'(a)$.

Démonstration. Pour $x \neq a$, on divise par $x - a$ et il reste $\tau_a(x) - m$, qui tend vers $f'(a) - m$; ce quotient tend vers 0 si et seulement si $m = f'(a)$. $\square$

Toute autre droite passant par A commet donc une erreur **proportionnelle** à $x - a$, de coefficient $f'(a) - m$: elle s'écarte de la courbe linéairement, alors que la tangente s'en écarte moins vite que linéairement. C'est le sens précis de « la tangente est la droite qui épouse le mieux la courbe ».

3.2 L'approximation affine, sur un exemple chiffré

De l'écriture 3 découle l'**approximation affine** : pour h petit, $f(a + h) \approx f(a) + f'(a)h$. Le symbole $\approx$ n'est pas une égalité et n'a aucune valeur dans une démonstration ; c'est une indication de calcul

numérique. Ce que la théorie garantit, c'est que l'erreur est négligeable devant h , sans dire de combien. Les formules de Taylor remplaceront ce $\approx$ par une majoration explicite, et c'est exactement leur raison d'être.

Exemple

Une valeur approchée de $\sqrt{1,02}$, et l'erreur commise. Prenons $f = \sqrt{}$ sur $]0, +\infty[$ et $a = 1$. On admet, ou l'on relit dans le tableau de la section 9, que $f'(x) = \frac{1}{2\sqrt{x}}$, donc $f(1) = 1$ et $f'(1) = \frac{1}{2}$. L'approximation affine s'écrit

$$\sqrt{1+h} \approx 1 + \frac{h}{2}.$$

Pour $h = 0,02$, elle donne $\sqrt{1,02} \approx 1,01$. La valeur exacte vaut $1,00995\dots$, et l'erreur est d'environ $4,95 \times 10^{-5}$.

D'où vient cette erreur. Elle se calcule sans aucune formule de Taylor, par quantité conjuguée :

$$\sqrt{1+h} - \left(1 + \frac{h}{2}\right) = \frac{(1+h) - \left(1 + \frac{h}{2}\right)^2}{\sqrt{1+h} + 1 + \frac{h}{2}} = \frac{-\frac{h^2}{4}}{\sqrt{1+h} + 1 + \frac{h}{2}}.$$

Le numérateur est en h^2 , et c'est ce qui rend l'approximation bonne. Pour $h = 0,02$, il vaut -10^{-4} , le dénominateur vaut environ $2,02$, et l'erreur environ $-4,95 \times 10^{-5}$: le calcul numérique est confirmé, signe compris.

Deux enseignements. Ce signe négatif dit que la tangente passe **au-dessus** de la courbe, et ce n'est pas un accident : c'est la concavité de la racine carrée, que la section sur la convexité expliquera. Et l'ordre de grandeur en h^2 se vérifie : pour $h = 0,2$, dix fois plus grand, l'approximation donne $1,1$ contre une valeur exacte de $1,09544\dots$, soit une erreur d'environ $4,55 \times 10^{-3}$, cent fois plus grande.

La lecture graphique des tangentes, autre face de cette section, est travaillée dans l'exercice 02.

4 4. Dérivabilité entraîne continuité, et la réciproque est fausse

Théorème

Dérivable en a entraîne continue en a . Si f est dérivable en a , alors f est continue en a .

Démonstration. Pour tout $x \in I$ distinct de a , on peut écrire l'identité

$$f(x) - f(a) = (x - a) \times \tau_a(x),$$

qui n'est que la définition de τ_a multipliée par $x - a$. Faisons tendre x vers a . Le premier facteur tend vers 0. Le second tend vers $f'(a)$, qui est un **réel**, par hypothèse de dérivabilité. Le théorème des opérations sur les limites, acquis au chapitre précédent, donne que le produit tend vers $0 \times f'(a) = 0$. Ainsi $f(x) - f(a)$ tend vers 0, c'est-à-dire $f(x) \rightarrow f(a)$: la fonction f est continue en a . $\square$

Où le caractère fini a servi. Exactement à une ligne, et il faut la voir : c'est le produit $0 \times f'(a)$. Si le taux tendait vers $+\infty$, le produit serait une forme indéterminée et la conclusion tomberait. C'est pourquoi la racine carrée, dont le taux explose en 0, reste néanmoins continue en 0 : sa continuité se démontre autrement, ce théorème ne la donne pas.

4.1 Le contre-exemple de la valeur absolue

Exemple

La fonction $x \mapsto |x|$ est continue en 0 et n'y est pas dérivable.

Continuité. La seconde inégalité triangulaire donne $||x| - |0|| \leq |x - 0|$, ce qui montre que $|x| \rightarrow 0$ quand $x \rightarrow 0$. La fonction est continue en 0, comme le chapitre précédent l'a établi.

Non-dérivabilité. Le taux d'accroissement en 0 vaut, pour $h \neq 0$,

$$\tau_0(h) = \frac{|h| - 0}{h} = \frac{|h|}{h},$$

qui vaut 1 si $h > 0$ et -1 si $h < 0$. Sa limite à droite en 0 vaut donc 1, sa limite à gauche vaut -1 : les deux limites latérales existent et diffèrent, donc le taux n'a pas de limite en 0, en vertu du théorème sur les limites latérales du chapitre précédent. La fonction $x \mapsto |x|$ n'est pas dérivable en 0. $\square$

Ce que ce contre-exemple prouve, et ce qu'il ne prouve pas. Il prouve que l'implication du théorème n'est pas une équivalence : la continuité est **strictement plus faible** que la dérivabilité. Il ne prouve pas que les défauts de dérivabilité soient rares ou localisés : il existe des fonctions continues sur $\mathbb{R}$ et dérivables en **aucun** point. Leur construction relève d'un cours ultérieur ; il suffit de savoir qu'elles existent, pour ne jamais raisonner sur l'idée qu'une fonction continue serait « presque partout » dérivable.

Méthode

Comment on rédige le lien entre continuité et dérivabilité. L'implication ne va que dans un sens, d'où trois usages et un interdit.

1. **Sens direct, pour obtenir la continuité.** « f est dérivable en a , donc f est continue en a . » Raccourci légitime, qui évite un calcul de limite ; il sert dans la démonstration de la dérivée d'un produit, en section 7.
2. **Contraposée, pour réfuter la dérivabilité.** « f n'est pas continue en a , donc f n'est pas dérivable en a . » Argument imparable et gratuit : devant une fonction par morceaux qui saute en a , il n'y a **aucun** taux d'accroissement à calculer.
3. **Jamais dans l'autre sens.** « f est continue en a donc dérivable en a » est faux, et c'est la faute la plus fréquente du chapitre. Elle vient d'une image mentale, celle d'une courbe « sans trou donc sans angle », qui n'a aucune valeur démonstrative : $x \mapsto |x|$ suffit à la réfuter.

La formulation qui trahit l'erreur est celle qui affirme la dérivabilité sans jamais nommer de limite. Pour la démontrer, il faut exhiber la limite du taux d'accroissement, ou invoquer un théorème d'opérations en vérifiant ses hypothèses. Il n'y a pas de troisième voie.

L'exercice 34 met six copies sous les yeux du lecteur et demande, pour chacune, ce qui est démontré et ce qui est seulement affirmé ; deux d'entre elles reposent sur la confusion ci-dessus.

5 5. Dérivées à gauche et à droite

5.1 Les deux définitions

Définition

Soit a un point de I .

- Si I contient un intervalle de la forme $[a, a + r[$ avec $r > 0$, on dit que f est **dérivable à droite en a** lorsque τ_a admet en a une limite finie à droite. On la note $f'_d(a)$.
- Si I contient un intervalle de la forme $]a - r, a]$ avec $r > 0$, on dit que f est **dérivable à gauche en a** lorsque τ_a admet en a une limite finie à gauche. On la note $f'_g(a)$.

Ces deux nombres se calculent avec le même quotient que la dérivée, en restreignant les x à un seul côté. Toute la difficulté pratique est de rédaction : il faut écrire « pour $x > a$ » ou « pour $h < 0$ » **avant** le calcul, sous peine d'employer la mauvaise expression quand la fonction est définie par morceaux.

Théorème

Caractérisation par les dérivées latérales. Soit a un point **intérieur** à I . La fonction f est dérivable en a si et seulement si $f'_g(a)$ et $f'_d(a)$ existent toutes deux et sont **égales**. Dans ce cas,

$$f'(a) = f'_g(a) = f'_d(a).$$

Démonstration. Il suffit d'appliquer à la fonction τ_a , au point a , le théorème sur les limites latérales démontré au chapitre précédent. L'hypothèse « a intérieur à I » garantit que τ_a est définie de part et d'autre de a , ce que ce théorème exige. La limite de τ_a en a existe donc si et seulement si les deux limites latérales existent et coïncident, et le caractère fini se transmet de la même façon. $\square$

C'est le premier outil de **réfutation** de la dérivabilité, et le seul vraiment maniable au tableau : deux dérivées latérales qui diffèrent interdisent la dérivabilité, sans autre calcul.

5.2 Trois façons d'échouer

Point anguleux. Lorsque $f'_g(a)$ et $f'_d(a)$ existent, sont finies, et **différent**, la courbe présente en a un **point anguleux** : elle admet deux demi-tangentes, de pentes $f'_g(a)$ et $f'_d(a)$, qui forment un angle. La valeur absolue en 0 en est l'exemple minimal, avec les pentes -1 et 1 . Moins artificiel : $f : x \mapsto |x^2 - 1|$ au point 1. Pour $x > 1$ proche de 1, on a $f(x) = x^2 - 1$, donc $\tau_1(x) = x + 1 \rightarrow 2$; pour $x < 1$ proche de 1, on a $f(x) = 1 - x^2$, donc $\tau_1(x) = -(x + 1) \rightarrow -2$. Ainsi $f'_g(1) = -2$ et $f'_d(1) = 2$.

Tangente verticale. Lorsque le taux d'accroissement tend vers $+\infty$ ou vers $-\infty$, la fonction n'est **pas** dérivable en a , la limite n'étant pas finie. Pourtant la courbe possède bel et bien une tangente en ce point : la droite **verticale** d'équation $x = a$, position limite des sécantes dont la pente devient infinie. Les deux affirmations « f n'est pas dérivable en a » et « $\mathcal{C}_f$ admet une tangente en A » sont donc **simultanément vraies**, sans contradiction, parce que la seconde parle de la courbe et non de la fonction. C'est un piège de partiel classique.

Deux exemples valent d'être distingués.

- La racine carrée en 0 : le taux vaut $\frac{1}{\sqrt{x}}$ pour $x > 0$ et tend vers $+\infty$. Comme le domaine ne contient rien à gauche de 0, on parle d'une **demi-tangente verticale**.
- La **racine cubique**, réciproque de la bijection $x \mapsto x^3$ de $\mathbb{R}$ sur $\mathbb{R}$ construite au chapitre précédent, notée $\sqrt[3]{x}$. En 0, le taux d'accroissement vaut, pour $x \neq 0$,

$$\frac{\sqrt[3]{x}}{x} = \frac{\sqrt[3]{x}}{(\sqrt[3]{x})^3} = \frac{1}{(\sqrt[3]{x})^2},$$

qui tend vers $+\infty$ des deux côtés. La courbe a une **tangente verticale** en l'origine, qu'elle traverse.

Remarque

Une notation à ne pas confondre. On écrit $\sqrt[3]{x}$, et non $x^{1/3}$. La puissance d'exposant réel a été définie au chapitre précédent par $x^\alpha = e^{\alpha \ln x}$, donc **uniquement pour** $x > 0$: l'écriture $x^{1/3}$ n'a aucun sens pour x négatif, alors que la racine cubique est définie sur $\mathbb{R}$ tout entier. Employer une fonction là où on ne l'a pas définie est une faute qu'un correcteur relève sans hésiter, et elle se glisse d'autant mieux que les deux notations coïncident sur $]0, +\infty[$. Une convention étendue existe cependant, sous laquelle $x^{1/3}$ désigne la racine cubique sur $\mathbb{R}$; elle est licite à condition d'être annoncée, et la section 9 dit exactement à quelles conditions.

Oscillation. Troisième mode d'échec, le plus radical : le taux d'accroissement n'a **aucune** limite, ni finie ni infinie. C'est le cas de la fonction définie par $f(x) = x \sin \frac{1}{x}$ pour $x \neq 0$ et $f(0) = 0$, dont le taux en 0 vaut $\sin \frac{1}{h}$ et oscille indéfiniment entre -1 et 1 . Aucune tangente, aucune demi-tangente, et pourtant f est continue en 0, puisque $|f(x)| \leq |x|$.

L'exercice 04 travaille ce troisième mode, l'oscillation, sur la famille $x^\alpha \sin \frac{1}{x}$: chacune de ses questions a pour réponse un seuil sur α , et les seuils sont différents.

5.3 Étudier une fonction définie par morceaux

Méthode

Comment on rédige l'étude d'une fonction définie par morceaux. Le plan est toujours le même, et l'ordre des étapes n'est pas négociable.

1. **Les points où rien ne se passe.** Sur chaque morceau **ouvert**, la fonction coïncide avec une expression usuelle : une phrase suffit, et elle nomme la structure. « Sur $] -\infty, 1[$, f coïncide avec une fonction polynomiale, donc f y est continue et dérivable. » On ne calcule aucune limite là.
2. **La continuité au point de raccord, d'abord.** Pour chaque point a de raccord, trois quantités : la valeur $f(a)$, lue sur la ligne de la définition qui contient le cas $x = a$; la limite à gauche $f(a^-)$, calculée avec l'expression valable pour $x < a$; la limite à droite $f(a^+)$, avec celle valable pour $x > a$.
3. **Si f n'est pas continue en a , on s'arrête** : la contraposée de la section 4 conclut, f n'est pas dérivable en a , et aucun taux d'accroissement n'est à écrire.
4. **La dérivabilité ensuite, et seulement ensuite.** On calcule les deux taux latéraux, chacun avec l'expression de son côté et avec la valeur $f(a)$ qu'on vient d'établir, puis on compare $f'_g(a)$ et $f'_d(a)$.
5. **La conclusion, en toutes lettres.** « Les deux dérivées latérales existent, sont finies et égales à ℓ , donc f est dérivable en a et $f'(a) = \ell$ », ou bien « elles diffèrent, donc f n'est pas dérivable en a : la courbe y présente un point anguleux ».

Attention à ce qu'on n'a PAS le droit de faire. Dériver chaque morceau, calculer $\lim_{x \rightarrow a^-} f'(x)$ et $\lim_{x \rightarrow a^+} f'(x)$, les comparer, et conclure. Ce raisonnement **ressemble** au précédent sans être le même : il porte sur les limites de la fonction dérivée, pas sur les dérivées latérales de f en a . Un théorème permet parfois de passer des unes aux autres, le **théorème de la limite de la dérivée**, mais il arrive en section 17, il a des hypothèses précises, et il ne s'applique pas toujours : il existe des fonctions dérivables en a dont la dérivée n'a aucune limite en a . Tant qu'on ne l'a pas énoncé et vérifié, la seule voie sûre est le taux d'accroissement.

Exemple

Un recollement à deux paramètres. Soient α et β deux réels, et f définie sur $\mathbb{R}$ par $f(x) = x^2$ si $x \leq 1$ et $f(x) = \alpha x + \beta$ si $x > 1$. Pour quelles valeurs de α et β la fonction f est-elle dérivable sur $\mathbb{R}$?

Ailleurs qu'en 1. Sur $] -\infty, 1[$ et sur $]1, +\infty[$, la fonction coïncide avec une fonction polynomiale : elle y est dérivable, quels que soient α et β .

Continuité en 1. La valeur est donnée par la première ligne, qui contient le cas $x = 1$: $f(1) = 1$. On a $f(1^-) = 1$ et $f(1^+) = \alpha + \beta$. La continuité en 1 équivaut donc à $\alpha + \beta = 1$; sans cette égalité, f n'est pas continue en 1, donc pas dérivable, et l'étude s'arrête.

Dérivabilité en 1. Supposons désormais $\beta = 1 - \alpha$. Pour $x < 1$, le taux vaut $\tau_1(x) = \frac{x^2 - 1}{x - 1} = x + 1$, qui tend vers 2 : ainsi $f'_g(1) = 2$. Pour $x > 1$, le numérateur devient $\alpha x + 1 - \alpha - 1 = \alpha(x - 1)$, donc $\tau_1(x) = \alpha$ et $f'_d(1) = \alpha$.

Conclusion. La fonction f est dérivable en 1 si et seulement si $\alpha = 2$, et alors $\beta = -1$. Le couple $(2, -1)$ est le seul qui convienne ; pour tout autre couple vérifiant $\alpha + \beta = 1$, la courbe présente en 1 un point anguleux de pentes 2 et α .

Notons où la continuité a servi : sans elle, le numérateur $\alpha x + \beta - 1$ ne tend pas vers 0, le quotient tend vers l'infini, et le taux ne dit plus rien d'utile. Ce n'est pas une commodité de plan, c'est une nécessité de calcul.

L'exercice 03 recolle **trois** morceaux en **deux** points de jonction, ce qui change la nature de la question : il faut compter les degrés de liberté contre les contraintes avant de calculer. Et l'exercice 20 pousse la même question un cran plus loin en demandant la **classe** du raccord obtenu, notion introduite en section 18.

6 6. Dérivabilité sur un intervalle, fonction dérivée

Définition

Soit J un intervalle contenu dans I .

- Si J est **ouvert**, on dit que f est **dérivable sur J** lorsqu'elle est dérivable en chaque point de J .
- Si $J = [a, b]$ est un **segment**, on dit que f est dérivable sur $[a, b]$ lorsqu'elle est dérivable en chaque point de $]a, b[$, dérivable à droite en a et dérivable à gauche en b . On convient alors de noter $f'(a) = f'_d(a)$ et $f'(b) = f'_g(b)$.

Dans les deux cas, l'application qui à chaque point associe le nombre dérivé s'appelle la **fonction dérivée** de f , et se note f' .

La convention sur les extrémités est commode, et il faut savoir qu'elle en est une : $f'(a)$ y désigne une dérivée à droite, pas la limite d'un objet bilatéral qui n'existe pas.

6.1 Les notations

La notation de Lagrange, f' , désigne la fonction dérivée et $f'(a)$ sa valeur en a : c'est celle de ce cours. La notation de Leibniz, $\frac{df}{dx}$, insiste sur la variable par rapport à laquelle on dérive, ce qui devient indispensable dès qu'il y en a plusieurs ; la valeur en a s'écrit $\frac{df}{dx}(a)$. Enfin Df apparaît en algèbre linéaire, quand on voit la dérivation comme une application entre espaces de fonctions.

Attention : $\frac{df}{dx}$ n'est pas un quotient de deux nombres, mais un symbole d'un seul tenant, hérité d'une notation où df et dx désignaient des accroissements infiniment petits. Les manipulations

du type « on multiplie par dx » se justifient autrement, et n'ont pas leur place dans une démonstration de première année.

6.2 L'ensemble de dérivabilité

Définition

L'**ensemble de dérivabilité** de f est l'ensemble des points de son domaine où f est dérivable. C'est le domaine de la fonction dérivée f' .

Cet ensemble **se détermine, il n'est pas donné**, et il est en général strictement plus petit que le domaine de f . La valeur absolue est définie sur $\mathbb{R}$ et dérivable sur $\mathbb{R}^*$ seulement ; la racine carrée est définie sur $[0, +\infty[$ et dérivable sur $]0, +\infty[$ seulement ; la racine cubique est définie sur $\mathbb{R}$ et dérivable sur $\mathbb{R}^*$. Un énoncé qui commence par « soit f dérivable sur I » a donc déjà fait ce travail, et un énoncé qui n'en dit rien le laisse à faire.

Trois questions distinctes se posent ainsi devant une fonction, et il faut les garder séparées :

- a. Où f est-elle définie ? b. Où f est-elle dérivable ? c. La fonction f' est-elle continue ?

La troisième est la plus surprenante, et sa réponse est **non en général** : une fonction peut être dérivable en tout point d'un intervalle sans que sa dérivée y soit continue. Une fonction dérivable dont la dérivée est de plus continue est dite de **classe C^1** , et cette distinction structure le bloc des sections 15 à 18. La dérivée n'est pourtant pas libre de tout : le théorème de Darboux montrera qu'elle vérifie la propriété des valeurs intermédiaires, continue ou non.

Enfin, rien n'interdit de recommencer : si f' est elle-même dérivable, sa dérivée se note f'' , puis $f^{(3)}$. Ces **dérivées successives**, objet de la section 18, sont le préalable indispensable aux formules de Taylor. L'exercice 04 fait déjà déterminer trois ensembles de dérivabilité qui diffèrent du domaine.

7. Opérations sur les dérivées

Toutes les dérivées calculées en pratique le sont par ce théorème, et jamais par un taux d'accroissement. Encore faut-il pouvoir citer ses hypothèses.

Théorème

Opérations algébriques. Soient f et g deux fonctions définies sur I , dérivables en a , et λ un réel. Alors $f + g$, λf et fg sont dérivables en a , avec

$$(f + g)'(a) = f'(a) + g'(a), \quad (\lambda f)'(a) = \lambda f'(a),$$

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

Si de plus $g(a) \neq 0$, alors $\frac{1}{g}$ et $\frac{f}{g}$ sont dérivables en a , avec

$$\left(\frac{1}{g}\right)'(a) = -\frac{g'(a)}{g(a)^2}, \quad \left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

7.1 La somme, en une ligne

Pour $x \neq a$, le taux d'accroissement de $f + g$ est la somme des taux de f et de g , puisque le quotient est linéaire en son numérateur :

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

où l'exposant indique de quelle fonction le taux est celui. Le théorème des opérations sur les limites conclut, et le cas de λf est identique.

7.2 Le produit, démontré

Démonstration de la formule du produit. Soit $x \in I$ avec $x \neq a$. Le numérateur $f(x)g(x) - f(a)g(a)$ ne se factorise pas tel quel : on lui ajoute et on lui retranche le terme croisé $f(a)g(x)$, ce qui est le seul geste de la démonstration. Il vient

$$f(x)g(x) - f(a)g(a) = (f(x) - f(a))g(x) + f(a)(g(x) - g(a)).$$

En divisant par $x - a$, qui est non nul, on obtient

$$\frac{(fg)(x) - (fg)(a)}{x-a} = \tau_a^f(x)g(x) + f(a)\tau_a^g(x).$$

Faisons tendre x vers a et examinons les trois facteurs variables. Le taux $\tau_a^f(x)$ tend vers $f'(a)$ et le taux $\tau_a^g(x)$ tend vers $g'(a)$, par hypothèse. Quant à $g(x)$, il tend vers $g(a)$: c'est ici, et seulement ici, qu'intervient le théorème de la section 4, puisque g est dérivable en a , donc continue en a . Le théorème des opérations sur les limites donne alors

$$\lim_{x \rightarrow a} \frac{(fg)(x) - (fg)(a)}{x-a} = f'(a)g(a) + f(a)g'(a),$$

limite finie : la fonction fg est dérivable en a , de nombre dérivé annoncé. $\square$

Cette démonstration explique l'ordre des sections : « dérivable entraîne continue » devait précéder les opérations, sans quoi le facteur $g(x)$ resterait sans limite. La dissymétrie du calcul, où l'on a gardé $g(x)$ et $f(a)$, ne se retrouve d'ailleurs pas dans le résultat, symétrique en f et g ; le choix inverse aurait donné la même formule.

7.3 Le quotient, et son hypothèse cachée

Il suffit de traiter $\frac{1}{g}$, puisque $\frac{f}{g} = f \times \frac{1}{g}$ et que le produit est démontré. Or l'écriture même de $\frac{1}{g(x)}$ suppose $g(x) \neq 0$, alors que l'hypothèse ne porte que sur le point a . C'est le théorème de **conservation locale du signe** du chapitre précédent qui comble le trou : g est continue en a et $g(a) \neq 0$, donc g garde le signe de $g(a)$, et en particulier ne s'annule pas, sur tout un voisinage de a . Le taux d'accroissement s'y calcule alors :

$$\frac{1}{x-a} \left(\frac{1}{g(x)} - \frac{1}{g(a)} \right) = \frac{g(a) - g(x)}{(x-a)g(x)g(a)} = -\frac{1}{g(x)g(a)} \times \tau_a^g(x).$$

Le premier facteur tend vers $-\frac{1}{g(a)^2}$ par continuité de g en a , le second vers $g'(a)$, d'où la formule. Celle du quotient s'obtient ensuite en développant le produit et en réduisant au même dénominateur.

7.4 La composée, et pourquoi la preuve naïve échoue

Théorème

Dérivée d'une composée. Soient f définie sur I et dérivable en a , et g définie sur un intervalle J contenant $f(I)$ et dérivable au point $b = f(a)$. Alors $g \circ f$ est dérivable en a et

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

C'est la formule que tout le monde récite et que presque personne ne sait démontrer, parce que la démonstration « évidente » est fautive. Elle consisterait à écrire, pour $x \neq a$,

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

et à passer à la limite. Le problème saute aux yeux : **rien ne garantit que** $f(x) \neq f(a)$, et l'on peut avoir divisé par zéro. L'objection n'est pas théorique. Si f est constante, le quotient n'a jamais de sens ; et il existe des fonctions non constantes, dérivables en a , qui prennent la valeur $f(a)$ en des points arbitrairement proches de a . Ainsi la fonction φ définie par $\varphi(x) = x^2 \sin \frac{1}{x}$ pour $x \neq 0$ et $\varphi(0) = 0$ est dérivable en 0, son taux d'accroissement valant $x \sin \frac{1}{x}$, majoré en valeur absolue par $|x|$; et elle s'annule en tous les points $\frac{1}{n\pi}$ pour n entier non nul, aussi proches de 0 qu'on veut.

On passe donc par l'écriture 3 de la section 2, **qui ne divise jamais**. C'est la voie propre, et c'est celle-ci qu'il faut savoir refaire.

Démonstration. Posons $b = f(a)$. La dérivabilité de f en a fournit une fonction ε_1 , de limite nulle en 0, telle que

$$f(a + h) = f(a) + f'(a)h + h\varepsilon_1(h).$$

Notons $w(h) = f(a + h) - b$ l'accroissement de f , de sorte que

$$w(h) = h(f'(a) + \varepsilon_1(h)).$$

Cette écriture donne deux informations : $w(h)$ tend vers 0 quand h tend vers 0, et le quotient $\frac{w(h)}{h}$ tend vers $f'(a)$.

La dérivabilité de g en b fournit de même une fonction ε_2 , de limite nulle en 0, telle que $g(b + k) = g(b) + g'(b)k + k\varepsilon_2(k)$ pour tout k tel que $b + k \in J$. **Convenons de poser** $\varepsilon_2(0) = 0$: l'égalité reste vraie pour $k = 0$, ses deux membres valant $g(b)$, et la fonction ε_2 devient ainsi **continue** en 0, ce qui va être décisif.

Appliquons cette égalité avec $k = w(h)$, ce qui est licite puisque $b + w(h) = f(a + h)$ appartient bien à J :

$$g(f(a + h)) - g(f(a)) = g'(b)w(h) + w(h)\varepsilon_2(w(h)).$$

Divisons par $h \neq 0$:

$$\frac{g(f(a + h)) - g(f(a))}{h} = g'(b) \frac{w(h)}{h} + \frac{w(h)}{h} \varepsilon_2(w(h)).$$

Le premier terme tend vers $g'(b)f'(a)$. Dans le second, le facteur $\frac{w(h)}{h}$ tend vers $f'(a)$, donc reste borné au voisinage de 0 ; et $\varepsilon_2(w(h))$ tend vers 0, car $w(h)$ tend vers 0 et ε_2 est **continue** en 0 avec $\varepsilon_2(0) = 0$. Le second terme tend donc vers 0, et le taux d'accroissement de $g \circ f$ en a tend vers $g'(b)f'(a)$, limite finie. $\square$

Remarque

Où l'on retrouve le piège de la composition des limites. La convention $\varepsilon_2(0) = 0$ n'est pas une commodité d'écriture : sans elle, on affirmerait que $\varepsilon_2(w(h))$ tend vers 0 en composant deux limites, alors que w peut parfaitement prendre la valeur 0 au voisinage de 0. C'est exactement la situation que le théorème de composition des limites du chapitre précédent interdit de traiter à la légère. En rendant ε_2 continue en 0, on se place sous sa seconde hypothèse, celle qui autorise la conclusion sans rien savoir de la fonction intérieure : une ligne de convention remplace une hypothèse invérifiable.

7.5 Les corollaires du calcul courant

Soit u une fonction dérivable sur un intervalle. Les formules suivantes sont la composée appliquée aux fonctions usuelles, et il faut les lire avec leur condition de validité.

- a. $(u^n)' = n u^{n-1} u'$ pour $n \in \mathbb{Z}$, avec u ne s'annulant pas si $n < 0$ d. $(\ln u)' = \frac{u'}{u}$, sous réserve que $u > 0$
- b. $(u^\alpha)' = \alpha u^{\alpha-1} u'$ pour α réel, sous réserve que $u > 0$ e. $(\sqrt{u})' = \frac{u'}{2\sqrt{u}}$, sous réserve que $u > 0$
- c. $(e^u)' = u' e^u$ f. $(\sin u)' = u' \cos u$ et $(\cos u)' = -u' \sin u$

Méthode

Comment on rédige un calcul de dérivée composée. Trois phrases, dans cet ordre.

1. **Nommer le domaine de dérivabilité, avant tout calcul.** « La fonction $\psi : x \mapsto \ln(1+x^2)$ est définie et dérivable sur $\mathbb{R}$ comme composée de $x \mapsto 1+x^2$, dérivable sur $\mathbb{R}$ et à valeurs dans $]0, +\infty[$, et de $\ln$, dérivable sur $]0, +\infty[$. » C'est la mention de l'ensemble d'arrivée de la fonction intérieure qui justifie la composition.
2. **Appliquer la formule**, dérivée extérieure évaluée en la fonction intérieure, multipliée par la dérivée intérieure : $\psi'(x) = \frac{2x}{1+x^2}$.
3. **Simplifier, et donner le résultat sur le domaine annoncé.**

Les trois erreurs qui reviennent. Oublier la dérivée intérieure : la dérivée de $x \mapsto \sin(2x)$ est $x \mapsto 2 \cos(2x)$, et non $x \mapsto \cos(2x)$. Inverser l'ordre au numérateur du quotient : c'est $f'g - fg'$, et le contrôle se fait en dérivant $\frac{1}{x}$, dont la dérivée doit sortir négative. Appliquer la formule b ou d là où la fonction intérieure change de signe : une dérivée calculée hors du domaine où l'expression a un sens est fausse, même si le calcul formel « marche ».

L'exercice 05 propose dix dérivées à calculer et signale, pour chacune, l'erreur qui lui est associée.

8 8. Dérivée d'une fonction réciproque

Le chapitre précédent a construit les fonctions réciproques par le théorème de la bijection, et démontré que la réciproque d'une fonction continue et strictement monotone sur un intervalle est continue et strictement monotone de même sens. Il ne restait qu'une question, celle de la dérivabilité, et elle se règle ici.

Théorème

Dérivée d'une bijection réciproque. Soit f continue et strictement monotone sur un intervalle I , donc bijective de I sur l'intervalle $J = f(I)$. Soient $a \in I$ et $b = f(a)$. On suppose que f est dérivable en a et que $f'(a) \neq 0$. Alors f^{-1} est dérivable en b et

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

La notation f^{-1} désigne ici la **bijection réciproque**, l'existence de la bijection étant une hypothèse ; elle ne se confond pas avec la notation d'image réciproque du chapitre « Logique, raisonnement, ensembles et applications », qui ne suppose rien.

Démonstration. Soit $y \in J$ avec $y \neq b$, et posons $x = f^{-1}(y)$. Comme f^{-1} est injective, $y \neq b$ entraîne $x \neq a$. On peut donc écrire

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

le dernier quotient ayant un sens puisque $f(x) \neq f(a)$, encore par injectivité.

Faisons maintenant tendre y vers b . La fonction f^{-1} est continue en b , d'après le théorème de la bijection du chapitre précédent, donc $x = f^{-1}(y)$ tend vers a . De plus $x \neq a$ dès que $y \neq b$, ce qui est précisément la première hypothèse du théorème de composition des limites : on a donc le droit de composer, et $\tau_a(x)$ tend vers $f'(a)$.

Comme $f'(a) \neq 0$, le quotient $\frac{1}{\tau_a(x)}$ tend vers $\frac{1}{f'(a)}$, qui est un réel. Le taux d'accroissement de f^{-1} en b admet donc une limite finie : f^{-1} est dérivable en b , de nombre dérivé $\frac{1}{f'(a)}$. $\square$

Propriété

Version globale. Si f est continue et strictement monotone sur I , dérivable sur I , et si f' ne s'annule en **aucun** point de I , alors f^{-1} est dérivable sur $J = f(I)$ et

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

C'est le théorème appliqué en chaque point. La formule est à lire attentivement : le dénominateur n'est pas f' , mais f' **composée avec** f^{-1} , ce qui est indispensable pour que les deux membres soient des fonctions de la même variable, celle qui parcourt J .

8.1 L'interprétation géométrique

Les courbes de f et de f^{-1} sont symétriques l'une de l'autre par rapport à la droite d'équation $y = x$: au point A de coordonnées (a, b) de $\mathcal{C}_f$ correspond le point A' de coordonnées (b, a) de $\mathcal{C}_{f^{-1}}$. Cette symétrie échange les coordonnées, donc elle transforme une droite de pente m non nulle en une droite de pente $\frac{1}{m}$: la formule du théorème en est la traduction analytique.

Elle explique du même coup l'hypothèse $f'(a) \neq 0$. Si $f'(a) = 0$, la tangente en A est **horizontale**, et son image par la symétrie est **verticale** : la courbe de f^{-1} a bien une tangente en A' , mais f^{-1} n'est pas dérivable en b , et l'on retrouve la situation de la section 5, le taux d'accroissement étant l'inverse d'une quantité qui tend vers 0. C'est exactement ce qui se passe pour $x \mapsto x^3$, dont la dérivée $x \mapsto 3x^2$

ne s'annule qu'en 0 : sa réciproque, la racine cubique, est dérivable sur $\mathbb{R}^*$ avec

$$(\sqrt[3]{})'(y) = \frac{1}{3(\sqrt[3]{y})^2},$$

et ne l'est pas en 0, où sa courbe présente une tangente verticale. Le calcul direct de la section 5 le montrait déjà ; le théorème l'explique.

Méthode

Comment on rédige la dérivée d'une réciproque. Quatre étapes, dont la deuxième est celle qu'on oublie.

1. **Établir la bijection**, en citant le théorème de la bijection avec ses trois hypothèses, intervalle, continuité, stricte monotonie, et nommer l'intervalle image J .
2. **Dériver f et déterminer où f' s'annule**. Ces points sont à exclure, et il faut dire lesquels.
3. **Écrire la formule** sous la forme $(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}$, en gardant $f^{-1}(y)$ tel quel : à ce stade on n'a pas d'expression explicite.
4. **Éliminer f^{-1} de l'écriture** au moyen d'une identité : c'est le seul vrai calcul. Pour les fonctions circulaires réciproques, cette identité est trigonométrique et **elle demande une justification de signe**, faute de quoi une racine carrée apparaît sans qu'on sache si elle est affectée du bon signe.

8.2 Les dérivées d'Arcsin, Arccos et Arctan

Ces trois fonctions ont été **construites** au chapitre précédent par le théorème de la bijection ; on ne les reconstruit pas, on les dérive.

Arcsinus. La restriction de $\sin$ à $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ est continue, strictement croissante, et réalise une bijection sur $[-1, 1]$ dont la réciproque est Arcsin. La dérivée de cette restriction est $\cos$, qui s'annule exactement aux deux extrémités $-\frac{\pi}{2}$ et $\frac{\pi}{2}$, images réciproques de -1 et 1 . Le théorème s'applique donc en tout point de l'**ouvert** $] -1, 1[$, et donne

$$\text{Arcsin}'(x) = \frac{1}{\cos(\text{Arcsin } x)}.$$

Reste à simplifier le dénominateur, et c'est là qu'un signe se joue. Posons $\theta = \text{Arcsin } x$, de sorte que $\sin \theta = x$ et, par construction de la fonction, $\theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. La relation fondamentale donne

$$\cos^2 \theta = 1 - \sin^2 \theta = 1 - x^2, \quad \text{donc} \quad |\cos \theta| = \sqrt{1 - x^2}.$$

C'est une valeur absolue, pas encore un cosinus. Pour la lever, on utilise l'intervalle où vit θ : sur $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, le cosinus est positif ou nul, et même strictement positif à l'intérieur. Comme $x \in] -1, 1[$, on a $\theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$, donc $\cos \theta > 0$ et $\cos \theta = +\sqrt{1 - x^2}$. D'où

$$\forall x \in] -1, 1[, \quad \text{Arcsin}'(x) = \frac{1}{\sqrt{1 - x^2}}.$$

Cette expression tend vers $+\infty$ quand x tend vers 1 ou vers -1 : la courbe d'Arcsin présente deux tangentes verticales aux extrémités de son domaine, reflet par la symétrie d'axe $y = x$ des deux tangentes horizontales du sinus en $\pm \frac{\pi}{2}$.

Arccosinus. Même travail avec la restriction de $\cos$ à $[0, \pi]$, strictement décroissante, dont la dérivée $-\sin$ s'annule exactement en 0 et en π , images réciproques de 1 et -1 . L'étape clé est encore le signe : en posant $\theta = \operatorname{Arccos} x \in]0, \pi[$ pour $x \in]-1, 1[$, on a $\sin \theta > 0$, donc $\sin \theta = +\sqrt{1-x^2}$, et

$$\forall x \in]-1, 1[, \quad \operatorname{Arccos}'(x) = \frac{1}{-\sin(\operatorname{Arccos} x)} = \frac{-1}{\sqrt{1-x^2}}.$$

Le signe moins vient de la dérivée du cosinus, et il était prévisible : Arccos est strictement décroissante.

Remarque

Une vérification qui ne conclut pas encore. Les deux résultats donnent $\operatorname{Arcsin}'(x) + \operatorname{Arccos}'(x) = 0$ sur $] -1, 1[$, ce qui s'accorde avec l'identité $\operatorname{Arcsin} x + \operatorname{Arccos} x = \frac{\pi}{2}$ du chapitre précédent. Attention cependant : passer d'une dérivée nulle à une fonction constante demande un théorème, celui qui découle de l'égalité des accroissements finis et exige un **intervalle**. Il est démontré plus loin ; tant qu'on ne l'a pas, la vérification est un contrôle de vraisemblance, pas une démonstration.

Arctangente. La restriction de $\tan$ à $]-\frac{\pi}{2}, \frac{\pi}{2}[$ est continue, strictement croissante, de limites $-\infty$ et $+\infty$ aux bornes, donc bijective sur $\mathbb{R}$. Sa dérivée se calcule par la formule du quotient de la section 7 :

$$\tan' = \left(\frac{\sin}{\cos} \right)' = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2} = 1 + \tan^2.$$

Cette dérivée ne s'annule **jamais**, puisqu'elle vaut au moins 1 : le théorème s'applique en tout point de $\mathbb{R}$, sans exclusion. L'étape clé est alors immédiate, car $\tan(\operatorname{Arctan} x) = x$ pour tout réel x :

$$\forall x \in \mathbb{R}, \quad \operatorname{Arctan}'(x) = \frac{1}{1 + \tan^2(\operatorname{Arctan} x)} = \frac{1}{1 + x^2}.$$

Aucune justification de signe n'est ici nécessaire, et la raison mérite d'être dite : on n'a extrait **aucune racine carrée**. Les discussions de signe des deux cas précédents venaient uniquement de là.

L'exercice 06 fait travailler la formule dans ses trois usages, et l'exercice 07 pousse le procédé jusqu'à une fonction dont la réciproque n'a aucune expression avec les fonctions usuelles, ce qui est le cas le plus fréquent en pratique et le meilleur argument en faveur du théorème.

9 9. Tableau des dérivées usuelles

Le tableau ci-dessous rassemble les dérivées à connaître. La troisième colonne fait partie de l'énoncé, au même titre que la deuxième.

Fonction	Dérivée	Ensemble de dérivabilité
$x \mapsto x^n, n \in \mathbb{N}^*$	nx^{n-1}	$\mathbb{R}$
$x \mapsto x^{-n}, n \in \mathbb{N}^*$	$-nx^{-n-1}$	$\mathbb{R}^*$
$x \mapsto x^\alpha, \alpha \in \mathbb{R}$	$\alpha x^{\alpha-1}$	$]0, +\infty[$
$x \mapsto \sqrt{x}$	$\frac{1}{2\sqrt{x}}$	$]0, +\infty[$
exp	exp	$\mathbb{R}$
ln	$x \mapsto \frac{1}{x}$	$]0, +\infty[$
sin	cos	$\mathbb{R}$
cos	$-\sin$	$\mathbb{R}$
tan	$1 + \tan^2 = \frac{1}{\cos^2}$	$\mathbb{R}$ privé des $\frac{\pi}{2} + k\pi$
$x \mapsto x $	$x \mapsto \frac{x}{ x }$	$\mathbb{R}^*$
Arcsin	$x \mapsto \frac{1}{\sqrt{1-x^2}}$	$] -1, 1[$
Arccos	$x \mapsto \frac{-1}{\sqrt{1-x^2}}$	$] -1, 1[$
Arctan	$x \mapsto \frac{1}{1+x^2}$	$\mathbb{R}$
ch	sh	$\mathbb{R}$
sh	ch	$\mathbb{R}$
th	$1 - \text{th}^2 = \frac{1}{\text{ch}^2}$	$\mathbb{R}$

Trois commentaires, et ils portent tous sur la troisième colonne.

Les fonctions hyperboliques ne demandent rien de neuf. Elles ont été définies au chapitre précédent par $\text{ch } x = \frac{e^x + e^{-x}}{2}$ et $\text{sh } x = \frac{e^x - e^{-x}}{2}$: leurs dérivées se calculent en deux lignes avec la composée et le facteur -1 de l'exposant, et celle de $\text{th} = \frac{\text{sh}}{\text{ch}}$ par la formule du quotient, exactement comme pour la tangente, le dénominateur ne s'annulant jamais.

Les bornes du domaine sont là où tout se joue. Le tableau donne $(\sqrt{x})' = \frac{1}{2\sqrt{x}}$ sur $]0, +\infty[$ et **ne dit rien en 0**, ce qui est correct : la fonction y est définie, continue, et non dérivable. De même pour Arcsin et Arccos en ± 1 , et pour la valeur absolue en 0. Un tableau ne dispense jamais d'examiner les extrémités.

La ligne des puissances réelles est un piège de notation. La formule $(x^\alpha)' = \alpha x^{\alpha-1}$ vaut sur $]0, +\infty[$ **seulement**, parce que $x^\alpha = e^{\alpha \ln x}$ n'y est défini que là. Pour un exposant entier, la première et la deuxième ligne prennent le relais sur $\mathbb{R}$ ou $\mathbb{R}^*$; pour la racine cubique, il faut passer par la réciproque, comme en section 8. Écrire $(x^{1/3})' = \frac{1}{3}x^{-2/3}$ pour $x < 0$ est une faute **tant qu'on s'en tient à cette définition**, même si le second membre a l'air de vouloir dire quelque chose.

La convention étendue, et à quelle condition elle est licite. Certains textes — et, dans ce chapitre, les énoncés des exercices 01 et 05, qui le disent en toutes lettres avant leur première question — posent au contraire que $x^{1/3}$ désigne la racine cubique, donc une fonction définie sur $\mathbb{R}$ tout entier, et que $x^{2/3}$ désigne $(x^{1/3})^2$. Sous cette convention, l'égalité ci-dessus devient **vraie sur $\mathbb{R}^*$** , puisque $x^{-2/3}$ y vaut $1/(\sqrt[3]{x})^2$, qui est strictement positif des deux côtés de 0. Les deux lectures sont défendables ; ce qui ne l'est pas, c'est de changer de convention en cours de copie sans le dire. **La règle est donc : on écrit $\sqrt[3]{x}$ par défaut, et si l'on emploie $x^{1/3}$ hors de $]0, +\infty[$, on annonce la convention avant de s'en servir.**

10.10. Extremum local et lemme de Fermat

10.1 Les définitions, et le rôle des quantificateurs

Définition

Soit f définie sur une partie D de $\mathbb{R}$ et $c \in D$.

- f admet un **maximum global** en c lorsque $\forall x \in D, f(x) \leq f(c)$.
- f admet un **maximum local** en c lorsqu'il existe $r > 0$ tel que

$$\forall x \in D \cap]c - r, c + r[, \quad f(x) \leq f(c).$$

- On définit de même **minimum global** et **minimum local** en renversant l'inégalité, et l'on parle d'**extremum** pour désigner indifféremment un maximum ou un minimum.
- L'extremum est dit **strict** lorsque l'inégalité est stricte pour $x \neq c$.

Tout est dans le $\exists r > 0$. Un maximum local n'affirme rien au-delà d'un voisinage de c , dont le rayon peut être minuscule et n'a pas à être précisé. Un maximum global le remplace par « pour tout x du domaine », ce qui est infiniment plus fort. L'implication va donc dans un seul sens : **global entraîne local**, jamais l'inverse. Une fonction peut avoir une infinité de maximums locaux et aucun maximum global, comme $x \mapsto x + \sin x$ n'en a aucun tout en ayant une infinité de points à tangente horizontale.

Il faut enfin un mot pour dire qu'un point n'est pas au bord.

Définition

Un point c est **intérieur** à I lorsqu'il existe $r > 0$ tel que $]c - r, c + r[\subset I$.

Pour un segment $[a, b]$, l'ensemble des points intérieurs est exactement $]a, b[$. La notion dit une seule chose, et c'est celle qui va servir : **il y a des points de I des deux côtés de c .**

10.2 Le lemme de Fermat

Théorème

Lemme de Fermat. Soit f définie sur un intervalle I et c un point **intérieur** à I . Si f admet un extremum **local** en c et si f est **dérivable** en c , alors

$$f'(c) = 0.$$

Démonstration. Traitons le cas d'un maximum local ; le cas d'un minimum s'en déduit en appliquant le résultat à $-f$, qui admet alors un maximum local en c et dont la dérivée est $-f'$.

Deux réels strictement positifs sont donnés par les hypothèses : un rayon r_1 tel que $]c - r_1, c + r_1[\subset I$, par l'intériorité de c , et un rayon r_2 tel que $f(x) \leq f(c)$ pour tout x de I dans $]c - r_2, c + r_2[$, par la définition du maximum local. Posons $r = \min(r_1, r_2)$, qui est strictement positif. Sur l'intervalle $]c - r, c + r[$, les deux propriétés sont vraies en même temps : il est contenu dans I , et f y est majorée par $f(c)$.

À droite. Soit $x \in]c, c + r[$. Alors $f(x) - f(c) \leq 0$ et $x - c > 0$, donc le quotient est négatif ou nul :

$$\tau_c(x) = \frac{f(x) - f(c)}{x - c} \leq 0.$$

Comme f est dérivable en c , ce taux tend vers $f'(c)$ quand x tend vers c par valeurs supérieures. Le passage à la limite dans une inégalité large, démontré au chapitre précédent, donne $f'(c) \leq 0$.

À gauche. Soit $x \in]c - r, c[$. Alors $f(x) - f(c) \leq 0$ et cette fois $x - c < 0$, donc le quotient est positif ou nul : $\tau_c(x) \geq 0$. Le même passage à la limite donne $f'(c) \geq 0$.

Les deux inégalités ensemble donnent $f'(c) = 0$. $\square$

Trois points de cette démonstration méritent d'être relus.

Le **minimum des deux rayons** est le geste standard du chapitre précédent, celui qui rend deux propriétés valables simultanément ; sans lui, on manipulerait deux voisinages différents sans droit de les confondre.

L'**intérieurité** sert exactement une fois, et de façon décisive : elle fournit des points x **des deux côtés** de c , donc les deux inégalités. C'est tout son rôle, et c'est pourquoi elle ne peut pas être affaiblie.

Enfin, on n'obtient que des inégalités **larges**, parce que le passage à la limite ne conserve pas la stricte inégalité. Ce n'est pas une faiblesse de la rédaction : c'est la règle rappelée au chapitre précédent, et la conclusion $f'(c) = 0$ ne demande rien de plus.

10.3 Les deux réserves

Réserve 1 : l'hypothèse « intérieur » est indispensable. Considérons $f : x \mapsto x$ sur $[0, 1]$. Elle atteint son minimum global en 0 et son maximum global en 1, et pourtant f' vaut constamment 1 : elle ne s'annule nulle part. Il n'y a aucune contradiction, puisque 0 et 1 ne sont pas intérieurs à $[0, 1]$.

Ce que le lemme dit encore aux bornes, c'est **une inégalité sur une seule dérivée latérale**, en reprenant la moitié de la démonstration. Et il faut prendre le signe au sérieux, car il dépend de deux choses à la fois : de la nature de l'extremum, et du côté où l'on peut calculer.

En un **minimum** atteint en a , extrémité gauche, seul le calcul à droite est possible : pour $x > a$ on a $f(x) - f(a) \geq 0$ et $x - a > 0$, donc le taux est positif, et $f'_d(a) \geq 0$. En un **maximum** atteint en b , extrémité droite, seul le calcul à gauche est possible : pour $x < b$ on a $f(x) - f(b) \leq 0$ et $x - b < 0$, donc le taux est encore **positif**, et $f'_g(b) \geq 0$. Sur l'exemple $x \mapsto x$, où $f' = 1 > 0$, ces deux inégalités sont bien vérifiées — et l'on voit du même coup que le signe attendu n'est pas celui que le mot « maximum » suggère.

Les deux autres cas se traitent de la même façon et donnent l'inégalité inverse : un **maximum** en a donne $f'_d(a) \leq 0$, un **minimum** en b donne $f'_g(b) \leq 0$. L'exercice 08 fait dresser le tableau complet de ces quatre cas.

Réserve 2 : la réciproque est fausse. La fonction $x \mapsto x^3$ vérifie $f'(0) = 0$ et n'admet en 0 **aucun** extremum, pas même local : elle est strictement croissante sur $\mathbb{R}$, donc elle prend des valeurs strictement inférieures à $f(0)$ à gauche et strictement supérieures à droite, dans tout voisinage de 0. Un point où la dérivée s'annule s'appelle un **point critique**, et ce vocabulaire est prudent à dessein : il ne dit pas « extremum ».

Méthode

Comment on rédige la recherche d'un extremum. La règle tient en une phrase : $f'(c) = 0$ est une condition **nécessaire**, jamais suffisante, et elle ne prouve rien à elle seule.

Ce qu'on écrit. « La fonction f est dérivable sur $]a, b[$; si elle admet un extremum local en un point c de $]a, b[$, alors $f'(c) = 0$ d'après le lemme de Fermat. Les points critiques sont donc $c_1, \dots, c_p$, et ce sont les **seuls candidats** intérieurs. »

Ce qu'on écrit ensuite, et qui est le vrai travail : on **conclut** par un autre argument. Le signe de f' de part et d'autre du point, obtenu à la section 15, transforme le candidat en extremum ou le disqualifie ; sur un segment, la comparaison des valeurs de f aux points critiques et aux deux extrémités désigne le maximum et le minimum globaux, dont l'existence est garantie par le théorème des bornes atteintes du chapitre précédent.

La faute type, et elle se paie cher : « $f'(c) = 0$ donc f admet un extremum en c ». C'est la réciproque du lemme, et $x \mapsto x^3$ suffit à la réfuter.

La faute jumelle, plus discrète : chercher les extremums d'une fonction sur un segment en ne regardant que les points critiques, et oublier les deux bornes. Sur $[0, 1]$, la fonction $x \mapsto x$ n'a aucun point critique et deux extremums.

L'exercice 08 fait démontrer le lemme, puis demande où l'hypothèse d'intériorité a servi **exactement** dans la preuve, et construit les situations où elle tombe en faisant dire ce qui subsiste dans chacune.

11 11. Théorème de Rolle

Théorème

Théorème de Rolle. Soient a et b deux réels avec $a < b$, et f une fonction définie sur $[a, b]$ vérifiant les **trois** hypothèses suivantes :

1. **(continuité)** f est continue sur le **segment** $[a, b]$;
2. **(dérivabilité)** f est dérivable sur l'**intervalle ouvert** $]a, b[$;
3. **(valeurs égales)** $f(a) = f(b)$.

Alors il existe au moins un réel $c \in]a, b[$ tel que $f'(c) = 0$.

Trois précisions avant la démonstration, parce qu'elles conditionnent tous les usages.

C'est un théorème d'existence. Il affirme qu'un tel c existe, il n'en donne aucun, et il n'affirme pas qu'il soit unique : une fonction constante vérifie les hypothèses et **tout** point de $]a, b[$ convient. On écrit donc « il existe $c \in]a, b[$ tel que », et jamais « le point c ».

La conclusion place c dans l'ouvert. C'est essentiel pour les applications : le point trouvé est strictement entre a et b , ce qui permet d'itérer le théorème sur des sous-intervalles.

Les trois hypothèses sont indépendantes, et retirer n'importe laquelle fait tomber la conclusion. En voici une seule illustration, celle de l'hypothèse 2, qui est la plus instructive : sur $[-1, 1]$, la valeur absolue est continue sur le segment et vérifie $f(-1) = f(1)$, donc les hypothèses 1 et 3 ; mais elle n'est pas dérivable en 0, et sa dérivée vaut -1 à gauche et $+1$ à droite, sans jamais s'annuler. La conclusion tombe.

Les deux autres réfutations se construisent de la même façon. L'exercice 09 fait reprendre les trois, celle-ci comprise, et c'est voulu : la rédiger soi-même après l'avoir lue est le meilleur moyen de voir ce qu'un contre-exemple doit contenir. Le point de rigueur y est essentiel : un contre-exemple ne vaut que si l'on **vérifie que les deux hypothèses restantes sont bien satisfaites**, sans quoi il ne prouve rien sur celle qu'on voulait mettre en cause.

11.1 La démonstration

Démonstration. La fonction f est continue sur le segment $[a, b]$: le **théorème des bornes atteintes**, démontré au chapitre précédent, s'applique. Il fournit deux points x_m et x_M de $[a, b]$ tels que

$$f(x_m) = m = \min_{[a,b]} f \quad \text{et} \quad f(x_M) = M = \max_{[a,b]} f.$$

Notons v la valeur commune $f(a) = f(b)$, donnée par l'hypothèse 3. Deux cas se présentent.

Premier cas : $m = M$. Alors f est comprise entre m et M en tout point, donc constante sur $[a, b]$, égale à m . Son taux d'accroissement est nul en tout point, donc $f' = 0$ sur $]a, b[$. Comme $a < b$, l'intervalle $]a, b[$ n'est pas vide : n'importe lequel de ses points convient comme c .

Second cas : $m < M$. On a toujours $m \leq v \leq M$, puisque v est une valeur de f . Ces deux inégalités ne peuvent pas être **toutes deux** des égalités, sans quoi $m = M$. Donc $m < v$, ou bien $M > v$, et il suffit de traiter l'un des deux, l'autre étant symétrique.

Supposons $M > v$, et considérons le point x_M . Comme $f(x_M) = M \neq v = f(a) = f(b)$, ce point n'est ni a ni b : il appartient donc à $]a, b[$, et il est **intérieur** à $[a, b]$. Par ailleurs f admet en x_M un maximum global sur $[a, b]$, donc en particulier un maximum local. Enfin f est dérivable en x_M , par l'hypothèse 2, puisque $x_M \in]a, b[$. Les trois hypothèses du lemme de Fermat sont réunies, et il donne $f'(x_M) = 0$. Le réel $c = x_M$ convient.

Si c'est $m < v$ qui a lieu, le même raisonnement appliqué au point x_m , où f admet un minimum local intérieur, donne $f'(x_m) = 0$, et $c = x_m$ convient. $\square$

Le cas où maximum et minimum sont tous deux atteints aux bords est celui que la disjonction ci-dessus règle sans le nommer, et il vaut la peine de le voir explicitement. Si x_m et x_M sont l'un et l'autre dans $\{a, b\}$, alors m et M sont l'un et l'autre égaux à v , donc $m = M = v$: on est dans le premier cas, la fonction est constante, et la conclusion est immédiate. C'est précisément pourquoi la démonstration commence par ce cas au lieu de le traiter comme une exception : il n'est pas dégénéré, il est le seul où le lemme de Fermat ne sert pas.

11.2 Pourquoi les hypothèses sont dissymétriques

Un fermé pour la continuité, un ouvert pour la dérivabilité : cette dissymétrie surprend, et elle se lit directement dans la démonstration.

La **continuité sur le segment fermé** est ce qui permet d'appeler le théorème des bornes atteintes, lequel exige un segment : c'est le caractère borné qui empêche f de s'échapper, et le caractère fermé qui garantit que les points x_m et x_M appartiennent bien à l'ensemble.

La **dérivabilité sur l'ouvert** suffit, parce qu'elle n'est utilisée qu'en un seul point, celui que produit le théorème des bornes atteintes, et que ce point est intérieur dès qu'on n'est pas dans le cas constant. Exiger la dérivabilité en a et en b serait une hypothèse plus forte et parfaitement inutile.

Cette économie n'est pas gratuite : elle **autorise des tangentes verticales aux extrémités**, et c'est ce qui rend le théorème applicable à des situations très courantes.

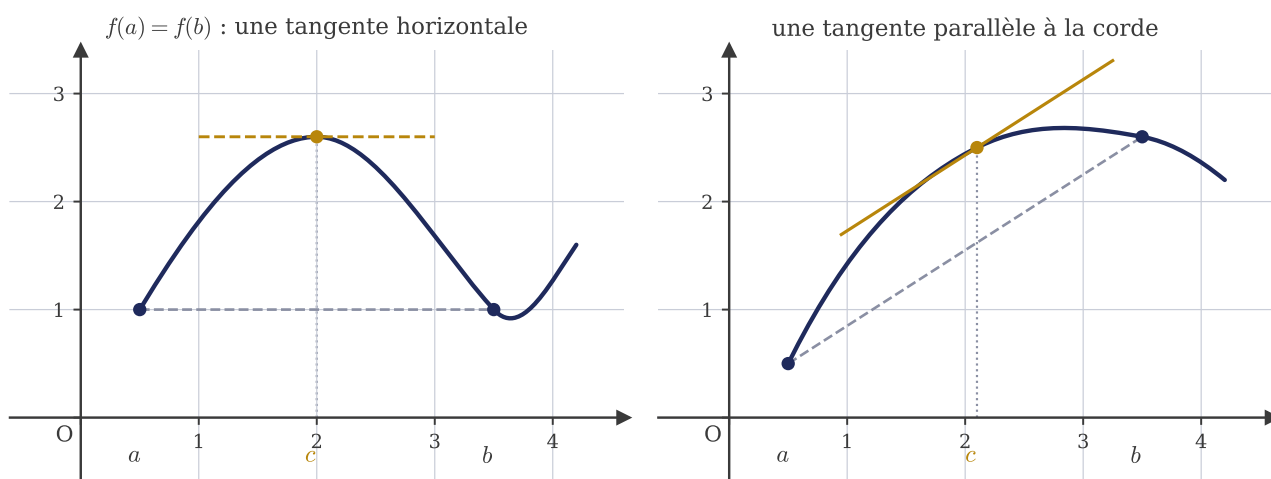
Exemple

Un demi-cercle. Soit $f : x \mapsto \sqrt{1-x^2}$ sur $[-1, 1]$. Elle est continue sur $[-1, 1]$ comme composée de fonctions continues, la quantité $1-x^2$ y étant positive. Elle est dérivable sur $] -1, 1[$, où $1-x^2 > 0$, avec

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

Enfin $f(-1) = f(1) = 0$. Les trois hypothèses sont vérifiées, et le théorème de Rolle s'applique : le point $c = 0$ convient, puisque $f'(0) = 0$.

Or f n'est **pas** dérivable en -1 ni en 1 : la dérivée ci-dessus tend vers $\pm\infty$ aux bornes, et la courbe, qui est un demi-cercle, y présente des tangentes verticales. Si le théorème avait exigé la dérivabilité sur $[-1, 1]$, il ne se serait pas appliqué ici.



La figure comporte deux panneaux. Celui de **gauche** illustre le théorème de Rolle : la corde qui joint les points d'abscisses a et b est horizontale, puisque $f(a) = f(b)$, et la courbe admet en un point c strictement compris entre a et b une tangente horizontale, qui est la conclusion. Celui de **droite** anticipe la section 12 : la corde n'y est plus horizontale, et la tangente en c lui est **parallèle** au lieu d'être horizontale. C'est l'égalité des accroissements finis, qui sera obtenue à partir du théorème de Rolle par une fonction auxiliaire, et qui contient Rolle comme le cas particulier où la corde est horizontale.

11.3 Compter les racines avec Rolle

L'usage le plus fréquent du théorème n'est pas de trouver un point c , dont on ne sait rien, mais de **dénombrer** les solutions d'une équation. Le principe tient dans un corollaire d'une ligne.

Propriété

Soit f continue sur un intervalle I et dérivable sur son intérieur. **Entre deux racines de f , il y a au moins une racine de f' .**

Par conséquent, si f' admet au plus k racines sur I , alors f en admet au plus $k+1$.

Démonstration. Soient $\alpha < \beta$ deux racines de f dans I . La fonction f est continue sur $[\alpha, \beta]$, dérivable sur $] \alpha, \beta[$, et $f(\alpha) = f(\beta) = 0$: le théorème de Rolle donne un point de $] \alpha, \beta[$ où f' s'annule.

Pour la seconde phrase, raisonnons par l'absurde. Si f avait $k+2$ racines distinctes, on pourrait les ranger par ordre croissant ; entre deux racines consécutives, la première partie fournit une racine de f' , et ces $k+1$ racines de f' sont deux à deux distinctes, puisqu'elles vivent dans des intervalles ouverts disjoints. Cela contredit l'hypothèse. $\square$

Méthode

Comment on rédige un comptage de racines. Le raisonnement se fait toujours en deux temps, et les deux ne se démontrent pas du tout de la même façon.

1. **Au moins tant de racines** : c'est le théorème des valeurs intermédiaires du chapitre précédent, appliqué sur des intervalles où f change de signe. Ce sens ne fait jamais intervenir la dérivée.
2. **Au plus tant de racines** : c'est le corollaire ci-dessus, ou bien la stricte monotonie obtenue par le signe de f' en section 15. Ce sens ne fait jamais intervenir le TVI.

Les deux ensemble donnent le nombre exact. Une copie qui n'en produit qu'un a démontré une inégalité, pas une égalité, et il faut le dire.

L'exercice 10 applique ce corollaire à une équation polynomiale à paramètre et à une équation transcendante, et fait comparer les deux façons de compter : par ce corollaire, qui MAJORE le nombre de racines, et par le tableau de variations, qui donne le compte EXACT.

Le théorème de Rolle a par ailleurs un usage qui dépasse le comptage, et c'est le plus fécond : appliqué non pas à f , mais à une **fonction auxiliaire** construite pour la circonstance, il produit des résultats qui ne ressemblent plus du tout à son énoncé. Les sections 12, 16 et 19 en donnent trois exemples, et l'exercice 11 en fait travailler un.

Le théorème de Rolle est le dernier énoncé de cette première partie, et il en est aussi le pivot : tout le bloc qui commence à la section 12, de l'égalité des accroissements finis jusqu'à la formule de Taylor-Lagrange, s'obtient en l'appliquant à des fonctions auxiliaires construites pour l'occasion.

12 12. L'égalité des accroissements finis

Le théorème de Rolle, établi à la section 11, souffre d'une hypothèse très particulière : il exige $f(a) = f(b)$. Autant dire qu'il ne s'applique presque jamais tel quel. Pourtant, c'est de lui que descend tout le reste du chapitre, et le passage se fait par un geste unique, qu'il faut voir une fois pour ne plus l'oublier : **on retranche à f ce qui l'empêche de prendre la même valeur aux deux bornes.**

12.1 L'énoncé

Théorème

Égalité des accroissements finis. Soient a et b deux réels tels que $a < b$, et f une fonction définie sur $[a, b]$ à valeurs réelles, telle que :

1. f est **continue sur le segment** $[a, b]$;
2. f est **dérivable sur l'intervalle ouvert** $]a, b[$.

Alors il existe au moins un réel $c \in]a, b[$ tel que

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

Les deux hypothèses sont **séparées**, et elles ne portent pas sur le même ensemble : la continuité est exigée sur le segment fermé, la dérivabilité seulement sur l'ouvert. Ce n'est pas une coquetterie. Sur $[0, 1]$, la fonction $x \mapsto \sqrt{x}$ est continue partout, y compris en 0, mais elle n'est pas dérivable en 0 : son taux d'accroissement en 0 tend vers $+\infty$, et sa courbe y admet une tangente verticale, comme la section 5 l'a décrit. L'égalité des accroissements finis s'y applique malgré tout, puisqu'elle ne demande la dérivabilité qu'à l'intérieur, et elle fournit un c dans $]0, 1[$ tel que $\frac{1}{2\sqrt{c}} = 1$, c'est-à-dire $c = \frac{1}{4}$. Une version du théorème qui exigerait la dérivabilité sur $[0, 1]$ tout entier laisserait cette fonction de côté pour rien.

Dans la suite, on abrège ce théorème en **EAF**, comme le fait la plupart des cours de première année.

12.2 D'où vient la fonction auxiliaire

Personne ne devine la fonction auxiliaire d'une démonstration ; on la **fabrique**, et on peut dire exactement comment.

On veut appliquer Rolle. Rolle demande une fonction qui prend la même valeur en a et en b . La fonction f ne le fait pas : elle passe de $f(a)$ à $f(b)$. On lui retranche donc la fonction affine qui fait exactement le même trajet, c'est-à-dire **la corde** qui joint les points $A(a, f(a))$ et $B(b, f(b))$ de la courbe. Cette corde a pour pente

$$p = \frac{f(b) - f(a)}{b - a},$$

et pour équation $y = f(a) + p(x - a)$, puisqu'elle passe par A .

La différence entre la courbe et sa corde est donc la fonction

$$\varphi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} (x - a),$$

qui mesure, en chaque abscisse x , l'**écart vertical** entre le point de la courbe et le point de la corde. Cet écart est nul en a et nul en b , pour la raison la plus simple du monde : en ces deux abscisses, la courbe et la corde se rencontrent. C'est exactement l'hypothèse de Rolle, et elle est obtenue par construction.

12.3 La démonstration

Démonstration. Posons, pour tout $x \in [a, b]$,

$$\varphi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a} (x - a).$$

Vérifions une par une les trois hypothèses du théorème de Rolle.

Continuité sur $[a, b]$. La fonction $x \mapsto f(a) + \frac{f(b) - f(a)}{b - a} (x - a)$ est une fonction affine, donc continue sur $\mathbb{R}$; le coefficient $\frac{f(b) - f(a)}{b - a}$ est un nombre réel fixé, puisque $b - a \neq 0$. La fonction f est continue sur $[a, b]$ par hypothèse. La différence de deux fonctions continues est continue : φ est continue sur $[a, b]$.

Dérivabilité sur $]a, b[$. Même argument, avec le théorème des opérations de la section 7 : f est dérivable sur $]a, b[$ par hypothèse, la fonction affine l'est sur $\mathbb{R}$, donc φ est dérivable sur $]a, b[$, et pour tout x de cet intervalle

$$\varphi'(x) = f'(x) - \frac{f(b) - f(a)}{b - a},$$

la dérivée d'une fonction affine étant son coefficient directeur.

Égalité des valeurs aux bornes. On calcule les deux :

$$\varphi(a) = f(a) - f(a) - \frac{f(b) - f(a)}{b - a} \times 0 = 0,$$

et

$$\varphi(b) = f(b) - f(a) - \frac{f(b) - f(a)}{b - a} \times (b - a) = (f(b) - f(a)) - (f(b) - f(a)) = 0.$$

Donc $\varphi(a) = \varphi(b) = 0$.

Les trois hypothèses sont vérifiées. Le théorème de Rolle, établi à la section 11, fournit donc un réel $c \in]a, b[$ tel que $\varphi'(c) = 0$, c'est-à-dire

$$f'(c) - \frac{f(b) - f(a)}{b - a} = 0, \quad \text{soit} \quad \frac{f(b) - f(a)}{b - a} = f'(c).$$

En multipliant par $b - a$, on obtient $f(b) - f(a) = f'(c)(b - a)$. $\square$

Remarque

Rolle et l'EAF sont le même théorème. L'EAF contient Rolle : il suffit d'y faire $f(a) = f(b)$ pour obtenir $f'(c) = 0$. Et la démonstration ci-dessus montre que Rolle contient l'EAF. Les deux énoncés sont donc équivalents, et l'on démontre Rolle parce que c'est celui dont la preuve est directe, à partir du lemme de Fermat et du théorème des bornes atteintes.

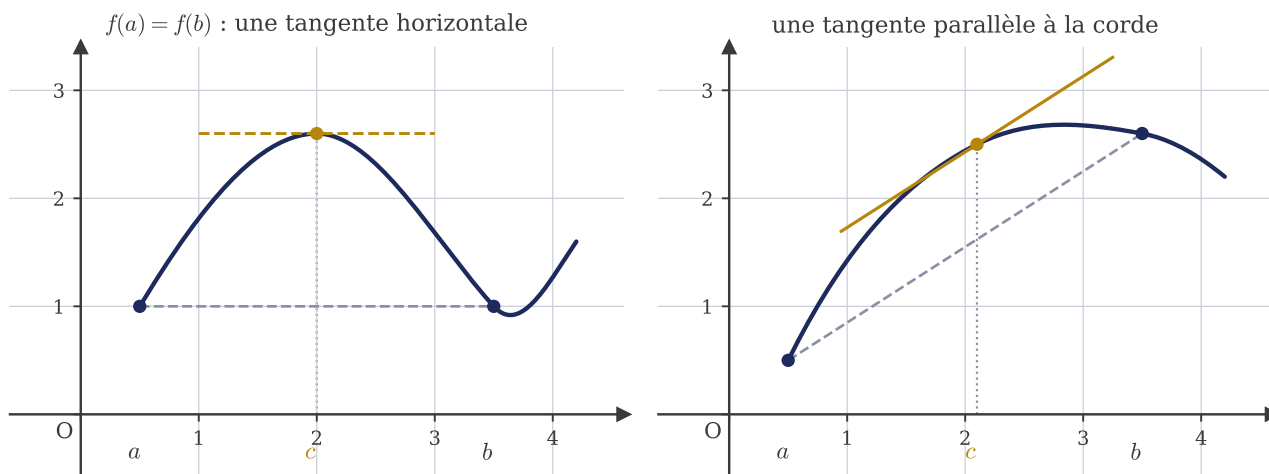
Ce qu'il faut retenir n'est pas cette équivalence, c'est le **procédé** : pour appliquer Rolle à une situation qui ne s'y prête pas, on retranche à f une fonction simple, choisie pour annuler l'obstacle. La section 16 refait exactement ce geste, avec une autre fonction simple.

Une seconde écriture de la conclusion sert très souvent. En posant $h = b - a > 0$, le point c appartient à $]a, a + h[$, donc s'écrit $c = a + \theta h$ avec $\theta \in]0, 1[$, et l'EAF devient

$$f(a + h) = f(a) + h f'(a + \theta h).$$

Sous cette forme, on voit ce que l'égalité ajoute à l'approximation affine de la section 3 : là où l'approximation affine écrivait $f(a + h) \approx f(a) + h f'(a)$ avec une erreur non contrôlée, l'EAF écrit une **égalité exacte**, au prix d'un point $a + \theta h$ dont on ne sait rien d'autre que sa position entre a et $a + h$.

12.4 Ce que dit la figure



C'est la figure déjà employée à la section 11, et elle comporte deux panneaux : celui de gauche illustre Rolle, celui de droite illustre l'EAF.

Sur le panneau de droite, la droite en pointillés est la corde qui joint $A(a, f(a))$ à $B(b, f(b))$: sa pente vaut $\frac{f(b) - f(a)}{b - a}$. Le segment tracé en un point intérieur est la tangente à la courbe au point d'abscisse c : sa pente vaut $f'(c)$. La conclusion du théorème dit que ces deux pentes sont égales, autrement dit :

il existe au moins un point de la courbe où la tangente est parallèle à la corde.

Le panneau de gauche est le cas particulier où la corde est horizontale, c'est-à-dire $f(a) = f(b)$: la tangente parallèle à la corde y est une tangente horizontale, et l'on retrouve Rolle. Passer d'un panneau

à l'autre, c'est **incliner la figure**, et c'est très exactement ce que fait la soustraction de la corde dans la démonstration.

12.5 L'interprétation cinématique, et le radar tronçon

Supposons que $f(t)$ soit la position d'un véhicule à l'instant t , sur un axe. Alors $f'(t)$ est sa **vitesse instantanée** à l'instant t , celle qu'affiche le compteur, et le quotient

$$\frac{f(b) - f(a)}{b - a} = \frac{\text{distance parcourue}}{\text{durée du trajet}}$$

est sa **vitesse moyenne** sur le trajet. L'EAF affirme alors qu'à un instant au moins, la vitesse instantanée a été exactement égale à la vitesse moyenne.

C'est le principe du **radar tronçon**, et il vaut la peine de le dire dans ces termes : un véhicule qui parcourt 30 km en 15 minutes a roulé à 120 km/h de moyenne ; l'EAF garantit qu'il existe un instant précis du trajet où son compteur affichait exactement 120 km/h. L'automobiliste ne peut donc pas se défendre en disant qu'il n'a jamais dépassé 110 : s'il n'avait jamais dépassé 110, sa vitesse moyenne aurait été inférieure à 110, ce qui est faux. Ce raisonnement est d'ailleurs la première application de l'**inégalité** des accroissements finis, qui fait l'objet de la section 13.

12.6 Ce que le théorème ne donne pas

C'est ici que se joue la différence entre un étudiant qui a compris l'énoncé et un étudiant qui l'a recopié. Quatre choses ne sont **pas** dans la conclusion.

1. Le point c n'est pas construit, il est seulement affirmé. La démonstration remonte à Rolle, qui remonte au lemme de Fermat et au théorème des bornes atteintes, qui remonte lui-même au théorème de Bolzano-Weierstrass du chapitre « Suites réelles ». À aucun moment on ne fabrique c : on démontre qu'il ne peut pas ne pas exister. Une démonstration d'existence pure ne fournit jamais l'objet.

2. Le point c n'est pas unique. Sur $[0, 2\pi]$, la fonction $\sin$ vérifie $\sin(2\pi) - \sin(0) = 0$; l'EAF fournit donc un c tel que $\cos c = 0$, et il y en a **deux** dans $]0, 2\pi[$: $c = \frac{\pi}{2}$ et $c = \frac{3\pi}{2}$. L'énoncé dit « il existe au moins un », jamais « il existe un unique ».

3. Le point c n'est en général pas calculable. Il arrive qu'on sache le déterminer : sur $[0, 1]$, la fonction $\exp$ donne $e - 1 = e^c$, donc $c = \ln(e - 1)$, soit environ 0,5413. C'est l'exception. Prenons $f(x) = e^x + x^3$ sur $[0, 1]$: la relation $f(1) - f(0) = f'(c)$ s'écrit $e^c + 3c^2 = e$, une équation qui mêle une exponentielle et un polynôme, et dont aucune méthode ne donne la solution sous forme close. Le réel c existe pourtant, et c'est tout ce dont on a besoin.

4. Le point c dépend de a et de b . C'est la conséquence la plus souvent oubliée. Si l'on applique l'EAF successivement sur $[a, b]$ puis sur $[a, b']$, on obtient **deux** points, et rien ne dit qu'ils sont égaux. Réutiliser le même nom c pour les deux, ou traiter c comme une fonction de b qu'on pourrait dériver, sont deux fautes du même ordre. Quand la confusion menace, on note le point c_b , ou l'on requantifie explicitement à chaque application.

Méthode

Comment on rédige une application de l'EAF. Quatre lignes, et elles s'écrivent toujours dans cet ordre.

1. **Nommer le segment, dans le bon sens.** « Appliquons l'égalité des accroissements finis à f sur le segment $[x, y]$, où $x < y$. » Si l'on ne sait pas lequel de x et y est le plus petit, on le dit et l'on traite les deux cas, ou l'on passe d'emblée aux valeurs absolues, comme le fait la section 13.
2. **Vérifier les deux hypothèses séparément, avec la bonne nature d'intervalle.** « f est continue sur $[x, y]$ » et « f est dérivable sur $]x, y[$ » sont deux phrases, pas une. Chacune se justifie, en général par les théorèmes d'opérations.
3. **Écrire la conclusion sous forme quantifiée.** « Il existe $c \in]x, y[$ tel que $f(y) - f(x) = f'(c)(y - x)$. » Le mot « existe » n'est pas décoratif : il dit que c n'est pas à notre disposition.
4. **N'utiliser ensuite de c que ce que l'on sait de lui**, c'est-à-dire son **appartenance** à $]x, y[$. En pratique, cela signifie une seule chose : on majore ou l'on minore $f'(c)$ à l'aide d'un encadrement de f' **valable sur tout l'intervalle**. C'est le seul usage licite du point c , et c'est aussi le plus fécond : les sections 13, 14, 15 et 17 ne font rien d'autre.

L'exercice 12 fait travailler l'énoncé pour lui-même : on y détermine explicitement le ou les points c dans les rares cas où le calcul aboutit, on y compte les solutions, et l'on y examine ce qui se passe quand une hypothèse est retirée.

Remarque

Une frontière à connaître. Tout ce qui précède suppose f à valeurs **réelles**. Pour une fonction à valeurs complexes, l'**égalité** des accroissements finis est fausse : la fonction $t \mapsto e^{it}$ définie sur $[0, 2\pi]$ y prend la même valeur aux deux bornes, alors que sa dérivée $t \mapsto ie^{it}$ ne s'annule jamais, son module valant constamment 1. Aucun point c ne peut donc convenir, et Rolle tombe avec l'EAF. Seule l'**inégalité** de la section 13 survit dans ce cadre, avec une démonstration différente. On ne développe pas ce sujet ici.

13 13. L'inégalité des accroissements finis, et les fonctions lipschitziennes

L'EAF fournit un point dont on ne sait rien. L'inégalité des accroissements finis consiste à s'en débarrasser tout de suite, en le majorant, et c'est ce qui la rend utilisable partout. On l'abrège en **IAF**.

13.1 L'énoncé et sa démonstration

Théorème

Inégalité des accroissements finis. Soit I un **intervalle** de $\mathbb{R}$ et f une fonction dérivable sur I . On suppose qu'il existe un réel $M \geq 0$ tel que

$$\forall t \in I, \quad |f'(t)| \leq M.$$

Alors

$$\forall (x, y) \in I^2, \quad |f(x) - f(y)| \leq M |x - y|.$$

Démonstration. Soient x et y dans I . Si $x = y$, les deux membres sont nuls et l'inégalité est vraie. Sinon, remarquons que les deux membres sont **symétriques** en x et y : échanger les noms ne change ni $|f(x) - f(y)|$ ni $|x - y|$. On peut donc supposer $x < y$, quitte à échanger.

Comme I est un intervalle et que x et y lui appartiennent, le segment $[x, y]$ est **contenu dans** I . C'est le seul endroit où l'hypothèse « intervalle » sert, et elle y est indispensable.

La fonction f est dérivable sur I , donc sur $[x, y]$; elle y est en particulier continue, la dérivabilité entraînant la continuité (section 4). Les deux hypothèses de l'EAF sont donc vérifiées sur $[x, y]$, et il existe $c \in]x, y[$ tel que

$$f(y) - f(x) = f'(c)(y - x).$$

Passons aux valeurs absolues : $|f(y) - f(x)| = |f'(c)| \times (y - x)$, puisque $y - x > 0$. Le point c appartient à I , donc $|f'(c)| \leq M$, et

$$|f(y) - f(x)| \leq M(y - x) = M|y - x|. \quad \square$$

Regardons ce qui vient de se passer. **On n'a rien su de c , et on n'avait besoin de rien** : la majoration de $|f'|$ vaut sur tout I , donc en particulier au point inconnu. C'est le mode d'emploi standard du point c , annoncé à la fin de la section 12, et c'est pour cela que l'IAF est utilisable là où l'EAF ne l'est pas.

Propriété

Version encadrée, quand le signe compte. Si f est dérivable sur l'intervalle I et s'il existe deux réels m et M tels que $m \leq f'(t) \leq M$ pour tout $t \in I$, alors, pour tous $x < y$ de I ,

$$m(y - x) \leq f(y) - f(x) \leq M(y - x).$$

La démonstration est la même, sans les valeurs absolues : on encadre $f'(c)$ au lieu de le majorer en module, puis on multiplie par $y - x$, qui est strictement positif et ne retourne donc pas les inégalités. Cette version sert dès qu'on veut un encadrement **orienté** et non une simple majoration d'écart, par exemple pour comparer f à une fonction affine.

13.2 Les fonctions lipschitziennes

Définition

Soit f définie sur une partie A de $\mathbb{R}$ et $k \geq 0$. On dit que f est **k -lipschitzienne sur A** lorsque

$$\forall (x, y) \in A^2, \quad |f(x) - f(y)| \leq k|x - y|,$$

et **lipschitzienne** lorsqu'il existe un tel k , appelé **constante** ou **rapport** de Lipschitz.

Cette définition n'est pas nouvelle : le chapitre « Fonctions d'une variable réelle : limites et continuité » l'a posée, et il y a démontré que **toute fonction lipschitzienne est uniformément continue**, donc continue, avec le $\delta = \frac{\varepsilon}{k}$ explicite. On ne le redémontre pas ici. Ce que la dérivation apporte, c'est un **moyen de vérifier** cette propriété sans revenir à la définition.

Propriété

Le lien. Soit I un intervalle et f dérivable sur I . Alors, pour $k \geq 0$:

$$f \text{ est } k\text{-lipschitzienne sur } I \iff |f'| \leq k \text{ sur } I.$$

Démonstration. Le sens direct de droite à gauche est exactement l'IAF, déjà démontrée.

Réciproquement, supposons f k -lipschitzienne sur I , et soit $a \in I$. Pour tout $x \in I$ distinct de a , la définition donne

$$\left| \frac{f(x) - f(a)}{x - a} \right| = \frac{|f(x) - f(a)|}{|x - a|} \leq k.$$

Le taux d'accroissement est donc majoré en module par k sur tout un voisinage épointé de a . Comme il tend vers $f'(a)$ quand x tend vers a , le passage à la limite dans une inégalité **large** (chapitre précédent) donne $|f'(a)| \leq k$. $\square$

Deux conséquences valent d'être notées. D'abord, pour une fonction dérivable sur un intervalle, **la meilleure constante de Lipschitz est** $\sup_I |f'|$: aucune valeur plus petite ne convient, et cette valeur convient dès qu'elle est finie. Ensuite, le passage à la limite se fait dans une inégalité large, et c'est pour cela qu'une fonction dont la dérivée vérifie $|f'| < k$ **partout** n'a aucune raison d'être k' -lipschitzienne pour un $k' < k$: le supremum peut valoir k sans être atteint. Cette nuance décide le sort de la section 14.

Exemple

Trois majorations qui reviennent constamment.

a. $\sin$ est 1-lipschitzienne sur $\mathbb{R}$, car $|\sin'| = |\cos| \leq 1$. rivée vaut $\frac{1}{1+x^2} \leq 1$.

b. $\cos$ est 1-lipschitzienne sur $\mathbb{R}$, pour la même raison. **d.** $x \mapsto \sqrt{x}$ est $\frac{1}{2}$ -lipschitzienne sur $[1, +\infty[$, car $\frac{1}{2\sqrt{x}} \leq \frac{1}{2}$ pour $x \geq 1$.

c. Arctan est 1-lipschitzienne sur $\mathbb{R}$, car sa dé-

Le point **a** mérite un mot : appliqué avec $y = 0$, il donne $|\sin x| \leq |x|$ pour tout réel x , inégalité que le chapitre précédent posait comme un acquis. Elle est ici **démontrée**, en une ligne, et l'on voit d'où elle vient : le sinus ne peut pas s'écarter de l'origine plus vite que sa pente maximale ne l'y autorise.

Le point **d** montre au passage pourquoi l'intervalle compte : sur $]0, 1]$, la dérivée $\frac{1}{2\sqrt{x}}$ n'est pas majorée, et $x \mapsto \sqrt{x}$ n'y est lipschitzienne pour aucune constante.

13.3 Contrôler une erreur

C'est l'usage le plus concret de tout le chapitre, et sans doute le seul que retiendra un étudiant qui ne fera pas de mathématiques. Le problème est le suivant : on mesure une grandeur x , on n'obtient qu'une valeur approchée x_0 , et l'on calcule $f(x_0)$ au lieu de $f(x)$. De combien se trompe-t-on ?

Méthode

Propager une incertitude. Trois temps, et le deuxième est celui qu'on saute.

1. **Écrire l'incertitude d'entrée et nommer l'intervalle.** L'énoncé « x_0 est connu à η près » signifie $|x - x_0| \leq \eta$, c'est-à-dire $x \in J$ où

$$J = [x_0 - \eta, x_0 + \eta].$$

Cet intervalle est l'objet sur lequel on va travailler ; on l'écrit noir sur blanc.

2. **Majorer $|f'|$ sur J tout entier**, et non en x_0 seulement. C'est l'étape décisive : l'IAF exige un majorant valable sur l'intervalle, et calculer $|f'(x_0)|$ ne donne rien du tout. En pratique, on étudie brièvement les variations de $|f'|$ sur J , et l'on prend la plus grande valeur, arrondie **par excès**.
3. **Conclure par l'IAF** : $|f(x) - f(x_0)| \leq M\eta$, et présenter le résultat sous la forme « $f(x_0)$ à $M\eta$ près », avec un nombre de décimales cohérent avec cette précision.

Le nombre M est le **facteur d'amplification** de l'incertitude. Il est proche de $|f'(x_0)|$ quand η est petit et f' raisonnable, mais c'est un résultat, pas un point de départ.

Exemple

Une aire mesurée. On mesure le côté d'un carré : $x_0 = 2,000$ m, à 10^{-3} m près. On veut son aire, $f(x) = x^2$.

L'intervalle. $|x - 2| \leq 10^{-3}$, donc $x \in J = [1,999, 2,001]$.

La majoration. f est dérivable sur J et $f'(x) = 2x$. Sur J , la fonction $|f'|$ est croissante et positive, donc majorée par sa valeur en 2,001 : $|f'(x)| \leq 2 \times 2,001 = 4,002 \leq 4,01$. On pose $M = 4,01$.

La conclusion. L'IAF sur l'intervalle J donne

$$|f(x) - f(2)| \leq 4,01 \times 10^{-3} \leq 4,1 \times 10^{-3}.$$

L'aire vaut donc $4,000 \text{ m}^2$ à $5 \times 10^{-3} \text{ m}^2$ près.

Lecture. L'incertitude a été multipliée par environ 4, qui est $f'(2)$. En termes **relatifs**, l'incertitude d'entrée valait $\frac{10^{-3}}{2} = 0,05 \%$ et celle de sortie vaut $\frac{4 \times 10^{-3}}{4} = 0,1 \%$: elle a **doublé**. C'est la règle bien connue selon laquelle élever au carré double l'incertitude relative, et l'on vient de la démontrer au lieu de la réciter.

Exemple

Un angle, puis un logarithme. Deux situations opposées.

Le cas confortable. Un angle est mesuré à $x_0 = 0,500$ rad, à 2×10^{-3} rad près, et l'on calcule $\sin x$. Comme $|\sin'| = |\cos| \leq 1$ sur $\mathbb{R}$ tout entier, l'IAF donne directement

$$|\sin x - \sin(0,5)| \leq 1 \times 2 \times 10^{-3} = 2 \times 10^{-3}.$$

L'erreur ne s'amplifie jamais, et c'est exactement ce que signifie « $\sin$ est 1-lipschitzienne ». On peut faire un peu mieux en travaillant sur $J = [0,498, 0,502]$: la fonction $\cos y$ est positive et décroissante, donc $|\cos| \leq \cos(0,498) \leq 0,879$, ce qui donne une erreur inférieure à $1,8 \times 10^{-3}$. Le gain est mince : c'est le signe que la majoration grossière était déjà bonne.

Le cas dangereux. Une concentration est mesurée à $x_0 = 0,010$, à 10^{-3} près, et l'on calcule $\ln x$. L'intervalle est $J = [0,009, 0,011]$, et sur J la dérivée $\frac{1}{x}$ est décroissante et positive, donc majorée par $\frac{1}{0,009} \leq 112$. L'IAF donne

$$|\ln x - \ln(0,010)| \leq 112 \times 10^{-3} \leq 0,12.$$

Pour un logarithme qui vaut environ $-4,605$, une incertitude de $0,12$ n'est pas négligeable, et l'on comprend pourquoi : une incertitude **relative** de 10 % en entrée devient une incertitude **absolue** du même ordre en sortie, parce que $\ln$ transforme précisément les rapports en différences.

La morale. Le majorant se prend sur l'intervalle, et c'est **le point le plus défavorable qui commande**. Ici, majorer $|f'|$ par $\frac{1}{0,010} = 100$, valeur au centre, aurait sous-estimé l'erreur de plus de 10 %.

L'exercice 13 traite plusieurs propagations d'incertitude de ce type, avec des fonctions où le facteur d'amplification réserve des surprises. L'exercice 07, lui, met l'IAF au service d'une fonction **réci-proque** dont on ne sait écrire aucune formule : c'est là que la combinaison de la section 8 et de la présente section devient vraiment payante, puisqu'on contrôle une fonction qu'on ne sait pas calculer.

14 14. Suites récurrentes contrôlées par l'inégalité des accroissements finis

Voici la section la plus rentable du chapitre, au sens le plus terre à terre : c'est celle dont la méthode se retrouve, presque mot pour mot, dans un partiel sur deux. Le cadre est celui d'une suite définie par une relation de récurrence $u_{n+1} = f(u_n)$, objet que le chapitre « Suites réelles » a introduit et étudié par la monotonie. On y ajoute ici un outil qui donne bien plus : une **vitesse**.

14.1 Le cadre et le théorème

Théorème

Convergence géométrique vers un point fixe. Soit I un intervalle de $\mathbb{R}$, f une fonction dérivable sur I , et (u_n) la suite définie par la donnée de u_0 et la relation $u_{n+1} = f(u_n)$. On suppose :

1. I est stable par f , c'est-à-dire $f(I) \subset I$, et $u_0 \in I$;
2. f admet un **point fixe** ℓ dans I , c'est-à-dire $f(\ell) = \ell$;
3. il existe $k \in [0, 1[$ tel que $|f'(t)| \leq k$ pour tout $t \in I$.

Alors la suite (u_n) est bien définie, tous ses termes appartiennent à I , et

$$\forall n \in \mathbb{N}, \quad |u_n - \ell| \leq k^n |u_0 - \ell|.$$

En particulier (u_n) converge vers ℓ , et ℓ est le **seul** point fixe de f dans I .

Démonstration.

La suite est bien définie et reste dans I . Par récurrence. Le terme u_0 appartient à I par hypothèse. Si $u_n \in I$, alors $f(u_n)$ a un sens, et $u_{n+1} = f(u_n) \in f(I) \subset I$. Donc $u_n \in I$ pour tout n . C'est l'hypothèse de stabilité qui sert ici, et elle sert à quelque chose de précis : sans elle, un terme pourrait sortir de I , et l'hypothèse 3 ne s'appliquerait plus à lui.

La majoration. Par récurrence sur n . Au rang 0, l'inégalité s'écrit $|u_0 - \ell| \leq |u_0 - \ell|$, qui est vraie. Supposons-la vraie au rang n . Les deux réels u_n et ℓ appartiennent à I , et $|f'| \leq k$ sur I : l'inégalité des accroissements finis s'applique à ce couple de points. Or, **et c'est tout le mécanisme**, la relation $f(\ell) = \ell$ permet de récrire l'écart au rang suivant comme un écart d'images :

$$u_{n+1} - \ell = f(u_n) - f(\ell).$$

L'IAF donne alors

$$|u_{n+1} - \ell| = |f(u_n) - f(\ell)| \leq k |u_n - \ell| \leq k \times k^n |u_0 - \ell| = k^{n+1} |u_0 - \ell|,$$

la seconde inégalité étant l'hypothèse de récurrence, multipliée par $k \geq 0$. La propriété est héréditaire, donc vraie pour tout n .

La convergence. Comme $0 \leq k < 1$, la suite géométrique (k^n) tend vers 0 (chapitre « Suites réelles »), donc $k^n |u_0 - \ell|$ aussi. L'encadrement $0 \leq |u_n - \ell| \leq k^n |u_0 - \ell|$ et le théorème des gendarmes donnent $|u_n - \ell| \rightarrow 0$, c'est-à-dire $u_n \rightarrow \ell$.

L'unicité du point fixe. Soit ℓ' un point fixe de f dans I . L'IAF appliquée à ℓ et ℓ' donne

$$|\ell - \ell'| = |f(\ell) - f(\ell')| \leq k |\ell - \ell'|,$$

donc $(1 - k)|\ell - \ell'| \leq 0$. Comme $1 - k > 0$, il vient $|\ell - \ell'| \leq 0$, donc $\ell = \ell'$. **C'est ici, et seulement ici, que la stricte inégalité $k < 1$ est consommée dans cette partie de la preuve** : avec $k = 1$, la dernière ligne ne dirait plus rien. $\square$

14.2 La voie courte, et celle que beaucoup prennent

La démonstration ci-dessus applique l'IAF **entre u_n et ℓ** . C'est le geste à retenir, et c'est celui que la plupart des copies ne font pas : elles appliquent l'IAF entre deux termes consécutifs, ce qui donne

$$|u_{n+1} - u_n| = |f(u_n) - f(u_{n-1})| \leq k |u_n - u_{n-1}|, \quad \text{d'où} \quad |u_{n+1} - u_n| \leq k^n |u_1 - u_0|.$$

Cette majoration n'est pas fausse, et elle mène quelque part : elle prouve que la suite est de Cauchy, donc convergente. Mais il faut ensuite un argument **supplémentaire** pour identifier la limite, à savoir

la continuité de f et le passage à la limite dans la relation $u_{n+1} = f(u_n)$. Et surtout, elle ne donne aucune majoration de l'écart à ℓ : elle dit que la suite se stabilise, pas où.

La voie courte donne les trois choses d'un coup — convergence, valeur de la limite, vitesse — au prix d'une seule condition : **connaître ℓ à l'avance**. Or on le connaît presque toujours, parce qu'on l'a obtenu en résolvant l'équation $f(x) = x$ avant de commencer, ou parce que le théorème des valeurs intermédiaires en a donné l'existence. C'est pour cette raison qu'on cherche le point fixe **avant** de lancer l'étude de la suite, et jamais après.

14.3 Le rang qui garantit une précision

C'est l'apport propre de cette méthode, et il est spectaculaire. On veut $|u_n - \ell| \leq \varepsilon$ pour un $\varepsilon > 0$ donné. Il **suffit** que $k^n |u_0 - \ell| \leq \varepsilon$. Supposons $k > 0$ (si $k = 0$, la fonction est constante sur I et $u_1 = \ell$ déjà) et $u_0 \neq \ell$. En appliquant le logarithme, qui est strictement croissant, la condition devient

$$n \ln k \leq \ln \varepsilon - \ln |u_0 - \ell|.$$

Or $\ln k < 0$, puisque $0 < k < 1$. **En divisant par $\ln k$, l'inégalité se retourne**, et l'on obtient

$$n \geq \frac{\ln \varepsilon - \ln |u_0 - \ell|}{\ln k}.$$

Le retournement est l'erreur la plus fréquente de tout ce calcul, et elle est facile à détecter : si le résultat annonce qu'il faut prendre n **plus petit** qu'un certain nombre pour être plus précis, c'est qu'on a oublié de retourner.

Exemple

La suite $u_{n+1} = \cos(u_n)$. Posons $f = \cos$, $I = [0, 1]$ et $u_0 = 1$. Vérifions les trois hypothèses, dans l'ordre.

Stabilité. La fonction $\cos$ est décroissante sur $[0, 1]$, qui est contenu dans $[0, \pi]$. Donc $\cos([0, 1]) = [\cos 1, \cos 0] = [\cos 1, 1]$. Comme $0 < \cos 1$ et $1 \leq 1$, on a bien $\cos([0, 1]) \subset [0, 1]$. Et $u_0 = 1 \in I$.

Point fixe. Posons $g(x) = \cos x - x$, continue sur $[0, 1]$. On a $g(0) = 1 > 0$ et $g(1) = \cos 1 - 1 < 0$, puisque $\cos 1 < 1$. Le théorème des valeurs intermédiaires, acquis au chapitre précédent, fournit $\ell \in]0, 1[$ tel que $g(\ell) = 0$, c'est-à-dire $\cos \ell = \ell$. **On ne sait pas calculer** ℓ , et ce n'est pas nécessaire.

Contraction. Pour $t \in [0, 1]$, $|f'(t)| = |\sin t| = \sin t$ car $\sin$ est positive sur $[0, 1] \subset [0, \pi]$; et $\sin$ y est croissante, donc $\sin t \leq \sin 1$. Comme $\sin 1 \approx 0,8415$, on peut poser $k = 0,85$, majorant simple et strictement inférieur à 1.

Conclusion. Le théorème s'applique : $u_n \rightarrow \ell$ et, comme $\ell \in]0, 1[$ et $u_0 = 1$ donnent $|u_0 - \ell| \leq 1$,

$$|u_n - \ell| \leq 0,85^n.$$

Ici encore, on n'a pas eu besoin de la valeur de ℓ : il a suffi de majorer $|u_0 - \ell|$ par la **longueur de I** . C'est un réflexe à prendre.

Le rang. Pour obtenir $|u_n - \ell| \leq 10^{-6}$, il suffit que $0,85^n \leq 10^{-6}$, soit

$$n \geq \frac{\ln(10^{-6})}{\ln(0,85)} = \frac{-13,8155 \dots}{-0,16252 \dots} \approx 85,01.$$

L'entier $n = 86$ convient donc. On sait, **avant d'avoir calculé le moindre terme**, que 86 itérations garantissent six décimales.

Les premiers termes. Ils valent, arrondis à 10^{-4} :

n	0	1	2	3	4	5	6
u_n	1,0000	0,5403	0,8576	0,6543	0,7935	0,7014	0,7640

Deux lectures. D'abord, la suite **oscille** autour de sa limite, environ 0,7391 : elle n'est pas monotone, et c'était prévisible puisque f est décroissante sur I . L'argument par monotonie du chapitre « Suites réelles » est donc inapplicable tel quel, alors que l'IAF n'a rien demandé de tel. Ensuite, l'écart réel au rang 6 vaut environ 0,025, alors que la majoration garantit seulement $0,85^6 \approx 0,377$: elle est **grossière d'un facteur 15**. Cela n'a aucune importance. Une majoration garantie vaut mieux qu'une observation numérique, qui ne prouve rien.

14.4 Ce que l'inégalité apporte de plus que la monotonie

Le chapitre « Suites réelles » démontre la convergence d'une suite récurrente par l'argument classique : quand f est croissante et I stable, la suite (u_n) est monotone, son sens étant décidé par le signe de $u_1 - u_0$; bornée par I , elle converge ; et sa limite vérifie $\ell = f(\ell)$ par continuité de f . **Cet argument reste valable et reste utile** : il ne demande aucune hypothèse sur f' , seulement la croissance et la stabilité.

Mais il donne exactement une chose : l'**existence** de la limite. Il ne dit pas à quelle vitesse on s'en approche, donc il ne permet jamais de répondre à la question « combien de termes dois-je calculer pour avoir six décimales ? ». Pour cela il faudrait comparer u_n à ℓ , ce que la monotonie seule ne permet pas.

L'inégalité des accroissements finis répond à cette question, et elle y répond **à l'avance**. C'est ce qui la rend indispensable dès qu'un calcul doit être programmé : c'est elle qui justifie le critère d'arrêt d'un algorithme, et c'est sur elle que repose l'étude de la méthode de Newton, en fin de chapitre, où l'on obtiendra une convergence bien plus rapide encore.

Méthode

Les trois vérifications à écrire avant d'appliquer la méthode. Aucune ne peut être sautée, et chacune correspond à une manière précise de se tromper.

1. **Un intervalle stable, et on le démontre.** On exhibe I , puis on établit $f(I) \subset I$, le plus souvent par une étude des variations de f sur I (section 15) qui donne $f(I)$ explicitement. Écrire « I est clairement stable » ne vaut rien : la stabilité est ce qui fait exister la suite.
2. $u_0 \in I$. Une ligne, souvent absente. Sans elle, la récurrence n'a pas d'initialisation, et tout ce qui suit s'effondre.
3. **Un k constant, strictement inférieur à 1, valable sur tout I .** L'énoncé « $|f'(t)| < 1$ pour tout $t \in I$ » ne suffit **pas** : il faut un nombre k fixé qui majore. Contre-exemple à connaître : sur $I = [0, +\infty[$, la fonction $f(x) = x + e^{-x}$ vérifie $f'(x) = 1 - e^{-x}$, donc $0 \leq f'(x) < 1$ partout. Pourtant f n'a **aucun** point fixe, puisque $f(x) = x$ équivaudrait à $e^{-x} = 0$; la suite $u_{n+1} = f(u_n)$ est strictement croissante, et elle tend vers $+\infty$. L'obstacle est que $\sup_I f' = 1$, sans être atteint : aucun $k < 1$ ne majore.

Et une quatrième, plus discrète : **le point fixe doit être dans I** . Sans point fixe, le théorème n'a rien à dire, comme le contre-exemple ci-dessus le montre.

L'exercice 14 déroule cette méthode entièrement sur un exemple, puis fait varier l'intervalle de départ pour montrer que le choix de I n'est pas neutre : mal choisi, il ne donne aucun $k < 1$, alors qu'un intervalle plus petit et bien placé en donne un.

15. Variations et signe de la dérivée

Tout le lycée admet le lien entre le signe de la dérivée et les variations d'une fonction. Le voici démontré, et l'outil est l'égalité des accroissements finis. C'est le moment où l'EAF paie ce qu'elle a coûté.

15.1 Les énoncés

Théorème

Variations et signe de la dérivée. Soit I un intervalle de $\mathbb{R}$ et f une fonction dérivable sur I .

1. f est croissante sur $I \iff f'(t) \geq 0$ pour tout $t \in I$.
2. f est décroissante sur $I \iff f'(t) \leq 0$ pour tout $t \in I$.
3. f est constante sur $I \iff f'(t) = 0$ pour tout $t \in I$.
4. Si $f'(t) > 0$ pour tout $t \in I$, alors f est **strictement** croissante sur I . La réciproque est fausse.
5. Si $f'(t) \geq 0$ pour tout $t \in I$ et si f' ne s'annule qu'en un nombre **fini** de points de I , alors f est strictement croissante sur I .

Démonstration du point 1, sens direct de la droite vers la gauche. Supposons $f' \geq 0$ sur I , et soient $x < y$ dans I . Comme I est un intervalle, le segment $[x, y]$ est contenu dans I ; f y est dérivable, donc continue. L'égalité des accroissements finis fournit $c \in]x, y[$ tel que

$$f(y) - f(x) = f'(c)(y - x).$$

Le point c appartient à I , donc $f'(c) \geq 0$; et $y - x > 0$. Le produit est donc positif ou nul : $f(x) \leq f(y)$. La fonction f est croissante sur I . $\square$

Réciproque du point 1. Supposons f croissante sur I et soit $a \in I$. Pour $x \in I$ distinct de a , le taux d'accroissement $\frac{f(x) - f(a)}{x - a}$ est positif ou nul : si $x > a$, numérateur et dénominateur sont tous deux

positifs ; si $x < a$, ils sont tous deux négatifs. Le passage à la limite dans une inégalité large donne $f'(a) \geq 0$. $\square$

Notons que cette réciproque n'emploie **pas** l'EAF : elle ne demande que la définition de la dérivée. L'asymétrie est instructive. Le sens qui coûte cher, celui qui a besoin de tout l'arsenal, est celui qui va de la dérivée vers la fonction, c'est-à-dire du local vers le global — et c'est précisément le sens qu'on utilise.

Points 2 et 3. Le point 2 s'obtient en appliquant le point 1 à $-f$, dont la dérivée est $-f'$, et en remarquant que $-f$ est croissante si et seulement si f est décroissante. Le point 3 s'obtient en combinant les deux : si $f' = 0$, alors $f' \geq 0$ et $f' \leq 0$, donc f est à la fois croissante et décroissante, donc constante. La réciproque est immédiate, la dérivée d'une fonction constante étant nulle. $\square$

Point 4. La même démonstration qu'au point 1, à un mot près : $f'(c) > 0$ et $y - x > 0$ donnent un produit **strictement** positif, donc $f(x) < f(y)$. $\square$

La réciproque du point 4 est fautive, et le contre-exemple est celui qu'il faut savoir citer sans réfléchir : $x \mapsto x^3$ est strictement croissante sur $\mathbb{R}$ (démontré à la main au chapitre précédent), alors que sa dérivée $3x^2$ s'annule en 0. Une dérivée qui s'annule en un point isolé n'empêche donc pas la stricte croissance ; c'est ce que le point 5 précise.

Point 5. La fonction f est croissante sur I d'après le point 1. Supposons par l'absurde qu'elle ne soit pas strictement croissante : il existe alors $x < y$ dans I tels que $f(x) = f(y)$. Soit $t \in [x, y]$. La croissance donne $f(x) \leq f(t) \leq f(y) = f(x)$, donc $f(t) = f(x)$: la fonction f est **constante** sur le segment $[x, y]$. Par le point 3 appliqué à l'intervalle $[x, y]$, sa dérivée y est nulle, et en particulier f' s'annule en tout point de $]x, y[$, qui est un intervalle ouvert non vide, donc infini. Cela contredit l'hypothèse. La fonction f est donc strictement croissante sur I . $\square$

Cette démonstration montre au passage ce qu'est la vraie hypothèse : ce n'est pas « un nombre fini de points », c'est « **l'ensemble des zéros de f' ne contient aucun intervalle non trivial** ». Un nombre fini de points est simplement le cas le plus courant, et le plus facile à vérifier.

15.2 Le piège de l'intervalle

Aucun des cinq énoncés n'est vrai sur un ensemble qui n'est pas un intervalle, et cette phrase mérite d'être encadrée mentalement, parce qu'elle est violée dans presque toutes les copies au moins une fois.

Le contre-exemple, pour les points 1 et 2. Soit $f(x) = \frac{1}{x}$, définie sur $D = \mathbb{R}^*$, réunion des deux intervalles $] -\infty, 0[$ et $] 0, +\infty[$. Sa dérivée vaut $f'(x) = -\frac{1}{x^2}$, qui est **strictement négative en tout point de D** . Pourtant f n'est pas décroissante sur D : on a $-1 < 1$ alors que

$$f(-1) = -1 \quad \text{et} \quad f(1) = 1, \quad \text{donc} \quad f(-1) < f(1).$$

La dérivée est strictement négative partout, et la fonction augmente entre deux points de son domaine.

Où la démonstration casse, exactement. Reprenons le point 1 avec $x = -1$ et $y = 1$. La première phrase de la preuve dit : « comme I est un intervalle, le segment $[x, y]$ est contenu dans I ». Ici, le segment $[-1, 1]$ contient 0, qui n'appartient pas à D . La fonction f n'est donc ni définie ni continue sur tout $[-1, 1]$, l'hypothèse de l'EAF n'est pas satisfaite, et le théorème ne s'applique pas. **L'hypothèse « intervalle » n'est consommée qu'une fois dans toute la démonstration, mais elle l'est de manière irremplaçable.**

Le contre-exemple, pour le point 3. Il est encore plus net. Soit g définie sur $D = \mathbb{R}^*$ par $g(x) = 0$ si $x < 0$ et $g(x) = 1$ si $x > 0$. Alors g est dérivable sur D et $g' = 0$ partout, sans que g soit constante. Il n'y a aucune contradiction avec le théorème : celui-ci affirme que g est constante sur **chaque intervalle** contenu dans D , et c'est exactement ce qui se passe.

Ce point a une conséquence qui dépasse ce chapitre. « Deux fonctions de même dérivée sur un intervalle différent d'une constante » est un énoncé vrai sur un intervalle, faux sur une réunion d'intervalles disjoints — les fonctions g ci-dessus et la fonction nulle ont la même dérivée sur $\mathbb{R}^*$ sans différer d'une constante. C'est ce qui commandera, dans le chapitre « Calcul intégral élémentaire », la manière dont on écrit les primitives d'une fonction dont le domaine n'est pas un intervalle. L'exercice 16 est entièrement consacré à cette question.

15.3 Démontrer une inégalité

Méthode

Démontrer une inégalité par étude de fonction. Cinq étapes, et une seule est vraiment discriminante.

1. **Poser $h = (\text{membre de droite}) - (\text{membre de gauche})$** , et annoncer l'objectif : montrer que $h \geq 0$ sur l'intervalle considéré. Le sens du choix se retient ainsi : pour prouver « petit $\leq$ grand », on retranche le petit au grand, de sorte que la conclusion visée soit « $h \geq 0$ ».
2. **Préciser l'intervalle I sur lequel on travaille** et justifier que h y est dérivable, par les théorèmes d'opérations.
3. **Calculer h' et le simplifier avant d'en étudier le signe.** Une expression mise sur dénominateur commun et factorisée se lit ; une somme de trois quotients ne se lit pas. Si le signe de h' reste illisible, on dérive une seconde fois et l'on recommence sur h' : c'est la cascade, et elle s'arrête toujours, faute de quoi l'inégalité relève d'un autre outil.
4. **Dresser le tableau de variations de h , puis calculer la valeur de h au point où le minimum est atteint.** C'est cette valeur, et elle seule, qui conclut. Une copie qui écrit « h est croissante, donc $h \geq 0$ » ne démontre rien : croissante à partir de quelle valeur ?
5. **Conclure en revenant à l'inégalité de départ**, et traiter le cas d'égalité s'il est demandé : il correspond aux points où h s'annule, que le tableau désigne directement.

Exemple

Montrons que $\ln x \leq x - 1$ pour tout $x > 0$.

Étape 1. Posons $h(x) = (x - 1) - \ln x$ sur $I =]0, +\infty[$. L'inégalité à démontrer équivaut à $h(x) \geq 0$ sur I .

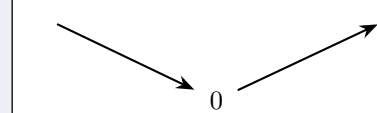
Étape 2. La fonction h est dérivable sur I comme différence d'une fonction affine et du logarithme, dérivable sur $]0, +\infty[$.

Étape 3. Pour $x > 0$,

$$h'(x) = 1 - \frac{1}{x} = \frac{x-1}{x}.$$

Le dénominateur x est strictement positif sur I : le signe de $h'(x)$ est donc celui de $x - 1$. Il est strictement négatif sur $]0, 1[$, nul en 1, strictement positif sur $]1, +\infty[$.

Étape 4. D'où le tableau de variations :

x	0	1	$+\infty$	
$h'(x)$		−	0	+
$h(x)$				

La fonction h atteint donc son minimum sur I au point 1, et ce minimum vaut

$$h(1) = (1 - 1) - \ln 1 = 0.$$

Étape 5. Pour tout $x > 0$, on a donc $h(x) \geq h(1) = 0$, c'est-à-dire $\ln x \leq x - 1$. De plus h' ne s'annule qu'en 1 : le point 5 du théorème donne la stricte décroissance sur $]0, 1]$ et la stricte croissance sur $[1, +\infty[$, donc $h(x) > 0$ pour $x \neq 1$. **L'égalité a lieu si et seulement si $x = 1$.**
□

Cette inégalité était posée comme un acquis au chapitre précédent ; elle est ici démontrée, et sa démonstration tient en six lignes.

Remarque

Le choix de l'outil, et c'est là que se gagne du temps. Trois routes mènent à une inégalité, et savoir les distinguer vaut mieux que savoir les parcourir.

L'étude de fonction, ci-dessus. Elle marche toujours, elle ne demande rien, et elle coûte un tableau complet. C'est le choix par défaut, et c'est le bon quand l'inégalité n'a pas de structure particulière.

L'inégalité des accroissements finis (section 13). Elle s'impose quand l'inégalité compare un **accroissement** $f(x) - f(y)$ à un multiple de $x - y$. Ainsi $|\sin x - \sin y| \leq |x - y|$ se démontre en une ligne par l'IAF, alors qu'une étude de fonction obligerait à fixer y et à étudier une fonction de x dépendant d'un paramètre : beaucoup plus long, pour le même résultat.

La convexité (section 24). Elle s'impose quand l'inégalité compare une fonction à une **tangente** ou à une **corde**. L'inégalité $e^x \geq 1 + x$ dit exactement que la courbe de $\exp$ est au-dessus de sa tangente au point d'abscisse 0 ; l'inégalité $\ln x \leq x - 1$ démontrée ci-dessus dit que la courbe de $\ln$ est au-dessous de sa tangente au point d'abscisse 1. Une fois la convexité acquise, ces deux inégalités deviendront des cas particuliers immédiats, sans le moindre calcul.

Le signe qui trahit la structure : si l'inégalité est de la forme « $f(x) \geq$ une expression affine » et que l'égalité a lieu en **un seul point**, c'est presque toujours une histoire de tangente, donc de convexité.

Quatre exercices entraînent cette section. L'exercice 15 propose quatre inégalités et demande explicitement, pour chacune, quel outil choisir et pourquoi. L'exercice 10 combine la stricte monotonie et le théorème des valeurs intermédiaires pour **compter** les solutions d'une équation, en s'appuyant sur le théorème de Rolle pour majorer leur nombre. L'exercice 16 travaille le piège de l'intervalle pour lui-même. L'exercice 18 mène une étude complète de fonction, du domaine au tracé.

16 16. Les accroissements finis généralisés, et la règle de L'Hôpital

Remarque

Selon les facs. Le théorème des accroissements finis généralisés, dit théorème de Cauchy, et la règle de L'Hôpital qui en découle figurent au premier semestre dans certaines maquettes, où ils servent de chemin standard vers la formule de Taylor-Lagrange ; ailleurs, ils sont repoussés au moment des développements limités, qui les rendent souvent inutiles. Ils sont traités ici, parce qu'un étudiant les rencontrera de toute façon dans un exercice ou dans la bouche d'un camarade, et parce que la fonction auxiliaire du théorème de Cauchy est le meilleur exercice de compréhension de la section 12.

16.1 Le théorème de Cauchy

Théorème

Théorème des accroissements finis généralisés (théorème de Cauchy). Soient a et b deux réels tels que $a < b$, et f et g deux fonctions continues sur $[a, b]$ et dérivables sur $]a, b[$. On suppose que g' ne s'annule en aucun point de $]a, b[$.

Alors $g(a) \neq g(b)$, et il existe $c \in]a, b[$ tel que

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Le quotient de gauche a un sens, et c'est la première chose à établir. Si l'on avait $g(a) = g(b)$, la fonction g vérifierait les trois hypothèses du théorème de Rolle sur $[a, b]$, qui fournirait un point de $]a, b[$ où g' s'annule : cela contredit l'hypothèse. Donc $g(b) - g(a) \neq 0$. Cette remarque, souvent sautée,

est précisément l'endroit où l'hypothèse « g' ne s'annule pas » sert une première fois.

D'où vient la fonction auxiliaire. Le raisonnement de la section 12 se reproduit, avec une différence : il n'y a plus de corde à retrancher, il y a deux fonctions. On cherche donc une combinaison de la forme

$$\psi = f - \lambda g,$$

où λ est un réel qu'on va **choisir** pour que la condition de Rolle, $\psi(a) = \psi(b)$, soit satisfaite. Cette condition s'écrit $f(a) - \lambda g(a) = f(b) - \lambda g(b)$, c'est-à-dire

$$\lambda(g(b) - g(a)) = f(b) - f(a), \quad \text{donc} \quad \lambda = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Le coefficient λ n'est donc pas un lapin sorti d'un chapeau : c'est le **seul** choix qui rend Rolle applicable, et il vaut exactement le membre de gauche de la conclusion visée. Une fonction auxiliaire ne se devine jamais ; elle se déduit de la condition qu'on veut satisfaire.

Démonstration. On note $\lambda = \frac{f(b) - f(a)}{g(b) - g(a)}$, qui a un sens d'après ce qui précède, et l'on pose, pour $x \in [a, b]$,

$$\psi(x) = f(x) - \lambda g(x).$$

On l'appelle ψ , et non φ , pour ne pas la confondre avec la fonction auxiliaire de la section 12, qui est une autre fonction.

Continuité et dérivabilité. Comme f et g sont continues sur $[a, b]$ et dérivables sur $]a, b[$, il en est de même de toute combinaison linéaire de f et g , en particulier de ψ , et $\psi'(x) = f'(x) - \lambda g'(x)$ pour $x \in]a, b[$.

Valeurs aux bornes. On calcule la différence, ce qui est plus court que les deux valeurs :

$$\psi(b) - \psi(a) = (f(b) - f(a)) - \lambda(g(b) - g(a)) = (f(b) - f(a)) - (f(b) - f(a)) = 0,$$

par définition de λ . Donc $\psi(a) = \psi(b)$.

Rolle. Le théorème de Rolle fournit $c \in]a, b[$ tel que $\psi'(c) = 0$, c'est-à-dire $f'(c) = \lambda g'(c)$. Comme $g'(c) \neq 0$ par hypothèse — c'est le second usage de cette hypothèse — on peut diviser :

$$\frac{f'(c)}{g'(c)} = \lambda = \frac{f(b) - f(a)}{g(b) - g(a)}. \quad \square$$

Propriété

L'EAF est le cas $g(x) = x$. Prenons pour g la fonction identité sur $[a, b]$: elle est continue, dérivable, et $g'(x) = 1$ ne s'annule jamais. On a $g(b) - g(a) = b - a$, et la conclusion du théorème de Cauchy devient

$$\frac{f(b) - f(a)}{b - a} = \frac{f'(c)}{1} = f'(c),$$

qui est exactement l'égalité des accroissements finis.

On a donc une chaîne d'énoncés emboîtés : Cauchy contient l'EAF, qui contient Rolle. Et comme Rolle a servi à démontrer les deux autres, les trois sont en fait **équivalents** : chacun se déduit des autres au prix d'une fonction auxiliaire. C'est Rolle qu'on démontre à la main, parce que c'est le seul dont la preuve soit directe, à partir des bornes atteintes et du lemme de Fermat.

Remarque

La fausse démonstration, celle qu'il faut savoir refuser. Il est tentant d'appliquer l'EAF séparément à f et à g , puis de diviser. On obtient alors

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c_1)(b - a)}{g'(c_2)(b - a)} = \frac{f'(c_1)}{g'(c_2)},$$

ce qui n'est pas la conclusion voulue : il y a **deux** points c_1 et c_2 , et rien ne dit qu'ils sont égaux. Tout le contenu du théorème de Cauchy est justement qu'**un seul et même point convient pour les deux fonctions**, et c'est la fonction auxiliaire ψ qui le produit. Cette erreur est le meilleur test de compréhension de l'énoncé.

16.2 La règle de L'Hôpital**Théorème**

Règle de L'Hôpital, forme $\frac{0}{0}$. Soit I un intervalle, a un point de I ou une borne de I , et f , g deux fonctions dérivables sur $I \setminus \{a\}$, avec g' ne s'annulant pas sur $I \setminus \{a\}$. On suppose :

1. $f(x) \rightarrow 0$ et $g(x) \rightarrow 0$ quand $x \rightarrow a$;
2. $\frac{f'(x)}{g'(x)} \rightarrow L$ quand $x \rightarrow a$, où L est un réel, ou $+\infty$, ou $-\infty$.

Alors g ne s'annule pas sur un voisinage épointé de a , et

$$\frac{f(x)}{g(x)} \xrightarrow{x \rightarrow a} L.$$

L'ordre des hypothèses est l'énoncé, et c'est là que les étudiants se blessent. Le théorème part de la limite du quotient des **dérivées**, et conclut sur la limite du quotient des **fonctions**. Il autorise donc à remplacer $\frac{f}{g}$ par $\frac{f'}{g'}$ **seulement si** cette seconde limite existe. Écrire d'emblée une égalité entre les deux limites, avant d'avoir montré que celle de droite existe, revient à évaluer un objet inconnu et un objet dont on ignore s'il existe. La rédaction correcte calcule d'abord $\lim \frac{f'}{g'}$, constate qu'elle existe, et **en déduit** la limite cherchée.

Démonstration, dans le cas où a est un réel. Quitte à poser $f(a) = 0$ et $g(a) = 0$, l'hypothèse 1 rend f et g continues en a , donc continues sur I . Soit $x \in I$ avec $x > a$. Sur le segment $[a, x]$, les fonctions f et g sont continues, dérivables sur $]a, x[$, et g' n'y annule pas : le théorème de Cauchy fournit $c_x \in]a, x[$ tel que

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

Quand $x \rightarrow a^+$, l'encadrement $a < c_x < x$ donne $c_x \rightarrow a$ par le théorème des gendarmes, et surtout $c_x \neq a$: c'est exactement la condition qui rend légitime le passage à la limite dans une composée, et le chapitre précédent a montré que l'oublier est le piège classique. L'hypothèse 2 donne alors $\frac{f'(c_x)}{g'(c_x)} \rightarrow$

L , donc $\frac{f(x)}{g(x)} \rightarrow L$ quand $x \rightarrow a^+$. Le raisonnement à gauche de a est identique, et les deux limites latérales coïncident. $\square$

Remarque

La règle possède des variantes que l'on emploie de la même manière : forme $\frac{\infty}{\infty}$, et limites en $+\infty$ ou en $-\infty$. Leur démonstration demande un travail supplémentaire, et on les **admet** ici. Les hypothèses et les précautions sont exactement les mêmes.

Exemple**Deux applications.**

Première. Calculons $\lim_{x \rightarrow 0} \frac{e^x - 1 - x}{x^2}$. Numérateur et dénominateur tendent vers 0 : la forme est indéterminée. Le quotient des dérivées vaut $\frac{e^x - 1}{2x}$, dont le numérateur et le dénominateur tendent encore vers 0 : nouvelle forme indéterminée, et l'on **re-vérifie** les hypothèses avant de recommencer. Le quotient des dérivées secondes vaut $\frac{e^x}{2}$, qui tend vers $\frac{1}{2}$.

On **remonte** alors : puisque $\frac{e^x}{2} \rightarrow \frac{1}{2}$, la règle donne $\frac{e^x - 1}{2x} \rightarrow \frac{1}{2}$; puisque cette limite existe, la règle donne à nouveau $\frac{e^x - 1 - x}{x^2} \rightarrow \frac{1}{2}$. Le sens de lecture est celui-là, et pas l'autre : on descend au brouillon, on remonte sur la copie.

Seconde. Calculons $\lim_{x \rightarrow 1} \frac{\ln x}{x^2 - 1}$. La forme est indéterminée en 1. Les dérivées donnent

$$\frac{1/x}{2x} = \frac{1}{2x^2} \xrightarrow{x \rightarrow 1} \frac{1}{2},$$

et le dénominateur dérivé $2x$ ne s'annule pas au voisinage épointé de 1. La limite cherchée vaut donc $\frac{1}{2}$.

16.3 Les deux cas d'échec

Ils sont obligatoires, au sens où un étudiant qui ne les connaît pas emploiera la règle à tort tôt ou tard.

Échec (a) : $\frac{f'}{g'}$ **n'a pas de limite, alors que** $\frac{f}{g}$ **en a une.** Prenons, sur $I = \mathbb{R}$ et en $a = 0$,

$$f(x) = x^2 \sin\left(\frac{1}{x}\right) \text{ pour } x \neq 0, \quad f(0) = 0, \quad \text{et} \quad g(x) = x.$$

Les deux tendent vers 0 en 0, et $g'(x) = 1$ ne s'annule jamais : les hypothèses de départ sont réunies. Le quotient des fonctions vaut $\frac{f(x)}{g(x)} = x \sin \frac{1}{x}$, dont la valeur absolue est majorée par $|x|$: il tend vers 0.

Mais le quotient des dérivées vaut

$$\frac{f'(x)}{g'(x)} = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right),$$

et il **n'a pas de limite en 0**. Pour le voir, on emploie la caractérisation séquentielle du chapitre précédent avec deux suites tendant vers 0 : en $x_n = \frac{1}{2n\pi}$, le sinus vaut 0 et le cosinus vaut 1, donc le quotient vaut -1 ; en $y_n = \frac{1}{(2n+1)\pi}$, le sinus vaut 0 et le cosinus vaut -1 , donc le quotient vaut 1. Deux suites de points tendant vers 0 donnent deux limites différentes : il n'y a pas de limite.

Ce que cet exemple démontre, c'est que la règle de L'Hôpital est une implication à **sens unique**. L'absence de limite pour $\frac{f'}{g'}$ n'autorise aucune conclusion sur $\frac{f}{g}$: ni qu'elle existe, ni qu'elle n'existe pas. Une copie qui écrit « le quotient des dérivées n'a pas de limite, donc le quotient de départ non plus » commet une faute de logique, pas une faute de calcul. Cette fonction $x^2 \sin(1/x)$ est étudiée en détail à l'exercice 04, avec toute la famille $x^\alpha \sin(1/x)$.

Échec (b) : l'application circulaire. Cherchons $\lim_{x \rightarrow +\infty} \frac{\sqrt{x^2+1}}{x}$, forme indéterminée $\frac{\infty}{\infty}$. Le quotient des dérivées vaut

$$\frac{x/\sqrt{x^2+1}}{1} = \frac{x}{\sqrt{x^2+1}},$$

c'est-à-dire l'**inverse** du quotient de départ. Une seconde application ramène au quotient initial, et ainsi de suite : la règle tourne en rond sans jamais conclure, et l'on peut l'appliquer indéfiniment sans progresser d'un pas.

La sortie est toujours la même : revenir au calcul direct. Pour $x > 0$,

$$\frac{\sqrt{x^2+1}}{x} = \frac{\sqrt{x^2+1}}{\sqrt{x^2}} = \sqrt{1 + \frac{1}{x^2}} \xrightarrow{x \rightarrow +\infty} 1.$$

Une factorisation par le terme dominant a suffi. La leçon est qu'une règle qui ne conclut pas au bout de deux applications est probablement le mauvais outil.

Méthode

Avertissement de rédaction. Quatre points, dans l'ordre où ils coûtent des points.

1. **Vérifier la forme indéterminée AVANT d'appliquer la règle**, et l'écrire. Appliquée sans réfléchir à $\frac{x+1}{x+2}$ en 0, la règle donnerait $\frac{1}{1} = 1$, alors que la limite vaut $\frac{1}{2}$. C'est l'erreur la plus fréquente, et la plus facile à éviter.
2. **Dériver le numérateur et le dénominateur séparément**, jamais le quotient. Le quotient des dérivées et la dérivée du quotient sont deux choses différentes, et la confusion se voit immédiatement à la présence d'un carré au dénominateur.
3. **Vérifier que g' ne s'annule pas** au voisinage épointé du point étudié, et le dire.
4. **La règle ne dispense jamais de réfléchir à l'outil.** Un développement limité, objet des sections 20 à 23, donne la même limite, souvent en moins de lignes, sans aucune hypothèse à vérifier, et il fournit **en plus** le terme suivant, donc la vitesse à laquelle la limite est atteinte. Sur le premier exemple ci-dessus, le développement limité de e^x à l'ordre 2 donne le résultat d'un trait. Quand les deux méthodes sont possibles, le développement limité est le choix sûr en partiel.

L'exercice 11 traite le théorème de Cauchy et la règle de L'Hôpital ensemble : on y démontre le théorème sur des cas particuliers, on y calcule des limites des deux manières, et l'on y reconstruit les deux cas d'échec.

17 17. Le théorème de la limite de la dérivée, et le théorème de Darboux

Remarque

Selon les faits. Ces deux théorèmes sont enseignés systématiquement dans certaines universités, notamment là où le cours d'analyse du premier semestre traite en détail les fonctions définies par morceaux ; ailleurs, ils apparaissent seulement en exercice, ou au second semestre. Le premier est d'un usage constant dès qu'on raccorde deux formules ; le second est plus culturel, mais il justifie une phrase qu'on écrit sans y penser, et c'est une bonne raison de le connaître.

17.1 Le théorème de la limite de la dérivée**Théorème**

Théorème de la limite de la dérivée. Soit I un intervalle, a un point de I , et f une fonction définie sur I telle que :

1. f est **continue en** a ;
2. f est **dérivable sur** $I \setminus \{a\}$;
3. $f'(x)$ tend vers un réel ℓ quand x tend vers a .

Alors f est **dérivable en** a , avec $f'(a) = \ell$, et f' est continue en a . En particulier, si f est de classe $\mathcal{C}^1$ sur $I \setminus \{a\}$, elle est de classe $\mathcal{C}^1$ sur I tout entier.

Démonstration par l'EAF. Elle est immédiate une fois l'égalité des accroissements finis acquise, et il vaut la peine de voir à quel point.

Notons d'abord que f est continue sur I : elle l'est en a par l'hypothèse 1, et en tout autre point par l'hypothèse 2, la dérivabilité entraînant la continuité.

Soit $x \in I$ avec $x \neq a$, et notons J_x le segment d'extrémités a et x . Comme I est un intervalle, $J_x \subset I$. La fonction f est continue sur J_x , et elle est dérivable sur l'intervalle **ouvert** d'extrémités a et x , qui ne contient pas a : les deux hypothèses de l'EAF sont donc satisfaites, malgré le fait qu'on ne suppose rien de la dérivabilité en a . C'est très exactement ce que la séparation des hypothèses de la section 12 autorise.

L'EAF fournit donc un réel c_x , strictement compris entre a et x , tel que

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

Quand $x \rightarrow a$, l'encadrement $|c_x - a| < |x - a|$ donne $c_x \rightarrow a$; et $c_x \neq a$, ce qui est la condition qui rend le passage à la limite légitime dans la composée. L'hypothèse 3 donne alors $f'(c_x) \rightarrow \ell$.

Le taux d'accroissement de f en a tend donc vers ℓ : par définition, f est dérivable en a et $f'(a) = \ell$. Enfin, l'hypothèse 3 se relit maintenant $f'(x) \rightarrow f'(a)$ quand $x \rightarrow a$, ce qui est la continuité de f' en a . $\square$

Exemple

Un prolongement. Soit f définie sur $\mathbb{R}$ par $f(x) = x^2 \ln|x|$ pour $x \neq 0$, et $f(0) = 0$.

Continuité en 0. Par croissances comparées (chapitre précédent), $x^2 \ln|x| \rightarrow 0$ quand $x \rightarrow 0$, donc f est continue en 0.

Dérivabilité hors de 0. Sur $\mathbb{R}^*$, f est de classe $\mathcal{C}^\infty$ comme produit de fonctions qui le sont, et

$$f'(x) = 2x \ln|x| + x.$$

Limite de la dérivée. Toujours par croissances comparées, $f'(x) \rightarrow 0$ quand $x \rightarrow 0$.

Conclusion. Le théorème donne : f est dérivable en 0, $f'(0) = 0$, et f est de classe $\mathcal{C}^1$ sur $\mathbb{R}$.

Ce que le théorème a réellement apporté. Ici, la voie directe fonctionne aussi : le taux d'accroissement en 0 vaut $\frac{x^2 \ln|x|}{x} = x \ln|x|$, qui tend vers 0. Elle donne la dérivabilité en 0 et la valeur $f'(0) = 0$ — mais **pas** la classe $\mathcal{C}^1$: il faudrait encore vérifier séparément que f' est continue en 0. Le théorème donne les deux d'un seul coup, et c'est là son vrai gain. Il gagne également, en volume de calcul, chaque fois que f' est simple à calculer alors que le taux d'accroissement demande une factorisation pénible, ce qui est fréquent pour un quotient ou une composée.

Remarque

Le cas où ℓ est infini. Si $f'(x) \rightarrow +\infty$ quand $x \rightarrow a$, le même calcul montre que le taux d'accroissement de f en a tend vers $+\infty$. La conclusion change alors de nature : f n'est **pas** dérivable en a , et sa courbe y admet une **tangente verticale**, situation décrite à la section 5.

Le témoin est $x \mapsto \sqrt{x}$ en 0, dont la dérivée $\frac{1}{2\sqrt{x}}$ tend vers $+\infty$. Le théorème donne donc une information dans ce cas aussi, mais ce n'est plus la dérivabilité : c'est la forme de la courbe.

La réciproque est fausse. Une fonction peut être dérivable en a sans que f' admette la moindre limite en a . Le témoin est la fonction déjà rencontrée à la section 16, définie par $f(x) = x^2 \sin(1/x)$ pour $x \neq 0$ et $f(0) = 0$: son taux d'accroissement en 0 vaut $x \sin(1/x)$, majoré en valeur absolue par $|x|$, donc elle est dérivable en 0 avec $f'(0) = 0$; et sa dérivée $2x \sin(1/x) - \cos(1/x)$ n'a pas de limite en 0, comme on l'a vu. **L'exercice 04 construit cet exemple en entier**, avec toute la famille $x^\alpha \sin(1/x)$ et le rôle que joue chaque valeur de α ; on ne le détaille pas ici.

Méthode

Le bon usage : une condition suffisante, jamais nécessaire.

- **Ce que le théorème autorise** : « f' tend vers ℓ fini en a , donc f est dérivable en a et $f'(a) = \ell$ ». C'est la phrase à écrire, avec les trois hypothèses vérifiées avant.
- **Ce qu'il n'autorise pas** : « f' n'a pas de limite en a , donc f n'est pas dérivable en a ». Ce serait la contraposée d'une implication qui n'existe pas, et la fonction $x^2 \sin(1/x)$ la réfute.
- **Que faire quand f' n'a pas de limite ?** Revenir au taux d'accroissement, c'est-à-dire à la définition de la section 2. C'est toujours possible, et c'est toujours la méthode qui tranche. Le théorème est un raccourci, pas un substitut à la définition.
- **Ne pas oublier l'hypothèse de continuité en a** , qui ne se déduit pas des deux autres. Sans elle, l'énoncé est faux : la fonction valant 0 sur $]-\infty, 0[$ et 1 sur $[0, +\infty[$ a une dérivée nulle sur $\mathbb{R}^*$, donc $f'(x) \rightarrow 0$ quand $x \rightarrow 0$, alors qu'elle n'est même pas continue en 0, donc encore moins dérivable.

L'exercice 17 est consacré à ce théorème : on y traite plusieurs raccords, on y détermine la classe exacte de la fonction obtenue, et l'on y rencontre au moins un cas où le théorème ne conclut pas et

où il faut revenir à la définition.

17.2 Le théorème de Darboux

Théorème

Théorème de Darboux (admis). Soit I un intervalle et f une fonction **dérivable sur** I . Alors f' vérifie la propriété des valeurs intermédiaires sur I : pour tous a et b dans I avec $a < b$, et pour tout réel y compris entre $f'(a)$ et $f'(b)$, il existe $c \in [a, b]$ tel que $f'(c) = y$.

Ce théorème est admis dans ce cours. Sa démonstration, courte mais astucieuse, consiste à appliquer le lemme de Fermat de la section 10 à la fonction $x \mapsto f(x) - yx$ sur $[a, b]$, en montrant que ses hypothèses forcent son extremum à être intérieur.

Ce qu'il dit, et pourquoi c'est surprenant. Le théorème des valeurs intermédiaires du chapitre précédent suppose la **continuité**. Darboux affirme qu'une dérivée possède la même propriété **sans qu'on suppose rien de sa continuité**. Et l'hypothèse manque vraiment : on vient de voir que la dérivée de $x^2 \sin(1/x)$ n'est pas continue en 0. Une dérivée peut donc être discontinue — mais alors, sa discontinuité est nécessairement d'un type sauvage, du genre oscillation, jamais un simple saut.

Ce que cela interdit. Reprenons la fonction signe, notée sgn , valant -1 sur $]-\infty, 0[$, 0 en 0 et 1 sur $]0, +\infty[$. Elle prend les valeurs -1 et 1 ; le réel $\frac{1}{2}$ est compris entre les deux, et pourtant sgn ne le prend jamais. Elle viole donc la propriété des valeurs intermédiaires sur tout intervalle contenant 0 , et le théorème de Darboux permet d'en conclure quelque chose de fort :

aucune fonction dérivable sur un intervalle contenant 0 n'a la fonction signe pour dérivée.

C'est un énoncé de non-existence, et il n'est démontrable par aucun calcul : seul un théorème sur la classe des dérivées peut le donner. Autrement dit, toutes les fonctions ne sont pas des dérivées, et la classe des dérivées est strictement plus petite que celle des fonctions quelconques.

Une conséquence qu'on utilise sans le savoir. Soit f dérivable sur un intervalle I , avec f' ne s'annulant en aucun point de I . Alors f' garde un **signe constant** sur I : si elle prenait une valeur strictement négative et une valeur strictement positive, Darboux lui ferait prendre la valeur 0 entre les deux, ce qui est exclu. Donc f est strictement monotone sur I , par le point 4 de la section 15.

C'est exactement la phrase « f' ne s'annule pas, donc f est strictement monotone », qu'on écrit couramment. Quand f' est continue, elle se justifie par le théorème des valeurs intermédiaires appliqué à f' . Quand on ne suppose pas f' continue — et rien ne l'impose ici — c'est Darboux qui la rend légitime, et rien d'autre.

18 18. Dérivées successives, classes de régularité et formule de Leibniz

Toute la suite du chapitre — formules de Taylor, développements limités, convexité — repose sur des dérivées d'ordre supérieur à un. Cette section pose le vocabulaire et l'outil de calcul.

18.1 Définitions

Définition

Dérivées successives. Soit I un intervalle et f une fonction définie sur I . On définit, quand cela a un sens, les **dérivées successives** de f par récurrence :

- $f^{(0)} = f$;
- si $f^{(p)}$ est définie et dérivable sur I , on pose $f^{(p+1)} = (f^{(p)})'$.

On dit que f est **p fois dérivable sur I** lorsque $f^{(p)}$ est définie sur I .

Les premières dérivées se notent f' , f'' , f''' ; au-delà, on écrit $f^{(4)}$, $f^{(5)}$, et l'on n'écrit jamais f'''' .

Remarque

Un piège de notation qui coûte des points. Les deux écritures $f^{(p)}$ et f^p ne désignent pas la même chose : la première, **avec les parenthèses**, est la dérivée p -ième ; la seconde, sans parenthèses, est la puissance p -ième, c'est-à-dire la fonction $x \mapsto (f(x))^p$. Pour $f(x) = x$, on a $f^{(2)} = 0$ et $f^2(x) = x^2$: les deux objets n'ont rien de commun. Une copie qui omet les parenthèses écrit autre chose que ce qu'elle veut dire, et le correcteur n'a aucun moyen de deviner.

Définition

Classes de régularité. Soit I un intervalle, f définie sur I et $p \in \mathbb{N}$.

- f est de **classe $\mathcal{C}^0$ sur I** lorsqu'elle est continue sur I .
- f est de **classe $\mathcal{C}^p$ sur I** lorsqu'elle est p fois dérivable sur I et que $f^{(p)}$ est continue sur I .
- f est de **classe $\mathcal{C}^\infty$ sur I** , ou **indéfiniment dérivable**, lorsqu'elle est de classe $\mathcal{C}^p$ pour tout entier p .

On note $\mathcal{C}^p(I)$ l'ensemble des fonctions de classe $\mathcal{C}^p$ sur I .

La lettre est muette : on écrit indifféremment $\mathcal{C}^k$, $\mathcal{C}^p$ ou $\mathcal{C}^n$. Dans cette section, on emploie p , afin de garder les lettres n et k disponibles pour la formule de Leibniz, où elles ont un autre rôle.

Propriété

La chaîne des inclusions. Pour tout $p \in \mathbb{N}$ et tout intervalle I :

$$\mathcal{C}^{p+1}(I) \subset \{f \text{ } (p+1) \text{ fois dérivable sur } I\} \subset \mathcal{C}^p(I) \subset \{f \text{ } p \text{ fois dérivable sur } I\}.$$

Chacune de ces inclusions est **stricte**.

Démonstration des inclusions. La deuxième est la plus instructive : si f est $(p+1)$ fois dérivable, alors $f^{(p)}$ admet une dérivée, à savoir $f^{(p+1)}$; or une fonction dérivable est continue (section 4) ; donc $f^{(p)}$ est continue, c'est-à-dire $f \in \mathcal{C}^p(I)$. La première et la troisième s'en déduisent : être de classe $\mathcal{C}^{p+1}$, c'est être $(p+1)$ fois dérivable **avec en plus** la continuité de $f^{(p+1)}$; et être de classe $\mathcal{C}^p$, c'est être p fois dérivable avec en plus une continuité. $\square$

Les trois témoins de stricte inclusion, qu'il faut savoir citer :

- La fonction $x \mapsto |x|$ est de classe $\mathcal{C}^0$ sur $\mathbb{R}$ sans y être dérivable : le point anguleux en 0 de la section 5.
- La fonction $x \mapsto x^2 \sin(1/x)$, prolongée par 0 en 0, est **dérivable sur $\mathbb{R}$ sans être de classe $\mathcal{C}^1$** : sa dérivée existe partout mais n'est pas continue en 0. Être p fois dérivable est donc strictement plus faible qu'être de classe $\mathcal{C}^p$.

- Pour chaque $p \geq 1$, la fonction $x \mapsto x^p|x|$ est de classe $\mathcal{C}^p$ sur $\mathbb{R}$ sans y être $(p+1)$ fois dérivable : une récurrence courte montre que sa dérivée p -ième vaut $(p+1)!|x|$, qui est continue mais non dérivable en 0. Les classes $\mathcal{C}^p$ forment donc une hiérarchie infinie, et jamais stationnaire.

18.2 Stabilité par les opérations

Propriété

Soient I un intervalle, $p \in \mathbb{N} \cup \{\infty\}$, et u, v deux fonctions de classe $\mathcal{C}^p$ sur I .

1. Pour tous réels λ et μ , la fonction $\lambda u + \mu v$ est de classe $\mathcal{C}^p$, et

$$(\lambda u + \mu v)^{(p)} = \lambda u^{(p)} + \mu v^{(p)}.$$

2. Le produit uv est de classe $\mathcal{C}^p$ (sa dérivée p -ième est donnée par la formule de Leibniz ci-dessous).
3. Le quotient $\frac{u}{v}$ est de classe $\mathcal{C}^p$ sur tout intervalle où v ne s'annule pas.
4. Si u est de classe $\mathcal{C}^p$ sur I à valeurs dans un intervalle J , et w de classe $\mathcal{C}^p$ sur J , alors $w \circ u$ est de classe $\mathcal{C}^p$ sur I .

Le point 1 se démontre par une récurrence immédiate à partir de la linéarité de la dérivation, établie à la section 7. Le point 4 mérite d'être écrit, parce que sa récurrence a une subtilité.

Démonstration du point 4, par récurrence sur p . L'énoncé à démontrer par récurrence est : « pour **toutes** fonctions u et w de classe $\mathcal{C}^p$ dont les domaines s'emboîtent, $w \circ u$ est de classe $\mathcal{C}^p$ ». La quantification sur toutes les fonctions fait partie de la propriété récurrente ; sans elle, l'hérédité ne pourrait pas s'appliquer aux fonctions dérivées, qui ne sont pas celles de départ.

Initialisation, $p = 0$. La composée de deux fonctions continues est continue (chapitre précédent).

Hérédité. Supposons la propriété vraie au rang p , et soient u et w de classe $\mathcal{C}^{p+1}$. Elles sont en particulier dérivables, et la formule de dérivation d'une composée (section 7) donne

$$(w \circ u)' = (w' \circ u) \times u'.$$

Or w' est de classe $\mathcal{C}^p$ (puisque w est $\mathcal{C}^{p+1}$), et u est de classe $\mathcal{C}^p$: l'hypothèse de récurrence, appliquée au couple (u, w') et non au couple de départ, donne que $w' \circ u$ est de classe $\mathcal{C}^p$. De plus u' est de classe $\mathcal{C}^p$. Le produit de deux fonctions $\mathcal{C}^p$ est $\mathcal{C}^p$ par le point 2 : donc $(w \circ u)'$ est de classe $\mathcal{C}^p$, ce qui signifie exactement que $w \circ u$ est de classe $\mathcal{C}^{p+1}$. $\square$

Ce que ces stabilités permettent d'écrire en copie. Les fonctions usuelles — polynômes, exp, ln, sin, cos, tan, Arcsin, Arccos, Arctan, $x \mapsto x^\alpha$ — sont de classe $\mathcal{C}^\infty$ sur leur domaine, ce qui se vérifie sur leurs dérivées. Par conséquent, toute fonction obtenue à partir d'elles par sommes, produits, quotients et composées est de classe $\mathcal{C}^\infty$ sur tout intervalle où elle est définie. Cela autorise la phrase qui remplace des pages de calcul :

« La fonction f est de classe $\mathcal{C}^\infty$ sur I comme quotient de fonctions de classe $\mathcal{C}^\infty$ dont le dénominateur ne s'annule pas sur I . »

Encore faut-il l'écrire, et vérifier l'annulation du dénominateur.

Exemple

Quatre dérivées n -ièmes de référence, toutes obtenues par une récurrence immédiate, et toutes réutilisées aux sections 19 à 23.

a. Pour $a \in \mathbb{R}$ et $f(x) = e^{ax} : f^{(n)}(x) = a^n e^{ax}.$ $\frac{n!}{(1-x)^{n+1}}.$

b. Pour $f = \sin : f^{(n)}(x) = \sin(x + n\frac{\pi}{2}).$

c. Pour $f(x) = \frac{1}{1-x}$ sur $] -\infty, 1[: f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}.$

d. Pour $f(x) = \ln(1+x)$ et $n \geq 1 : f^{(n)}(x) = \frac{(-1)^{n-1}(n-1)!}{(1+x)^n}.$

Le point **b** se démontre ainsi : la formule est vraie pour $n = 0$, et si elle l'est au rang n , alors

$$f^{(n+1)}(x) = \cos\left(x + n\frac{\pi}{2}\right) = \sin\left(x + n\frac{\pi}{2} + \frac{\pi}{2}\right) = \sin\left(x + (n+1)\frac{\pi}{2}\right).$$

La formule analogue pour $\cos$ s'obtient de la même manière.

18.3 La formule de Leibniz**Remarque**

Selon les facts. La formule de Leibniz figure au premier semestre dans plusieurs universités, où elle sert à calculer des dérivées n -ièmes et à établir des relations de récurrence sur les coefficients de Taylor ; d'autres maquettes ne la mentionnent qu'à l'occasion d'un exercice. Elle est traitée ici, avec sa démonstration : c'est l'un des rares résultats de ce chapitre dont la preuve soit purement combinatoire, et c'est un excellent entraînement au maniement des sommes.

Théorème

Formule de Leibniz. Soient I un intervalle, $n \in \mathbb{N}$, et u, v deux fonctions n fois dérivables sur I . Alors uv est n fois dérivable sur I et

$$(uv)^{(n)} = \sum_{k=0}^n \binom{n}{k} u^{(k)} v^{(n-k)}.$$

La ressemblance avec la formule du binôme de Newton n'est pas fortuite : mêmes coefficients binomiaux, les puissances étant remplacées par des ordres de dérivation. Les deux formules se démontrent d'ailleurs de la même façon, et par le même ingrédient combinatoire.

Notons aussi que la formule est **symétrique** en u et v : le changement d'indice $k \mapsto n - k$ échange leurs rôles, puisque $\binom{n}{n-k} = \binom{n}{k}$. On a donc le droit de choisir laquelle des deux fonctions joue le rôle de u , et l'on choisira toujours celle dont les dérivées s'arrêtent.

Démonstration, par récurrence sur n . La propriété à démontrer au rang n est : « pour **toutes** fonctions u et v n fois dérivables sur I , le produit uv est n fois dérivable et la formule est vraie ».

Initialisation. Au rang $n = 0$, la somme se réduit à son unique terme $\binom{0}{0}u^{(0)}v^{(0)} = uv$, et $(uv)^{(0)} = uv$. La formule est vraie. Au rang $n = 1$, elle donne $u'v + uv'$, qui est la dérivée d'un produit établie à la section 7 : c'est cohérent.

Hérédité. Supposons la formule vraie au rang n , et soient u et v deux fonctions $(n+1)$ fois dérivables sur I . Elles sont en particulier n fois dérivables, donc

$$(uv)^{(n)} = \sum_{k=0}^n \binom{n}{k} u^{(k)} v^{(n-k)}.$$

Dans cette somme, chaque terme est le produit de $u^{(k)}$ et de $v^{(n-k)}$, avec $k \leq n$ et $n-k \leq n$: ces deux fonctions sont dérivables, puisque u et v le sont $n+1$ fois. Chaque terme est donc dérivable, et la somme aussi. En dérivant terme à terme (linéarité) et en appliquant la dérivée d'un produit à chacun :

$$(uv)^{(n+1)} = \sum_{k=0}^n \binom{n}{k} \left(u^{(k+1)} v^{(n-k)} + u^{(k)} v^{(n-k+1)} \right) = S_1 + S_2,$$

où l'on a posé

$$S_1 = \sum_{k=0}^n \binom{n}{k} u^{(k+1)} v^{(n-k)} \quad \text{et} \quad S_2 = \sum_{k=0}^n \binom{n}{k} u^{(k)} v^{(n+1-k)}.$$

Le regroupement, étape décisive. Les deux sommes n'ont pas le même terme générique : dans S_1 , l'ordre de dérivation de u est décalé d'un cran. **On effectue donc un changement d'indice dans S_1 , et dans S_1 seulement.** Posons $j = k + 1$: quand k parcourt $0, 1, \dots, n$, l'indice j parcourt $1, 2, \dots, n+1$, et l'on a $u^{(k+1)} = u^{(j)}$ ainsi que $v^{(n-k)} = v^{(n-j+1)} = v^{(n+1-j)}$. D'où

$$S_1 = \sum_{j=1}^{n+1} \binom{n}{j-1} u^{(j)} v^{(n+1-j)}.$$

Dans S_2 , on renomme simplement k en j , sans aucun décalage :

$$S_2 = \sum_{j=0}^n \binom{n}{j} u^{(j)} v^{(n+1-j)}.$$

Les deux sommes portent maintenant sur le **même** terme générique $u^{(j)} v^{(n+1-j)}$, mais **pas sur les mêmes indices** : S_1 va de 1 à $n+1$, S_2 de 0 à n . On ne peut réunir deux sommes que si elles portent sur les mêmes indices : on isole donc le terme $j = n+1$ de S_1 et le terme $j = 0$ de S_2 .

$$\begin{aligned} S_1 + S_2 &= \binom{n}{n} u^{(n+1)} v^{(0)} + \sum_{j=1}^n \binom{n}{j-1} u^{(j)} v^{(n+1-j)} \\ &\quad + \sum_{j=1}^n \binom{n}{j} u^{(j)} v^{(n+1-j)} + \binom{n}{0} u^{(0)} v^{(n+1)}. \end{aligned}$$

Les deux sommes centrales portent désormais sur les mêmes indices $j = 1, \dots, n$: on les réunit terme à terme, et la **relation de Pascal**

$$\binom{n}{j-1} + \binom{n}{j} = \binom{n+1}{j},$$

acquise au chapitre « Logique, raisonnement, ensembles et applications », donne

$$S_1 + S_2 = u^{(n+1)} v + \sum_{j=1}^n \binom{n+1}{j} u^{(j)} v^{(n+1-j)} + u v^{(n+1)}.$$

Enfin $\binom{n+1}{n+1} = 1$ et $\binom{n+1}{0} = 1$: les deux termes isolés sont exactement les termes d'indices $j = n+1$ et $j = 0$ de la somme complète. En les y réintégrant,

$$(uv)^{(n+1)} = \sum_{j=0}^{n+1} \binom{n+1}{j} u^{(j)} v^{(n+1-j)},$$

qui est la formule au rang $n+1$. La récurrence est établie. $\square$

Méthode

Les trois gestes du regroupement, ceux qui décident si la démonstration aboutit.

1. **Le changement d'indice se fait dans une seule des deux sommes** : celle où l'ordre de dérivation de u a été décalé. Le faire dans les deux ramène au point de départ.
2. **Deux sommes ne se réunissent que si elles portent sur les mêmes indices**. D'où l'extraction des deux termes extrêmes, qui n'est pas une astuce mais une nécessité de forme.
3. **La relation de Pascal est le seul ingrédient combinatoire**. C'est exactement la même étape que dans la démonstration de la formule du binôme, et la reconnaître dispense de la redécouvrir.

Exemple

Un facteur polynomial, dont les dérivées s'arrêtent. Calculons la dérivée n -ième de $f(x) = x^2 e^x$ sur $\mathbb{R}$.

Le choix. On pose $u(x) = x^2$ et $v(x) = e^x$, et non l'inverse : c'est le polynôme qu'on met en u , parce que ses dérivées finissent par s'annuler, ce qui tronque la somme.

Les dérivées de u . $u'(x) = 2x$, $u''(x) = 2$, et $u^{(k)} = 0$ pour tout $k \geq 3$.

Les dérivées de v . $v^{(m)}(x) = e^x$ pour tout m .

Application de la formule. Seuls les termes $k = 0$, $k = 1$ et $k = 2$ sont non nuls :

$$f^{(n)}(x) = \binom{n}{0} x^2 e^x + \binom{n}{1} (2x) e^x + \binom{n}{2} \times 2 \times e^x.$$

Comme $\binom{n}{0} = 1$, $\binom{n}{1} = n$ et $\binom{n}{2} = \frac{n(n-1)}{2}$, il vient

$$f^{(n)}(x) = (x^2 + 2nx + n(n-1)) e^x.$$

Vérification. Pour $n = 1$, la formule donne $(x^2 + 2x)e^x$, et le calcul direct donne $(x^2 e^x)' = (2x + x^2)e^x$: c'est bien cela. Pour $n = 2$, elle donne $(x^2 + 4x + 2)e^x$, et la dérivation directe de $(x^2 + 2x)e^x$ donne $(2x + 2 + x^2 + 2x)e^x = (x^2 + 4x + 2)e^x$: c'est encore cela. **Une vérification sur les deux premiers rangs coûte trois lignes et détecte la quasi-totalité des erreurs de coefficients.**

Une remarque de forme. La formule reste valable pour $n = 0$ et $n = 1$, avec la convention $\binom{n}{k} = 0$ dès que $k > n$, qui est la convention usuelle.

L'exercice 19 pousse la formule plus loin : on y établit une **relation de récurrence** entre les dérivées successives d'une fonction, ce qui est le véritable usage de Leibniz en première année. L'exercice 20 travaille la classe exacte d'une fonction définie par raccord, en combinant cette section et le théorème de la limite de la dérivée de la section 17 — deux outils qu'on emploie presque toujours ensemble.

18.4 Ce que cette section prépare

Une dernière remarque, et elle ouvre tout le reste du chapitre. Pour une fonction suffisamment régulière au voisinage d'un point a , on dispose désormais d'une suite de nombres :

$$f(a), \quad f'(a), \quad f''(a), \quad \dots, \quad f^{(n)}(a).$$

La section 19 montre que ces $n + 1$ nombres **suffisent** à décrire le comportement de f au voisinage de a , à un reste négligeable près : ce sont eux qui deviennent les coefficients des formules de Taylor, puis des développements limités des sections 20 à 23. Savoir les calculer devient donc un enjeu de calcul à part entière — et c'est précisément ce que la formule de Leibniz permet quand f est un produit, ce que la stabilité par composition permet quand f est une composée.

19. Les formules de Taylor

Tout ce qui précède tournait autour d'une seule idée : remplacer f , au voisinage d'un point, par une droite. La dérivabilité en a s'écrivait $f(a+h) = f(a) + f'(a)h + h\varepsilon(h)$, la tangente était le meilleur objet du premier degré, et l'inégalité des accroissements finis de la section 13 chiffrait l'erreur commise. Les formules de Taylor font la même chose avec un **polynôme de degré n** au lieu d'une droite, et elles répondent aux deux questions qu'on se pose alors : quel polynôme, et quelle erreur.

Le premier point ne fait presque pas débat. Si l'on cherche un polynôme P de degré au plus n qui colle à f en a , la façon la plus naturelle de « coller » est d'imposer les mêmes valeurs de la fonction **et de ses n premières dérivées** en a . Cette condition détermine P entièrement, et le calcul est immédiat.

Définition

Soit I un intervalle, $a \in I$ et $f : I \rightarrow \mathbb{R}$ une fonction n fois dérivable en a . Le **polynôme de Taylor de f à l'ordre n au point a** est

$$\begin{aligned} P_n(x) &= \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \\ &= f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n. \end{aligned}$$

Le **reste d'ordre n** est la fonction $R_n = f - P_n$, définie sur I .

La notation P_n tait deux dépendances qu'il ne faut jamais perdre de vue : ce polynôme dépend de f **et du point a** . Quand deux fonctions ou deux points circulent dans un même exercice, on écrit $P_{n,a}$, ou l'on nomme les polynômes différemment. C'est une source d'erreur classique en copie, et elle ne coûte rien à éviter.

Deux vérifications valent d'être faites une fois pour toutes, et elles se font en dérivant terme à terme. D'abord, $P_n^{(j)}(a) = f^{(j)}(a)$ pour tout $j \leq n$: chaque dérivation abaisse d'un cran les puissances de $x-a$, et en $x=a$ tous les termes s'annulent sauf un. Ensuite, si f est elle-même polynomiale de degré au plus n , alors $P_n = f$ exactement, et le reste est nul : les formules de Taylor n'ont rien à dire sur les polynômes, elles servent à approcher les autres fonctions par des polynômes.

Reste la deuxième question, celle de l'erreur, et c'est elle qui donne trois théorèmes au lieu d'un. Les trois formules qui suivent ont exactement la même partie polynomiale P_n . Elles ne diffèrent que par ce qu'elles disent de R_n , et donc par ce qu'elles exigent de f . Ne les confondre jamais est le premier acquis attendu de cette section.

19.1 Taylor-Young : ce qui se passe tout près de a

Théorème

Formule de Taylor-Young. Soit I un intervalle, a un point de I , $n \geq 1$ un entier et $f : I \rightarrow \mathbb{R}$ une fonction n fois dérivable en a . Alors

$$f(x) = P_n(x) + o((x-a)^n) \quad \text{quand } x \rightarrow a.$$

Autrement dit, il existe une fonction ε_n , définie au voisinage de a et de limite nulle en a , telle que

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + (x-a)^n \varepsilon_n(x).$$

L'hypothèse mérite d'être lue lentement, parce qu'elle est plus fine qu'il n'y paraît. Dire que f est n fois dérivable en a suppose que $f^{(n-1)}$ existe sur tout un voisinage de a — sans quoi on ne pourrait pas la dériver une fois de plus — et que cette fonction $f^{(n-1)}$ est dérivable **au point** a . On ne demande rien de plus : ni la continuité de $f^{(n)}$, ni même son existence ailleurs qu'en a .

Le symbole $o((x-a)^n)$ est ici pris **quand x tend vers a** , et nulle part ailleurs. Cette précision n'est pas une coquetterie : dans la section 22, des restes pris en 0, en a et en $+\infty$ se croiseront dans le même calcul, et un o dont on ignore le point où il est pris ne veut rien dire.

Démonstration, par récurrence sur n . Notons $\mathcal{H}_n$ l'énoncé : « pour toute fonction g définie au voisinage de a et n fois dérivable en a , on a $g(x) = P_n^g(x) + o((x-a)^n)$ quand $x \rightarrow a$ », où P_n^g désigne le polynôme de Taylor de g .

Initialisation. Pour $n = 1$, l'énoncé $\mathcal{H}_1$ s'écrit $g(x) = g(a) + g'(a)(x-a) + o(x-a)$, ce qui est exactement la définition de la dérivabilité de g en a , sous la forme donnée dans les premières sections du chapitre. Il n'y a rien à démontrer.

Hérédité. Soit $n \geq 2$; supposons $\mathcal{H}_{n-1}$ vraie et soit f une fonction n fois dérivable en a . Posons, pour x au voisinage de a ,

$$R_n(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k.$$

Comme $n \geq 2$, la fonction f est au moins une fois dérivable au voisinage de a , donc R_n l'est aussi, et en dérivant terme à terme le polynôme :

$$R'_n(x) = f'(x) - \sum_{k=1}^n \frac{f^{(k)}(a)}{(k-1)!} (x-a)^{k-1} = f'(x) - \sum_{j=0}^{n-1} \frac{(f')^{(j)}(a)}{j!} (x-a)^j,$$

la seconde écriture s'obtenant par le changement d'indice $j = k-1$ et l'identité $f^{(j+1)} = (f')^{(j)}$. On reconnaît le reste d'ordre $n-1$ de la fonction f' . Or f' est $n-1$ fois dérivable en a : l'hypothèse de récurrence $\mathcal{H}_{n-1}$ s'applique à f' et donne

$$R'_n(x) = o((x-a)^{n-1}) \quad \text{quand } x \rightarrow a.$$

Le passage du reste à sa dérivée. Il s'agit maintenant de remonter de R'_n à R_n , et c'est la **règle de L'Hôpital de la section 16** qui le fait. Posons $D(x) = (x-a)^n$. Les deux fonctions R_n et D sont continues en a , de valeur nulle en a — pour R_n , parce que tous les termes du polynôme sauf $f(a)$ s'annulent en a — et dérivables au voisinage épointé de a , où $D'(x) = n(x-a)^{n-1}$ ne s'annule pas. Enfin,

$$\frac{R'_n(x)}{D'(x)} = \frac{1}{n} \cdot \frac{R'_n(x)}{(x-a)^{n-1}} \xrightarrow{x \rightarrow a} 0,$$

d'après ce qui précède. La règle de L'Hôpital conclut : $\frac{R_n(x)}{(x-a)^n} \rightarrow 0$, c'est-à-dire $R_n(x) = o((x-a)^n)$, ce qui est $\mathcal{H}_n$ pour f . $\square$

Pourquoi cette démonstration plutôt qu'une autre. On peut aussi déduire Taylor-Young de Taylor-Lagrange, qui vient juste après, et c'est plus court : trois lignes. Mais cette déduction exige f de classe $\mathcal{C}^n$, donc une hypothèse strictement plus forte, et l'on perdrait précisément le cas qui servira le plus dans les sections 20 à 23, celui d'une fonction seulement dérivable n fois **au point** considéré. La récurrence ci-dessus, elle, n'utilise que la définition de la dérivée et un théorème déjà démontré.

Remarque

La version « classe $\mathcal{C}^n$ », celle que beaucoup de cours énoncent. Si f est de classe $\mathcal{C}^n$ sur un intervalle ouvert contenant a , elle est en particulier n fois dérivable en a , et la formule ci-dessus s'applique. C'est cette version, plus commode à vérifier et amplement suffisante pour toutes les fonctions usuelles, que vous rencontrerez le plus souvent. Retenez que l'énoncé démontré ici la contient.

19.2 Taylor-Lagrange : un contrôle chiffré, sur tout un intervalle

Théorème

Formule de Taylor-Lagrange. Soient $a < b$ deux réels, $n \geq 0$ un entier, et $f : [a, b] \rightarrow \mathbb{R}$ telle que

1. f est de classe $\mathcal{C}^n$ sur le **segment** $[a, b]$;
2. $f^{(n)}$ est dérivable sur l'**intervalle ouvert** $]a, b[$.

Alors il existe un réel c vérifiant $a < c < b$ tel que

$$f(b) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c).$$

L'énoncé vaut aussi lorsque $b < a$, le point c étant alors strictement compris entre b et a : il suffit d'appliquer ce qui suit à la fonction $x \mapsto f(-x)$, ou de refaire la démonstration à l'identique. On l'emploiera dans les deux sens sans autre commentaire.

Le cas $n = 0$ est l'égalité des accroissements finis de la section 12 : la formule s'y réduit à $f(b) = f(a) + (b-a)f'(c)$. Taylor-Lagrange en est donc la généralisation exacte, et c'est aussi de Rolle qu'elle se démontre. La dissymétrie des hypothèses — classe $\mathcal{C}^n$ sur le fermé, une dérivée de plus sur l'ouvert — est la même que dans Rolle, et pour la même raison : on n'a besoin de dériver qu'à l'intérieur, mais il faut de la continuité jusqu'aux bords pour appliquer le théorème.

Méthode

D'où sort la fonction auxiliaire. Le geste est le même que pour l'égalité des accroissements finis, où l'on avait retranché à f la corde pour se ramener à Rolle. Ici, on veut faire apparaître un facteur $(b-a)^{n+1}$ dans le reste : on considère donc la quantité

$$\varphi(x) = \underbrace{f(b) - \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} (b-x)^k}_{\text{le reste, vu comme fonction du point de base } x} - K \frac{(b-x)^{n+1}}{(n+1)!},$$

où x — et non plus a — est la variable, et où K est une constante qu'on n'a pas encore choisie. Deux observations donnent tout le reste. D'une part $\varphi(b) = f(b) - f(b) - 0 = 0$, **quelle que soit** la valeur de K : le second bout est gratuit. D'autre part, $\varphi(a) = 0$ est une équation du premier degré en K , qu'on peut donc résoudre. On a alors deux zéros, et Rolle s'applique.

Démonstration. Elle se déroule en quatre temps.

Étape 1 : le choix de K . Posons

$$A = f(b) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k$$

le reste que l'on cherche à évaluer, et définissons la constante K par l'égalité

$$A = K \frac{(b-a)^{n+1}}{(n+1)!}, \quad \text{soit} \quad K = \frac{(n+1)! A}{(b-a)^{n+1}},$$

ce qui a un sens puisque $b-a \neq 0$. À ce stade K est un réel parfaitement défini, et **tout ce qui reste à démontrer est que ce réel est une valeur de $f^{(n+1)}$** .

Étape 2 : la fonction auxiliaire et ses deux zéros. Définissons $\varphi : [a, b] \rightarrow \mathbb{R}$ par

$$\varphi(x) = f(b) - \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} (b-x)^k - K \frac{(b-x)^{n+1}}{(n+1)!}.$$

En $x = b$, tous les termes de la somme s'annulent sauf celui d'indice $k = 0$, qui vaut $f(b)$, et le dernier terme est nul : $\varphi(b) = 0$. En $x = a$, la somme vaut $P_n(b)$, donc $\varphi(a) = A - K \frac{(b-a)^{n+1}}{(n+1)!} = 0$ par le choix de K .

Étape 3 : la dérivée se télescope. La fonction φ est continue sur $[a, b]$, car f y est de classe $\mathcal{C}^n$ et chaque $f^{(k)}$ pour $k \leq n$ y est donc continue. Elle est dérivable sur $]a, b[$, où $f^{(n)}$ l'est par hypothèse. Dérivons la somme, chaque terme étant un produit :

$$\frac{d}{dx} \left[\frac{f^{(k)}(x)}{k!} (b-x)^k \right] = \frac{f^{(k+1)}(x)}{k!} (b-x)^k - \frac{f^{(k)}(x)}{(k-1)!} (b-x)^{k-1},$$

le second terme étant absent pour $k = 0$. En sommant de $k = 0$ à n , chaque terme positif d'indice k s'annule avec le terme négatif d'indice $k+1$: **tout se télescope**, et il ne subsiste que le terme positif d'indice n . La dérivée de la somme vaut donc $\frac{f^{(n+1)}(x)}{n!} (b-x)^n$. Comme par ailleurs la dérivée de $-K \frac{(b-x)^{n+1}}{(n+1)!}$ vaut $K \frac{(b-x)^n}{n!}$, on obtient

$$\varphi'(x) = \frac{(b-x)^n}{n!} (K - f^{(n+1)}(x)) \quad \text{pour tout } x \in]a, b[.$$

Étape 4 : Rolle. La fonction φ est continue sur $[a, b]$, dérivable sur $]a, b[$, et $\varphi(a) = \varphi(b) = 0$. Le **théorème de Rolle, établi à la section 11**, fournit un réel c tel que $a < c < b$ et $\varphi'(c) = 0$. Comme $c < b$, le facteur $(b-c)^n$ est strictement positif, donc non nul : l'égalité $\varphi'(c) = 0$ impose $K = f^{(n+1)}(c)$. En reportant dans l'étape 1, on obtient exactement l'égalité annoncée. $\square$

Il n'y a aucune raison, dans cette démonstration, d'espérer connaître c : il est fourni par Rolle, qui est un théorème d'existence. **Un c dont on prétend calculer la valeur est presque toujours le signe d'une faute de raisonnement**, et c'est ce qui rend la formule inutilisable telle quelle. D'où la conséquence suivante, qui est la version dont on se sert réellement. L'exercice 24 refait cette démonstration pas à pas, en faisant déterminer la constante K , vérifier une par une les hypothèses de Rolle, et apparaître le télescopage au lieu de l'affirmer.

19.3 L'inégalité de Taylor-Lagrange

Théorème

Inégalité de Taylor-Lagrange. Sous les hypothèses de la formule ci-dessus, supposons de plus $f^{(n+1)}$ **bornée** sur $]a, b[$, et posons

$$M_{n+1} = \sup_{t \in]a, b[} |f^{(n+1)}(t)|.$$

Alors

$$\left| f(b) - \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (b-a)^k \right| \leq \frac{|b-a|^{n+1}}{(n+1)!} M_{n+1}.$$

La démonstration tient en une ligne : le point c appartient à $]a, b[$, donc $|f^{(n+1)}(c)| \leq M_{n+1}$, et il suffit de passer aux valeurs absolues dans la formule de Taylor-Lagrange. L'inconnu c a disparu, et il ne reste qu'une majoration effectivement calculable.

Dans la pratique, on ne cherche presque jamais la valeur exacte de M_{n+1} : **n'importe quel majorant de $|f^{(n+1)}|$ convient**, et un majorant grossier donne déjà une inégalité correcte, simplement moins fine. C'est un point de rédaction : écrire « $|f^{(n+1)}(t)| \leq 2$ pour tout t de l'intervalle » suffit, il est inutile de calculer un sup.

Remarque

Une quatrième formule existe, et elle n'est pas de ce chapitre. La formule de Taylor avec reste intégral exprime R_n non plus par une valeur inconnue $f^{(n+1)}(c)$, mais par une expression exacte faisant intervenir $f^{(n+1)}$ tout entière. Elle est plus précise que les deux précédentes et les redonne toutes les deux, mais elle relève du chapitre « Calcul intégral élémentaire », qui construit l'outil dont elle a besoin. On la mentionne pour que vous sachiez qu'elle existe ; rien de ce chapitre ne l'utilise.

19.4 Les trois formules, côte à côte

C'est le tableau à réviser, et la seule chose à en retenir tient en une phrase : **Taylor-Young est local et ne majore rien**, Taylor-Lagrange chiffre.

Formule	Hypothèses	Portée	Ce qu'elle donne	Ce qu'elle ne donne pas
Taylor-Young	f n fois dérivable en a	locale, en a seulement	la nature du reste : $o((x-a)^n)$	aucune majoration à x fixé
Taylor-Lagrange	$\mathcal{C}^n$ sur $[a, b]$, $f^{(n)}$ dérivable sur $]a, b[$	globale, sur tout $[a, b]$	une égalité avec un c inconnu	la valeur de c
Inégalité de T.-L.	les mêmes, plus $f^{(n+1)}$ bornée	globale, sur tout $[a, b]$	une majoration calculable	le signe de l'erreur

La ligne du haut est celle qu'on emploie mal le plus souvent. Écrire « $f(x) = P_n(x) + o((x-a)^n)$, donc $f(1,2) \simeq P_n(1,2)$ à 10^{-4} près » est une faute : le symbole o dit que le reste devient négligeable **lorsque x tend vers a** , et absolument rien sur la valeur de ce reste au point 1,2, qui est un point fixé. Pour un chiffre, il faut l'inégalité. C'est exactement ce que l'exercice 21 fait mesurer, en comparant le reste réel à ce que chaque formule autorisait à en dire.

19.5 Un encadrement numérique, entièrement traité

Exemple

Encadrer $\sqrt{e}$ à mieux que 10^{-3} . On applique l'inégalité de Taylor-Lagrange à $f = \exp$, entre $a = 0$ et $b = \frac{1}{2}$, en remarquant que $\sqrt{e} = e^{1/2} = f(b)$.

Les dérivées. La fonction $\exp$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$ et $f^{(k)} = \exp$ pour tout k , donc $f^{(k)}(0) = 1$: le polynôme de Taylor en 0 à l'ordre n est $P_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$.

Le majorant. Sur $\left]0, \frac{1}{2}\right]$, on a $e^t \leq e^{1/2}$ par croissance de l'exponentielle, et $e^{1/2} \leq 2$ puisque $e \leq 4$. On prend donc $M_{n+1} = 2$, quel que soit n . C'est un majorant grossier, et il suffira.

Le choix de n . L'inégalité donne

$$\left| \sqrt{e} - P_n\left(\frac{1}{2}\right) \right| \leq \frac{2}{2^{n+1}(n+1)!}.$$

On cherche le plus petit n rendant ce majorant inférieur à 10^{-3} . Pour $n = 3$, il vaut $\frac{2}{16 \times 24} \simeq 5,2 \times 10^{-3}$: insuffisant. Pour $n = 4$, il vaut $\frac{2}{32 \times 120} \simeq 5,2 \times 10^{-4}$: c'est gagné. **On retient $n = 4$** , et l'on remarque au passage que chaque cran supplémentaire divise l'erreur par $2(n+2)$ — ici par 10 en passant de $n = 3$ à $n = 4$, puis par 12 au cran suivant — ce qui explique qu'un ordre de plus suffise presque toujours.

Le calcul.

$$P_4\left(\frac{1}{2}\right) = 1 + \frac{1}{2} + \frac{1}{8} + \frac{1}{48} + \frac{1}{384} = \frac{211}{128} = 1,6484375.$$

La conclusion, rédigée. D'après l'inégalité de Taylor-Lagrange appliquée à $\exp$ à l'ordre 4 entre 0 et $\frac{1}{2}$,

$$1,6479 \leq \sqrt{e} \leq 1,6490,$$

en ayant arrondi les bornes dans le sens qui élargit l'encadrement. On peut donc affirmer $\sqrt{e} = 1,64$ à 10^{-2} près, et rien de plus : la troisième décimale n'est pas déterminée, puisque l'encadrement exact $1,647917 \leq \sqrt{e} \leq 1,648958$ laisse subsister 7 et 8. Il faudrait passer à $n = 5$, dont le majorant vaut $4,4 \times 10^{-5}$, pour la fixer.

Ce dernier paragraphe est le point de méthode de toute la section. **Un encadrement ne donne que ce qu'il donne**, et annoncer une décimale que les bornes ne séparent pas est une faute de rigueur, pas une approximation bénigne. L'exercice 22 reprend l'exercice sur trois valeurs, et l'exercice 23 pose la question dans l'autre sens : combien de termes faut-il pour atteindre une précision imposée à l'avance ?

20 20. Développements limités : définition et premières propriétés

Taylor-Young affirme que, sous une hypothèse de dérivabilité, f s'écrit près de a comme un polynôme plus un reste négligeable. On va maintenant **oublier d'où vient ce polynôme** et ne garder que la forme de l'écriture. Ce renversement est le contenu de tout le reste du chapitre : il donne un objet qu'on peut additionner, multiplier, composer, et surtout obtenir sans dériver n fois — ce qui, pour une fonction comme $x \mapsto \frac{\ln(1+x)}{\cos x}$, change tout.

Définition

Soit f une fonction définie au voisinage de a , sauf peut-être en a , et $n \in \mathbb{N}$. On dit que f admet un **développement limité d'ordre n en a** , noté $DL_n(a)$, s'il existe des réels $a_0, a_1, \dots, a_n$ tels que

$$f(x) = a_0 + a_1(x-a) + \dots + a_n(x-a)^n + o((x-a)^n) \quad \text{quand } x \rightarrow a.$$

Le polynôme $a_0 + a_1(x-a) + \dots + a_n(x-a)^n$ s'appelle la **partie régulière** du développement, et le terme $o((x-a)^n)$ son **reste**.

Le o est pris **quand** $x \rightarrow a$, toujours, et c'est la seule lecture possible de la définition. Dans un exercice où plusieurs points interviennent, on l'écrit en toutes lettres au moins une fois : « $o((x-1)^3)$ quand $x \rightarrow 1$ ».

Méthode

Se ramener systématiquement en 0. Poser $x = a + h$ transforme « $x \rightarrow a$ » en « $h \rightarrow 0$ » et $(x-a)^k$ en h^k . La fonction $g : h \mapsto f(a+h)$ admet alors un $DL_n(0)$ **si et seulement si** f admet un $DL_n(a)$, avec exactement les mêmes coefficients. Tous les développements de référence s'écrivent donc en 0, et l'on n'en apprend pas d'autres.

Le seul point de vigilance est la **conclusion** : si la question portait sur f en a , la réponse doit repartir en x . Un développement laissé en h sans revenir à la variable de l'énoncé est une réponse incomplète, et c'est la faute la plus fréquente de l'exercice 28.

20.1 L'unicité de la partie régulière

C'est la propriété qui rend l'objet utilisable : sans elle, « le » développement limité n'aurait pas de sens, et l'identification des coefficients — qui est le geste de calcul le plus employé des trois sections suivantes — serait illégitime.

Théorème

Unicité. Si f admet un $DL_n(a)$, sa partie régulière est unique : deux écritures de la forme ci-dessus ont les mêmes coefficients $a_0, \dots, a_n$.

Démonstration. Quitte à poser $x = a + h$, on se ramène à $a = 0$. Supposons

$$f(h) = \sum_{k=0}^n a_k h^k + o(h^n) = \sum_{k=0}^n b_k h^k + o(h^n) \quad (h \rightarrow 0).$$

En soustrayant, et en notant $c_k = a_k - b_k$, il vient

$$\sum_{k=0}^n c_k h^k = o(h^n) \quad (h \rightarrow 0).$$

Supposons par l'absurde que les c_k ne soient pas tous nuls, et notons p le plus petit indice tel que $c_p \neq 0$. Pour $h \neq 0$, on peut factoriser par h^p et diviser :

$$c_p + c_{p+1}h + \dots + c_n h^{n-p} = \frac{o(h^n)}{h^p} = o(h^{n-p}).$$

Faisons tendre h vers 0. Le membre de gauche tend vers c_p , car c'est une fonction polynomiale. Le membre de droite tend vers 0, puisque $n-p \geq 0$: si $n > p$ c'est un o d'une puissance qui tend vers 0, et si $n = p$ c'est un $o(1)$, qui tend vers 0 par définition. Par unicité de la limite, $c_p = 0$, ce qui contredit le choix de p . Tous les c_k sont donc nuls. $\square$

Le ressort de cette preuve est le geste à retenir : devant une égalité de la forme « polynôme = $o(h^n)$ », on divise par la plus petite puissance présente et l'on fait tendre h vers 0. Il ressortira à chaque fois qu'il faudra montrer qu'un terme est nul.

20.2 Ce que l'unicité fait tomber

Propriété

Soit f admettant un $DL_n(a)$ de partie régulière $\sum_{k=0}^n a_k(x-a)^k$.

1. **Troncature.** Pour tout $p \leq n$, f admet un $DL_p(a)$, obtenu en ne gardant que les termes de degré $\leq p$.
2. **Lecture des deux premiers coefficients.** Si f est définie en a , alors $a_0 = f(a)$, et f est continue en a . Si de plus $n \geq 1$, alors f est dérivable en a et $a_1 = f'(a)$.
3. **Parité.** Si $a = 0$ et si f est paire, tous les coefficients d'indice impair sont nuls ; si f est impaire, tous ceux d'indice pair le sont.

Démonstration du point 1. Les termes $a_{p+1}(x-a)^{p+1} + \dots + a_n(x-a)^n$ sont tous des $o((x-a)^p)$, puisque chacun s'écrit $(x-a)^p$ multiplié par une quantité qui tend vers 0 ; et le reste $o((x-a)^n)$ en est un aussi, pour la même raison. On les absorbe donc tous dans un unique $o((x-a)^p)$. $\square$

Attention au sens de cette propriété : elle ne marche que vers le bas. Un DL_5 donne gratuitement le DL_2 , mais un DL_2 ne dit **rien** de l'ordre 3, et un développement obtenu à un ordre insuffisant ne se « prolonge » pas après coup. C'est la source de la moitié des erreurs de la section 22.

Démonstration du point 2. Faisons tendre x vers a dans l'égalité : tous les termes de degré ≥ 1 tendent vers 0, le reste aussi, donc $f(x) \rightarrow a_0$. Comme f est définie en a , l'égalité prise en $x = a$ donne directement $f(a) = a_0$, d'où la continuité. Si $n \geq 1$, on isole ensuite le taux d'accroissement : pour $x \neq a$,

$$\frac{f(x) - f(a)}{x - a} = a_1 + a_2(x-a) + \dots + a_n(x-a)^{n-1} + \frac{o((x-a)^n)}{x-a},$$

et le dernier terme est un $o((x-a)^{n-1})$, donc tend vers 0. Le taux d'accroissement tend vers a_1 : f est dérivable en a , de nombre dérivé a_1 . $\square$

Démonstration du point 3, f paire. Écrivons le développement en 0 : $f(x) = \sum a_k x^k + o(x^n)$. En remplaçant x par $-x$, et puisque $o((-x)^n)$ est encore un $o(x^n)$ quand $x \rightarrow 0$,

$$f(-x) = \sum_{k=0}^n (-1)^k a_k x^k + o(x^n).$$

Or $f(-x) = f(x)$: les deux écritures sont deux $DL_n(0)$ de la même fonction. Par unicité, $a_k = (-1)^k a_k$ pour tout k , donc $2a_k = 0$ dès que k est impair. $\square$

Le point 3 est un **outil de vérification gratuit**. Un développement de $\cos$ où figure un x^3 , un développement d'une fonction impaire avec un terme constant : le calcul est faux, et l'on peut le dire avant même de le relire.

20.3 Les deux pièges de cette section

Remarque

Piège n°1 : un DL_n ne donne pas n dérivées. Le point 2 ci-dessus s'arrête à l'ordre 1, et ce n'est pas une paresse de rédaction : **c'est le mieux qu'on puisse dire**. L'existence d'un $DL_2(0)$ n'entraîne nullement que f soit deux fois dérivable en 0.

Le témoin est le membre d'indice 3 de la famille $f_\alpha : x \mapsto x^\alpha \sin(1/x)$ que l'exercice 04 a balayée pour $\alpha = 0, 1, 2, 3$, et c'est le même fil qui court d'un bout à l'autre du chapitre : prenons $f(x) = x^3 \sin(1/x)$ pour $x \neq 0$, prolongée par $f(0) = 0$. Comme $|f(x)| \leq |x|^3$, on a $f(x) = 0 + 0x + 0x^2 + o(x^2)$: c'est un $DL_2(0)$, de partie régulière nulle. Pourtant, le calcul de f'' en 0 échoue : f' existe partout, mais son taux d'accroissement en 0 oscille sans limite, et $f''(0)$ n'existe pas. L'exercice 04 a établi la même chose un cran plus bas, pour $\alpha = 2$, sur la continuité de f' ; on retrouve ici, un ordre plus loin, exactement le même mécanisme d'oscillation.

L'implication ne va donc que dans un sens, et il faut savoir lequel :

$$f \text{ } n \text{ fois dérivable en } a \implies f \text{ admet un } DL_n(a),$$

c'est Taylor-Young, et **la réciproque est fausse dès $n = 2$** . Conséquence pratique immédiate : lorsqu'un développement limité vous est **donné** par l'énoncé, vous pouvez en lire $f(a)$ et $f'(a)$, et pas $f''(a)$ — sauf si l'énoncé précise par ailleurs que f est deux fois dérivable, auquel cas Taylor-Young et l'unicité donnent $f''(a) = 2a_2$. C'est exactement le tri que fait faire l'exercice 25.

Remarque

Piège n°2 : un développement limité n'est pas une somme infinie. L'écriture $e^x = 1 + x + \frac{x^2}{2} + \dots$ que vous croiserez ailleurs, avec des points de suspension et sans reste, désigne une **série**, objet de deuxième année. Ce n'est pas ce qui est écrit ici. Un DL_n est une égalité **locale**, à un nombre **fini** de termes, valable au voisinage d'un point et assortie d'un reste.

Trois conséquences, toutes vérifiées en partiel. D'abord, un DL_n ne dit rien de f loin de a : le développement de $\frac{1}{1-x}$ en 0 ne renseigne pas sur $x = 5$. Ensuite, une fonction peut admettre un $DL_n(0)$ **pour tout** n sans être « égale à sa série » : la fonction $x \mapsto e^{-1/x^2}$ prolongée par 0 en 0 est de classe $\mathcal{C}^\infty$ et toutes ses dérivées sont nulles en 0, donc tous ses développements limités en 0 sont nuls, alors qu'elle ne s'annule qu'en 0. Enfin, on ne parle jamais de « convergence » d'un développement limité : le mot n'a pas de sens ici, et l'employer signale une confusion.

21 21. Les développements limités usuels, et la primitivation

21.1 Le tableau

Tous ces développements sont pris **en** 0, et tous les o le sont **quand** $x \rightarrow 0$.

$f(x)$	DL en 0
e^x	$1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + o(x^n)$
$\frac{1}{1-x}$	$1 + x + x^2 + \dots + x^n + o(x^n)$
$\frac{1}{1+x}$	$1 - x + x^2 - \dots + (-1)^n x^n + o(x^n)$
$\ln(1+x)$	$x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + \frac{(-1)^{n-1} x^n}{n} + o(x^n)$
$\cos x$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + \frac{(-1)^p x^{2p}}{(2p)!} + o(x^{2p+1})$
$\sin x$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + \frac{(-1)^p x^{2p+1}}{(2p+1)!} + o(x^{2p+2})$
$\operatorname{ch} x$	$1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots + \frac{x^{2p}}{(2p)!} + o(x^{2p+1})$
$\operatorname{sh} x$	$x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots + \frac{x^{2p+1}}{(2p+1)!} + o(x^{2p+2})$
$(1+x)^\alpha$	$1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!} x^n + o(x^n)$

Deux détails de ce tableau sont volontaires. Pour $\cos$, $\sin$, ch et sh , le reste est écrit **un cran plus loin** que le dernier terme visible : le développement de $\cos$ à l'ordre $2p$ est en réalité valable à l'ordre $2p+1$, car le coefficient de x^{2p+1} est nul par parité. C'est gratuit, et cela fait gagner un ordre dans les calculs. Pour $(1+x)^\alpha$, le cas $\alpha = -1$ redonne la ligne de $\frac{1}{1+x}$, et les cas $\alpha = \frac{1}{2}$ et $\alpha = -\frac{1}{2}$ sont ceux qu'on emploie le plus :

$$\sqrt{1+x} = 1 + \frac{x}{2} - \frac{x^2}{8} + o(x^2), \quad \frac{1}{\sqrt{1+x}} = 1 - \frac{x}{2} + \frac{3x^2}{8} + o(x^2).$$

21.2 Comment chacun s'obtient : un cas détaillé

Ils viennent tous de Taylor-Young, et le travail est toujours le même : calculer $f^{(k)}(0)$ pour tout k .

Exemple

Le développement de $\cos$. La fonction $\cos$ est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}$, donc Taylor-Young s'applique à tout ordre. Ses dérivées successives se répètent avec une période de quatre : $\cos' = -\sin$, $\cos'' = -\cos$, $\cos''' = \sin$, $\cos^{(4)} = \cos$. En 0, cela donne la suite de valeurs 1, 0, -1, 0, 1, 0, -1, ... En clair : $\cos^{(k)}(0) = 0$ si k est impair, et $\cos^{(2p)}(0) = (-1)^p$. Taylor-Young à l'ordre $2p$ donne alors

$$\cos x = \sum_{j=0}^p \frac{(-1)^j}{(2j)!} x^{2j} + o(x^{2p}),$$

et le reste se relève d'un ordre par parité, comme expliqué ci-dessus. $\square$

Les autres se traitent de la même façon : $\exp^{(k)} = \exp$ donne $f^{(k)}(0) = 1$; pour $\ln(1+x)$, une récurrence immédiate donne $f^{(k)}(x) = \frac{(-1)^{k-1}(k-1)!}{(1+x)^k}$, d'où $f^{(k)}(0) = (-1)^{k-1}(k-1)!$ et le coefficient $\frac{(-1)^{k-1}}{k}$ après division par $k!$. **Faites-en un, une fois**, et ne refaites pas les autres : c'est le rôle de l'exercice 26 que de vous faire choisir lequel.

21.3 Comment les retenir sans les apprendre par cœur

C'est le vrai point de méthode de la section, et il tient en une phrase : **deux développements suffisent, tout le reste s'en déduit en trois lignes**. Les deux à connaître sont celui de e^x et celui de $\frac{1}{1-x}$, l'un parce qu'il est le plus régulier, l'autre parce qu'il est la somme géométrique.

Méthode

Reconstruire les usuels.

1. **De $\frac{1}{1-x}$ à $\frac{1}{1+x}$** : on remplace x par $-x$. Les puissances paires ne changent pas, les impaires changent de signe : les signes alternent.
2. **De $\frac{1}{1+x}$ à $\ln(1+x)$** : on primitive terme à terme (théorème ci-dessous), puisque $\ln(1+x)$ a pour dérivée $\frac{1}{1+x}$. Chaque $(-1)^k x^k$ devient $\frac{(-1)^k x^{k+1}}{k+1}$, et la constante vaut $\ln(1) = 0$. **Le développement de $\ln(1+x)$ commence donc par x , pas par 1** — se tromper là-dessus est l'erreur la plus fréquente du tableau.
3. **De e^x à ch et sh** : par définition $\operatorname{ch} x = \frac{e^x + e^{-x}}{2}$ et $\operatorname{sh} x = \frac{e^x - e^{-x}}{2}$. Dans la somme, les puissances impaires s'annulent ; dans la différence, les paires. On obtient les deux d'un coup, sans aucun signe à mémoriser.
4. **De ch et sh à cos et sin** : ce sont les mêmes, **avec les signes qui alternent**. Un moyen de contrôle imparable : cos est paire donc ne contient que des puissances paires, sin est impaire donc ne contient que des impaires, et les deux premiers termes se retrouvent à la main par $\cos x \simeq 1 - \frac{x^2}{2}$ et $\sin x \simeq x - \frac{x^3}{6}$.
5. $(1+x)^\alpha$: il n'y a rien à retenir, le coefficient de x^k est le produit de k facteurs qui descendent d'un en un à partir de α , divisé par $k!$. On l'écrit toujours à l'ordre voulu et jamais au-delà.

21.4 Primitiver un développement limité

Théorème

Théorème de primitivation. Soit I un intervalle ouvert contenant 0 et $f : I \rightarrow \mathbb{R}$ une fonction dérivable sur I . Si f' admet un $\text{DL}_n(0)$

$$f'(x) = a_0 + a_1x + \cdots + a_nx^n + o(x^n),$$

alors f admet un $\text{DL}_{n+1}(0)$, obtenu en **primitivant terme à terme** la partie régulière, la constante étant $f(0)$:

$$f(x) = f(0) + a_0x + \frac{a_1}{2}x^2 + \cdots + \frac{a_n}{n+1}x^{n+1} + o(x^{n+1}).$$

Démonstration. Notons $Q(x) = f(0) + a_0x + \frac{a_1}{2}x^2 + \cdots + \frac{a_n}{n+1}x^{n+1}$ le polynôme annoncé, et posons $\psi = f - Q$ sur I . La fonction ψ est dérivable sur I , vérifie $\psi(0) = f(0) - f(0) = 0$, et

$$\psi'(x) = f'(x) - (a_0 + a_1x + \cdots + a_nx^n) = o(x^n) \quad (x \rightarrow 0)$$

par hypothèse. Il s'agit de montrer que $\psi(x) = o(x^{n+1})$.

Soit $\varepsilon > 0$. Le fait que $\psi'(x) = o(x^n)$ fournit un réel $\delta > 0$ tel que $|\psi'(t)| \leq \varepsilon |t|^n$ pour tout t vérifiant $|t| \leq \delta$, avec $[-\delta, \delta] \subset I$. Fixons x tel que $0 < |x| \leq \delta$. Sur le segment d'extrémités 0 et x , la fonction

ψ est dérivable, et pour tout t de ce segment on a $|t| \leq |x|$, donc

$$|\psi'(t)| \leq \varepsilon |t|^n \leq \varepsilon |x|^n.$$

L'inégalité des accroissements finis de la section 13, appliquée entre 0 et x avec ce majorant, donne

$$|\psi(x) - \psi(0)| \leq \varepsilon |x|^n \times |x - 0| = \varepsilon |x|^{n+1}.$$

Comme $\psi(0) = 0$, on obtient $|\psi(x)| \leq \varepsilon |x|^{n+1}$ dès que $|x| \leq \delta$. C'est exactement la définition de $\psi(x) = o(x^{n+1})$, donc $f = Q + \psi$ admet le développement annoncé. $\square$

Exemple

Le développement d'Arctan. La fonction Arctan est dérivable sur $\mathbb{R}$, de dérivée $\frac{1}{1+x^2}$. En remplaçant x par x^2 dans le développement de $\frac{1}{1+x}$ — substitution légitime, la composition étant traitée à la section suivante et la condition $u(0) = 0$ étant ici satisfaite :

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - \dots + (-1)^n x^{2n} + o(x^{2n}).$$

Le théorème s'applique, avec $\text{Arctan}(0) = 0$:

$$\text{Arctan } x = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + \frac{(-1)^n x^{2n+1}}{2n+1} + o(x^{2n+1}).$$

Il n'existe pas de chemin plus court : calculer $\text{Arctan}^{(7)}(0)$ à la main pour appliquer Taylor-Young est une punition, pas une méthode.

Le développement d'Arcsin s'obtient de la même façon, en partant de $\text{Arcsin}'(x) = \frac{1}{\sqrt{1-x^2}}$, qu'on développe par la ligne $(1+u)^{-1/2}$ du tableau avec $u = -x^2$. On trouve $\text{Arcsin } x = x + \frac{x^3}{6} + \frac{3x^5}{40} + o(x^5)$. C'est le travail de l'exercice 26.

21.5 On ne dérive pas un développement limité

Remarque

La réciproque du théorème est fausse, et l'erreur est sanctionnée. Un $DL_n(0)$ de f ne donne **aucun** développement limité de f' , pas même à l'ordre 0. Ce n'est pas une précaution d'énoncé : c'est un fait, et voici le témoin.

Soit $f(x) = x^2 \sin(1/x)$ pour $x \neq 0$, prolongée par $f(0) = 0$ — la cousine de la fonction rencontrée à l'exercice 04. De $|f(x)| \leq x^2$ on tire $f(x) = 0 + 0x + o(x)$: la fonction admet un $DL_1(0)$, de partie régulière nulle. Dériver cette partie régulière terme à terme donnerait « $f'(x) = 0 + o(1)$ », c'est-à-dire $f'(x) \rightarrow 0$ quand $x \rightarrow 0$. Or, pour $x \neq 0$,

$$f'(x) = 2x \sin(1/x) - \cos(1/x),$$

qui n'a aucune limite en 0, puisque le premier terme tend vers 0 et le second oscille entre -1 et 1 . La fonction f' n'admet donc **pas** de $DL_0(0)$, et l'écriture qu'on venait de produire est fausse.

Asymétrie à retenir telle quelle : on primitive un développement limité, on ne le dérive jamais. La raison profonde est que primitiver est une opération régularisante — elle moyenne les oscillations, et c'est l'inégalité des accroissements finis qui le traduit dans la démonstration ci-dessus — tandis que dériver les amplifie. Si l'on a besoin d'un développement de f' , on le calcule à partir de f' elle-même, jamais à partir de celui de f .

22.2. Opérations sur les développements limités

Voici le cœur technique du chapitre. Les règles elles-mêmes sont simples ; ce qui coûte des points, c'est **l'ordre auquel on tronque chaque morceau**. Développer trop court fait perdre le résultat, développer trop long fait perdre du temps et multiplie les erreurs de calcul. Toute cette section apprend à trancher cette question **avant** de calculer.

Dans tout ce qui suit, f et g admettent des $DL_n(0)$ de parties régulières respectives A et B , et tous les o sont pris quand $x \rightarrow 0$.

22.1 Somme et produit par un réel

Propriété

$f + g$ et λf (pour $\lambda \in \mathbb{R}$) admettent des $DL_n(0)$, de parties régulières $A + B$ et λA .

C'est immédiat, la somme de deux $o(x^n)$ étant un $o(x^n)$. **Le seul piège est de sommer deux développements d'ordres différents** : si l'un est à l'ordre 3 et l'autre à l'ordre 5, la somme n'est connue qu'à l'ordre 3, et il faut commencer par tronquer le second. On ne garde jamais un terme dont on n'est pas sûr qu'il soit exact.

22.2 Produit : l'ordre du produit est le minimum des ordres

Propriété

Produit. fg admet un $DL_n(0)$, dont la partie régulière est le produit AB tronqué au degré n .

Démonstration, en une ligne d'idée. On écrit $f = A + o(x^n)$ et $g = B + o(x^n)$, puis on développe le produit : $fg = AB + A o(x^n) + B o(x^n) + o(x^n) o(x^n)$. Les fonctions polynomiales A et B sont bornées au voisinage de 0, donc les trois derniers termes sont des $o(x^n)$. Reste AB , qui est un polynôme de degré au plus $2n$; ses termes de degré $> n$ sont des $o(x^n)$, et l'on tronque. $\square$

Méthode

À quel ordre développer chaque facteur. C'est la question à se poser avant d'écrire quoi que ce soit, et la réponse dépend de la **valuation** de chaque facteur, c'est-à-dire du degré du premier terme non nul de sa partie régulière.

Si f a une valuation p (elle commence par $a_p x^p$ avec $a_p \neq 0$) et si l'on veut le produit fg à l'ordre n , alors g n'a besoin d'être développé qu'à l'ordre $n-p$: tout terme de g de degré $> n-p$ sera multiplié par au moins x^p et sortira au-delà de l'ordre n . Symétriquement pour f .

Traduction pratique : chaque fois qu'un facteur « commence tard », il fait gagner des ordres sur l'autre. Un produit par $\sin x$, qui commence par x , autorise à développer l'autre facteur un cran moins loin ; un produit par $1 - \cos x$, qui commence par $\frac{x^2}{2}$, en fait gagner deux.

Exemple

Le développement de $x \mapsto e^x \sin x$ à l'ordre 3 en 0. La fonction $\sin$ a pour valuation 1 : il suffit donc de développer e^x à l'ordre 2, et $\sin x$ à l'ordre 3. Avec les deux lignes de la section 21 :

$$e^x \sin x = \left(1 + x + \frac{x^2}{2} + o(x^2)\right) \left(x - \frac{x^3}{6} + o(x^3)\right).$$

On développe en ne conservant que les monômes de degré au plus 3 : x , puis $x \cdot x = x^2$, puis $x \cdot \frac{x^2}{2} = \frac{x^3}{2}$ et $-\frac{x^3}{6}$. D'où

$$e^x \sin x = x + x^2 + \frac{x^3}{3} + o(x^3).$$

Contrôle de cohérence : la fonction n'est ni paire ni impaire, donc rien n'oblige les coefficients à se raréfier, et le premier terme est bien x , produit des valuations 0 et 1.

22.3 Composition : la condition $u(0) = 0$ n'est pas négociable**Propriété**

Composition. Soient u admettant un $DL_n(0)$ avec $u(0) = 0$, et v admettant un $DL_n(0)$ de partie régulière C . Alors $v \circ u$ admet un $DL_n(0)$, dont la partie régulière s'obtient en **substituant** la partie régulière de u dans C et en tronquant au degré n .

La condition $u(0) = 0$ est ce qui rend la substitution licite : le développement de v n'est valable qu'au voisinage de 0, et il faut donc que $u(x)$ soit proche de 0 quand x l'est. Si $u(0) = b \neq 0$, la composition n'est pas interdite, mais elle demande le développement de v en b , pas en 0. Substituer un u de terme constant non nul dans le développement de v en 0 est une faute grave, et elle produit des résultats faux sans le moindre signal d'alarme.

Le gain d'ordre fonctionne ici comme pour le produit, et il est encore plus rentable : si u a une valuation $p \geq 1$, alors u^k a une valuation kp , et il est inutile de développer v au-delà de l'ordre $\lfloor n/p \rfloor$. C'est ce qui a permis, à la section précédente, de développer $\frac{1}{1+x^2}$ à l'ordre $2n$ en ne connaissant $\frac{1}{1+u}$ qu'à l'ordre n .

Exemple

Le développement de $x \mapsto e^{\sin x}$ à l'ordre 3 en 0. Posons $u(x) = \sin x$, qui vérifie $u(0) = 0$ et a pour valuation 1 : on développe u à l'ordre 3 et l'exponentielle à l'ordre 3 également. Avec $u = x - \frac{x^3}{6} + o(x^3)$:

$$u^2 = x^2 + o(x^3),$$

$$u^3 = x^3 + o(x^3),$$

le terme croisé $2 \times x \times (-x^3/6)$ étant de degré 4. Il vient

$$e^{\sin x} = 1 + u + \frac{u^2}{2} + \frac{u^3}{6} + o(u^3) = 1 + x + \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^3}{6} + o(x^3) = 1 + x + \frac{x^2}{2} + o(x^3).$$

Le coefficient de x^3 est nul : c'est un résultat, pas une omission, et il faut l'écrire. Noter au passage la substitution $o(u^3) = o(x^3)$, licite parce que u et x ont même valuation.

22.4 Quotient : deux méthodes, une seule à retenir

Pour $\frac{f}{g}$ avec $g(0) \neq 0$, on ne mémorise aucune formule. Deux chemins mènent au résultat.

Le premier, et le plus sûr, ramène le quotient à une composition. On factorise $g(0)$ au dénominateur pour faire apparaître $1 + u$ avec $u(0) = 0$, puis on compose avec le développement de $\frac{1}{1+u}$ du tableau. C'est une méthode, pas une astuce : elle marche toujours, et elle réutilise ce qui est déjà connu.

Le second est la division suivant les puissances croissantes : on pose la division des deux parties régulières comme une division euclidienne, mais en commençant par les termes de plus bas degré, et l'on s'arrête dès que le quotient a atteint le degré n . Elle est plus rapide sur les fractions rationnelles, et plus risquée sur les autres. Choisissez-en une et tenez-vous-y.

Si $g(0) = 0$, il n'y a pas de développement limité du quotient en général — la fonction peut même ne pas être bornée. On factorise alors par la plus grande puissance de x possible au numérateur et au dénominateur, et l'on se retrouve dans le cas précédent, ou bien la fonction n'a pas de DL et il faut le dire.

22.5 Quand un ordre trop court échoue, et comment le voir**Méthode**

Une division par x^p coûte p ordres. C'est la règle qui sert tout le temps, et elle explique presque tous les échecs. Pour obtenir la limite d'un quotient dont le dénominateur est équivalent à x^p , il faut développer le numérateur à l'ordre p **au moins**, et souvent plus si les premiers termes se simplifient.

Le geste à prendre : **regarder d'abord le dénominateur**, en déduire l'ordre visé au numérateur, ajouter un cran de sécurité si des simplifications sont prévisibles, puis calculer. Un $o(x^p)$ divisé par x^p donne un $o(1)$, qui tend vers 0 : c'est ce qui permet de conclure. Un $o(x^p)$ divisé par x^{p+1} ne donne **rien du tout**, et c'est le signe visible qu'on a développé trop court.

Exemple

Calculer $\lim_{x \rightarrow 0} \frac{\cos x - \sqrt{1-x^2}}{x^4}$.

Première tentative, à l'ordre 2. On a $\cos x = 1 - \frac{x^2}{2} + o(x^2)$, et $\sqrt{1-x^2} = 1 - \frac{x^2}{2} + o(x^2)$ en substituant $u = -x^2$ dans $\sqrt{1+u}$. La différence vaut donc $o(x^2)$, et le quotient s'écrit $\frac{o(x^2)}{x^4}$: **on ne peut rien en dire**, ni la limite, ni même l'existence. L'échec est visible, et c'est ce qui doit déclencher la reprise du calcul à un ordre plus élevé, pas une conclusion hâtive « la limite est 0 ».

Deuxième tentative, à l'ordre 4. Le dénominateur est x^4 : c'est l'ordre à viser.

$$\begin{aligned}\cos x &= 1 - \frac{x^2}{2} + \frac{x^4}{24} + o(x^4), \\ \sqrt{1-x^2} &= 1 + \frac{-x^2}{2} - \frac{(-x^2)^2}{8} + o(x^4) = 1 - \frac{x^2}{2} - \frac{x^4}{8} + o(x^4).\end{aligned}$$

Les termes constants et en x^2 se simplifient exactement — c'était prévisible, et c'est bien pour cela qu'il fallait aller à l'ordre 4 :

$$\cos x - \sqrt{1-x^2} = \frac{x^4}{24} + \frac{x^4}{8} + o(x^4) = \frac{x^4}{6} + o(x^4).$$

En divisant par x^4 : $\frac{1}{6} + o(1)$, donc la limite vaut $\frac{1}{6}$.

Ce diagnostic — quel ordre, et pourquoi — est l'objet de l'exercice 27, qui fait chiffrer le coût de chaque troncature.

22.6 Développement en un point $a \neq 0$, et en l'infini**Méthode**

En un point a quelconque : on translate. On pose $x = a + h$, on développe la fonction $h \mapsto f(a+h)$ en 0 par les méthodes ci-dessus, puis **on revient à x** en remplaçant h par $x - a$. Tout le travail se fait en 0, il n'y a pas de second catalogue à apprendre.

En $+\infty$ ou $-\infty$: on inverse. On pose $x = \frac{1}{t}$, ce qui envoie $x \rightarrow +\infty$ sur $t \rightarrow 0^+$, et l'on développe en 0 la fonction $t \mapsto f(1/t)$. Le résultat se relit en x en remplaçant t par $\frac{1}{x}$: on obtient une expression en puissances de $\frac{1}{x}$, avec un reste $o(1/x^n)$ quand $x \rightarrow +\infty$. On parle alors souvent de **développement asymptotique**. Attention au signe quand $x \rightarrow -\infty$: l'égalité $\sqrt{x^2} = |x|$ vaut $-x$ de ce côté, et c'est l'oubli qui fait rater la moitié des exercices d'asymptote.

Exemple

Développement de $f(x) = x^2 \operatorname{Arctan} \frac{1}{x}$ en $+\infty$, à l'ordre 3 en $\frac{1}{x}$. Posons $t = \frac{1}{x}$, avec $t \rightarrow 0^+$, de sorte que

$$f\left(\frac{1}{t}\right) = \frac{1}{t^2} \operatorname{Arctan} t.$$

On divise par t^2 : il faut donc développer Arctan **deux crans plus loin** que l'ordre voulu, soit à l'ordre 5. Le tableau de la section 21 donne

$$\operatorname{Arctan} t = t - \frac{t^3}{3} + \frac{t^5}{5} + o(t^5),$$

d'où, en divisant par t^2 ,

$$\frac{\operatorname{Arctan} t}{t^2} = \frac{1}{t} - \frac{t}{3} + \frac{t^3}{5} + o(t^3).$$

C'est la règle « une division par t^p coûte p ordres », appliquée avec $p = 2$. En revenant à x :

$$x^2 \operatorname{Arctan} \frac{1}{x} = x - \frac{1}{3x} + \frac{1}{5x^3} + o\left(\frac{1}{x^3}\right) \quad (x \rightarrow \pm\infty).$$

Le développement vaut des **deux** côtés : Arctan est impaire et son développement est valable pour t tendant vers 0 par valeurs positives comme négatives, donc aucun signe ne se retourne ici — contrairement au cas d'une racine carrée.

Ces deux changements de variable sont l'objet de l'exercice 28, qui les fait pratiquer dans les deux sens et sanctionne les retours à la variable initiale oubliés.

23 23. Ce que les développements limités permettent de faire

Quatre usages, et pas un de plus. Chacun se reconnaît à la question posée, et chacun a une méthode propre.

23.1 1. Lever une forme indéterminée

Méthode

La méthode, en trois gestes. *Un* : identifier la forme indéterminée et repérer la puissance de x qui « commande » le dénominateur. *Deux* : développer numérateur et dénominateur au même ordre, celui du dénominateur. *Trois* : factoriser par cette puissance et simplifier ; il ne reste qu'un quotient de deux fonctions ayant chacune une limite non nulle, ou une limite lisible. Pour les formes 1^∞ , 0^0 et ∞^0 , on passe d'abord à l'exponentielle : $u^v = e^{v \ln u}$, et l'on développe l'exposant.

Exemple

Calculer $\lim_{x \rightarrow 0} \frac{\ln(1+x) - x}{1 - \cos x}$. Le dénominateur vaut $\frac{x^2}{2} + o(x^2)$: il est commandé par x^2 , on développe donc tout à l'ordre 2.

Au numérateur, $\ln(1+x) - x = \left(x - \frac{x^2}{2} + o(x^2)\right) - x = -\frac{x^2}{2} + o(x^2)$. Le quotient s'écrit alors

$$\frac{-\frac{x^2}{2} + o(x^2)}{\frac{x^2}{2} + o(x^2)} = \frac{x^2 \left(-\frac{1}{2} + o(1)\right)}{x^2 \left(\frac{1}{2} + o(1)\right)} \xrightarrow{x \rightarrow 0} -1.$$

La factorisation par x^2 **avant** la simplification est ce qu'attend le correcteur : elle rend visible que les $o(x^2)$ deviennent des $o(1)$, donc tendent vers 0.

Remarque

Développement limité ou règle de L'Hôpital ? Les deux répondent à la même question, et la comparaison honnête est la suivante.

La **règle de L'Hôpital de la section 16** est la plus courte quand une seule application suffit et que les dérivées sont simples : deux lignes, et c'est fini. Elle devient un piège dès qu'il faut l'appliquer plusieurs fois, car chaque application exige de **revérifier** que la nouvelle forme est encore indéterminée et que le dénominateur dérivé ne s'annule pas. Trois applications enchaînées sans vérification, c'est trois occasions de fausse manœuvre — et c'est la faute la plus régulièrement sanctionnée du chapitre.

Le **développement limité** demande un investissement fixe (choisir l'ordre, écrire deux ou trois développements) mais il donne bien plus qu'une limite : il donne le **terme suivant**, donc la vitesse, donc l'équivalent, donc la position de la courbe. Sur l'exemple ci-dessus, L'Hôpital demanderait deux applications, le développement limité une ligne, et le développement limité fournit en prime l'information que le quotient vaut $-1 + O(x)$.

La règle du pouce : une seule dérivation évidente, L'Hôpital ; dès la deuxième, ou dès qu'on veut autre chose que la valeur limite, le développement limité. L'exercice 29 fait poser ce diagnostic sur huit limites **avant** tout calcul, ce qui est exactement l'ordre dans lequel il faut travailler.

23.2 2. Obtenir un équivalent**Définition**

Soient f et g définies au voisinage de a , sauf peut-être en a , avec g ne s'annulant pas sur un voisinage épointé de a . On dit que f **est équivalente à g en a** , et l'on note $f \sim_a g$, lorsque

$$\frac{f(x)}{g(x)} \xrightarrow{x \rightarrow a} 1.$$

Le point a fait partie de la notation, et l'omettre est une faute : $x^2 + x$ est équivalent à x en 0 et à x^2 en $+\infty$. Écrire « $x^2 + x \sim x$ » sans préciser où ne veut rien dire.

Propriété

Lien avec les développements limités. Si f admet un $DL_n(a)$ dont la partie régulière est non nulle, et si $a_p(x-a)^p$ est son **premier terme non nul**, alors

$$f \underset{a}{\sim} a_p (x-a)^p.$$

Démonstration. On écrit $f(x) = a_p(x-a)^p \left(1 + \frac{a_{p+1}}{a_p}(x-a) + \dots\right) + o((x-a)^n)$, puis on divise par $a_p(x-a)^p$: le contenu de la parenthèse tend vers 1, et le reste divisé donne un $o((x-a)^{n-p})$, qui tend vers 0 puisque $n \geq p$. Le quotient tend donc vers 1. $\square$

C'est la façon normale d'obtenir un équivalent : on développe, on regarde le premier terme non nul, c'est fini. Les équivalents usuels en 0 — $\sin x \sim x$, $1 - \cos x \sim \frac{x^2}{2}$, $\ln(1+x) \sim x$, $e^x - 1 \sim x$, $(1+x)^\alpha - 1 \sim \alpha x$ — se relisent tous directement sur le tableau de la section 21, et il est inutile de les apprendre séparément.

Propriété

Les deux règles d'or.

1. **On multiplie, on divise, on élève à une puissance fixe.** Si $f_1 \underset{a}{\sim} g_1$ et $f_2 \underset{a}{\sim} g_2$, alors

$$f_1 f_2 \underset{a}{\sim} g_1 g_2, \quad \frac{f_1}{f_2} \underset{a}{\sim} \frac{g_1}{g_2}, \quad \text{et } f_1^\beta \underset{a}{\sim} g_1^\beta \text{ pour } \beta \text{ réel fixé (quand les puissances ont un sens).}$$

2. **On n'additionne jamais des équivalents, et on ne les compose jamais.**

Le point 1 se démontre en une ligne, en multipliant ou en divisant les quotients qui tendent vers 1. Le point 2 demande deux contre-exemples, et il faut les connaître.

Exemple

Contre-exemple à l'addition. En $+\infty$, on a $x + \sqrt{x} \sim x$ et $-x \sim -x$. Si l'addition était licite, la somme $\sqrt{x}$ serait équivalente à $x - x = 0$, ce qui n'a même pas de sens, la définition exigeant un dénominateur non nul. Et de fait, $\sqrt{x}$ n'est équivalent à rien de nul : le quotient $\frac{\sqrt{x}}{x}$ tend vers 0, pas vers 1.

La raison en une phrase : un équivalent ne retient que le terme dominant, et une soustraction peut précisément faire disparaître ce terme dominant, ce qui laisse apparaître un terme dont on avait jeté l'information. Le remède est toujours le même : **revenir aux développements limités**, qui gardent, eux, les termes suivants.

Contre-exemple à la composition. En $+\infty$, $x+1 \sim x$. Pourtant $e^{x+1} = e \cdot e^x$, et le quotient $\frac{e^{x+1}}{e^x}$ vaut constamment e , qui n'est pas 1 : on n'a pas $e^{x+1} \sim e^x$. Un équivalent contrôle un rapport, pas un écart, et l'exponentielle transforme les écarts en rapports. Même prudence avec le logarithme, où il faut une hypothèse supplémentaire pour composer.

23.3 3. Tangente et position de la courbe

Le développement limité à l'ordre 1 redonne la tangente : $f(x) = f(a) + f'(a)(x-a) + o(x-a)$, et l'équation de la tangente en a est $y = f(a) + f'(a)(x-a)$. **Ce qui est nouveau, c'est que les termes suivants disent de quel côté la courbe passe.**

Méthode

Position par rapport à la tangente. On écrit un $DL_n(a)$ avec n assez grand pour qu'un terme au-delà du degré 1 soit non nul, et l'on note $a_p(x-a)^p$ le **premier terme non nul de degré ≥ 2** . La différence entre la courbe et la tangente vaut alors

$$f(x) - [f(a) + f'(a)(x-a)] = a_p(x-a)^p + o((x-a)^p),$$

et son signe, au voisinage de a , est celui de $a_p(x-a)^p$ — le reste étant négligeable devant lui. Deux cas :

- p **pair** : $(x-a)^p > 0$ des deux côtés, le signe est celui de a_p ; la courbe reste **du même côté** de sa tangente. Au-dessus si $a_p > 0$, en dessous si $a_p < 0$.
- p **impair** : $(x-a)^p$ change de signe en a ; la courbe **traverse** sa tangente. Le point est un **point d'inflexion**.

Exemple

Deux cas, sur les deux fonctions les plus courantes.

Pour $f = \exp$ en $a = 0$: $e^x = 1 + x + \frac{x^2}{2} + o(x^2)$. La tangente est $y = 1 + x$, le premier terme suivant est $\frac{x^2}{2}$, de degré $p = 2$ **pair** et de coefficient positif : la courbe de l'exponentielle est **au-dessus** de sa tangente en 0, des deux côtés. On retrouve l'inégalité $e^x \geq 1 + x$, mais seulement au voisinage de 0 ; la version globale demandera l'argument de convexité de la section 24.

Pour $f = \sin$ en $a = 0$: $\sin x = x - \frac{x^3}{6} + o(x^3)$. La tangente est $y = x$, et le premier terme suivant est $-\frac{x^3}{6}$, de degré $p = 3$ **impair** : la courbe traverse sa tangente en 0, qui est donc un point d'inflexion. Concrètement, $\sin x < x$ pour $x > 0$ petit, et $\sin x > x$ pour $x < 0$ petit.

23.4 4. Branches infinies**Méthode**

Asymptote oblique par développement en l'infini. On développe $f(x)$ en $+\infty$ selon les puissances de $\frac{1}{x}$:

$$f(x) = \alpha x + \beta + \frac{\gamma}{x} + o\left(\frac{1}{x}\right).$$

La droite d'équation $y = \alpha x + \beta$ est alors **asymptote** à la courbe en $+\infty$, puisque $f(x) - (\alpha x + \beta) \rightarrow 0$. Et le terme suivant donne la **position** : le signe de $\frac{\gamma}{x}$, pour x grand, est celui de γ en $+\infty$ et celui de $-\gamma$ en $-\infty$. La courbe est au-dessus de son asymptote là où ce signe est positif. Si $\gamma = 0$, on pousse d'un ordre de plus : la position se lit sur le premier terme non nul suivant.

Exemple

La fonction $f(x) = x^2 \operatorname{Arctan} \frac{1}{x}$. Le développement obtenu à la section 22 donne

$$f(x) = x - \frac{1}{3x} + o\left(\frac{1}{x}\right).$$

La droite $y = x$ est donc asymptote à la courbe en $+\infty$ **et** en $-\infty$, avec $\alpha = 1$ et $\beta = 0$. Reste la position, que donne le signe du terme suivant : en $+\infty$, $-\frac{1}{3x} < 0$, la courbe est **en dessous** de l'asymptote ; en $-\infty$, $-\frac{1}{3x} > 0$, elle est **au-dessus**. Un seul développement a donné les deux branches et leurs deux positions, ce qu'aucun calcul de limite ne pouvait fournir : la limite de $f(x) - x$ vaut 0 des deux côtés et ne dit rien du signe.

Méthode

Quand il n'y a pas d'asymptote : la branche parabolique. Le développement en $\frac{1}{x}$ n'existe pas toujours, et il faut alors revenir au diagnostic en deux temps.

1. On calcule $\frac{f(x)}{x}$ quand $x \rightarrow +\infty$. S'il tend vers $\pm\infty$, la courbe admet une **branche parabolique de direction** (Oy) . S'il tend vers 0, c'est une branche parabolique de direction (Ox) .
2. S'il tend vers un réel $\alpha \neq 0$, on calcule $f(x) - \alpha x$. Si cette différence tend vers un réel β , il y a une **asymptote** $y = \alpha x + \beta$; si elle tend vers $\pm\infty$, il y a une **branche parabolique de direction** $y = \alpha x$.

Exemple

Deux branches paraboliques. Pour $f(x) = x + \sqrt{x}$ sur $]0, +\infty[$: $\frac{f(x)}{x} = 1 + \frac{1}{\sqrt{x}} \rightarrow 1$, donc $\alpha = 1$; mais $f(x) - x = \sqrt{x} \rightarrow +\infty$. Il n'y a **pas** d'asymptote, mais une branche parabolique de direction $y = x$. Pour $g(x) = x \ln x$: $\frac{g(x)}{x} = \ln x \rightarrow +\infty$, branche parabolique de direction (Oy) .

L'exercice 30 fait trancher entre asymptote et branche parabolique sur une série de fonctions choisies pour que le diagnostic ne se devine pas. Et l'**exercice 18** rassemble tout : une étude complète de fonction, où les variations viennent de la section 15, la position de la tangente et les branches infinies de celle-ci, et la convexité de la suivante.

24 24. Convexité

24.1 La définition, et sa lecture géométrique

Définition

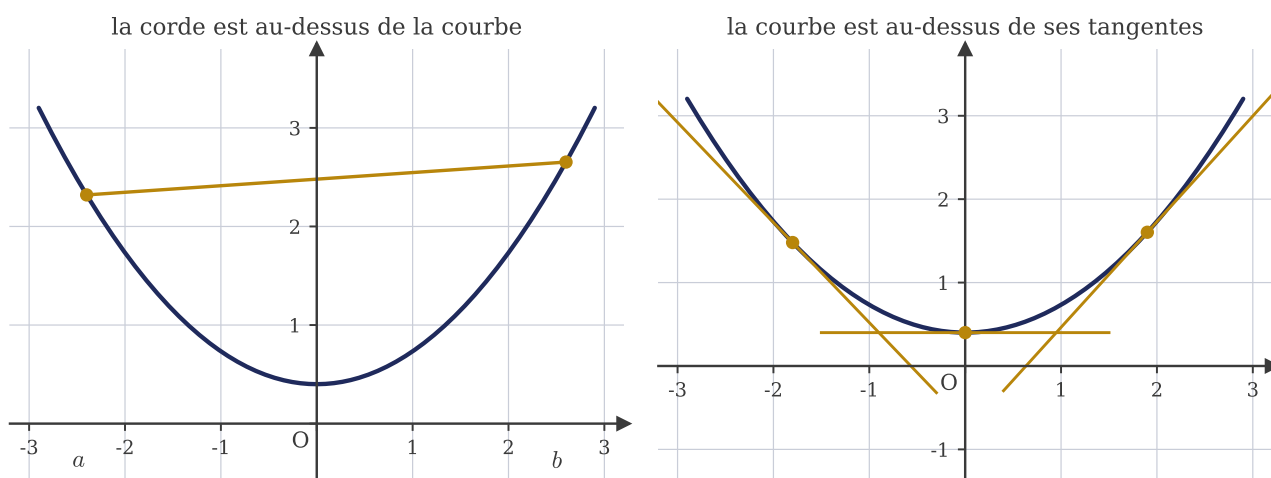
Soit I un intervalle et $f : I \rightarrow \mathbb{R}$. On dit que f est **convexe sur** I lorsque

$$\forall (x, y) \in I^2, \forall t \in [0, 1], \quad f(tx + (1 - t)y) \leq tf(x) + (1 - t)f(y).$$

On dit que f est **concave** lorsque $-f$ est convexe, c'est-à-dire lorsque l'inégalité ci-dessus est renversée.

Le point $tx + (1 - t)y$ décrit le segment $[x, y]$ quand t parcourt $[0, 1]$, et le nombre $tf(x) + (1 - t)f(y)$ est l'ordonnée, au-dessus de ce point, du segment joignant les deux points de la courbe. **La définition dit donc exactement une chose : entre deux points de la courbe, la corde est au-dessus de la courbe.** C'est le panneau de gauche de la figure ci-dessous.

Un mot sur I : la définition exige un **intervalle**, car il faut que $tx + (1 - t)y$ appartienne encore à l'ensemble de départ. Sur une réunion de deux intervalles disjoints, l'énoncé n'a pas de sens, exactement comme pour la monotonie à la section 15.

**24.2 L'inégalité des pentes**

C'est la reformulation la plus maniable de la définition, et c'est d'elle que descendent toutes les caractérisations qui suivent.

Théorème

Lemme des trois pentes. Soit f convexe sur I et $x < y < z$ trois points de I . Alors

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

En clair : les trois pentes des cordes croissent avec les points qui les portent. Ou encore, en fixant x : **la fonction pente** $t \mapsto \frac{f(t) - f(x)}{t - x}$ **est croissante** sur I privé de x .

Démonstration. Le point y est entre x et z : il s'écrit donc $y = tx + (1 - t)z$ avec

$$t = \frac{z - y}{z - x} \in]0, 1[, \quad 1 - t = \frac{y - x}{z - x}.$$

On le vérifie en développant. La convexité donne $f(y) \leq tf(x) + (1 - t)f(z)$.

Première inégalité. En retranchant $f(x)$ des deux côtés,

$$f(y) - f(x) \leq (t - 1)f(x) + (1 - t)f(z) = (1 - t)(f(z) - f(x)) = \frac{y - x}{z - x}(f(z) - f(x)).$$

On divise par $y - x$, qui est strictement positif, et l'on obtient l'inégalité de gauche.

Seconde inégalité. En retranchant cette fois l'inégalité de convexité à $f(z)$,

$$f(z) - f(y) \geq f(z) - tf(x) - (1-t)f(z) = t(f(z) - f(x)) = \frac{z-y}{z-x}(f(z) - f(x)),$$

et l'on divise par $z - y > 0$. $\square$

Ce lemme vaut **sans aucune hypothèse de dérivabilité** : c'est ce qui le rend précieux, et c'est aussi lui qui permet de démontrer qu'une fonction convexe sur un intervalle ouvert y est automatiquement continue. L'exercice 31 le fait lire sur une figure, puis l'applique à une simple table de valeurs : la convexité se teste sur des nombres, sans formule.

24.3 Caractérisations pour une fonction dérivable

Théorème

Soit f **dérivable** sur un intervalle I . Les trois énoncés suivants sont équivalents.

1. f est convexe sur I ;
2. f' est croissante sur I ;
3. la courbe de f est **au-dessus de toutes ses tangentes** : pour tous a et x dans I ,

$$f(x) \geq f(a) + f'(a)(x - a).$$

Démonstration de 1 implique 2. Soient $x < y$ dans I . D'après le lemme des trois pentes, pour tout t strictement compris entre x et y ,

$$\frac{f(t) - f(x)}{t - x} \leq \frac{f(y) - f(x)}{y - x} \leq \frac{f(y) - f(t)}{y - t}.$$

Dans l'inégalité de gauche, faisons tendre t vers x : le membre de gauche tend vers $f'(x)$, donc par passage à la limite dans une inégalité large, $f'(x) \leq \frac{f(y) - f(x)}{y - x}$. Dans celle de droite, faisons tendre

t vers y : on obtient $\frac{f(y) - f(x)}{y - x} \leq f'(y)$. En réunissant, $f'(x) \leq f'(y)$. $\square$

Démonstration de 2 implique 3, par l'égalité des accroissements finis. C'est le retour attendu du bloc B, et il est court. Soient a et x dans I , avec $x \neq a$.

Cas $x > a$. La fonction f est continue sur $[a, x]$ et dérivable sur $]a, x[$: l'égalité des accroissements finis de la section 12 fournit $c \in]a, x[$ tel que $f(x) - f(a) = f'(c)(x - a)$. Comme $c > a$ et que f' est croissante, $f'(c) \geq f'(a)$; en multipliant par $x - a > 0$, il vient $f(x) - f(a) \geq f'(a)(x - a)$.

Cas $x < a$. L'égalité des accroissements finis sur $[x, a]$ fournit $c \in]x, a[$ tel que $f(a) - f(x) = f'(c)(a - x)$. Cette fois $c < a$, donc $f'(c) \leq f'(a)$; en multipliant par $a - x > 0$, il vient $f(a) - f(x) \leq f'(a)(a - x)$, ce qui se réécrit $f(x) \geq f(a) + f'(a)(x - a)$.

Le cas $x = a$ est une égalité. Dans les deux cas, la courbe est au-dessus de la tangente en a . **Le sens de l'inégalité s'est retourné deux fois** — une fois sur $f'(c)$, une fois en repassant $f(x)$ à gauche — et c'est exactement là que les copies se trompent. $\square$

Démonstration de 3 implique 1. Soient $x < y$ dans I , $t \in]0, 1[$, et $m = tx + (1-t)y$, qui appartient à I . En appliquant l'hypothèse 3 au point $a = m$, d'abord en x puis en y :

$$f(x) \geq f(m) + f'(m)(x - m), \quad f(y) \geq f(m) + f'(m)(y - m).$$

Multiplions la première par $t > 0$, la seconde par $1 - t > 0$, et additionnons :

$$tf(x) + (1-t)f(y) \geq f(m) + f'(m)[t(x - m) + (1-t)(y - m)].$$

Or $t(x - m) + (1 - t)(y - m) = tx + (1 - t)y - m = 0$ par définition de m . Le crochet est nul, et il reste $tf(x) + (1 - t)f(y) \geq f(m)$: c'est l'inégalité de convexité. $\square$

Théorème

Caractérisation par la dérivée seconde. Soit f deux fois dérivable sur un intervalle I . Alors f est convexe sur I si et seulement si $f'' \geq 0$ sur I .

C'est le corollaire immédiat du point 2 : « f' croissante » équivaut à « $(f')' \geq 0$ », d'après la caractérisation des variations par le signe de la dérivée établie à la section 15. C'est le critère qu'on emploie en pratique, mais il faut se souvenir qu'il est **plus faible** que la définition, puisqu'il exige deux dérivations. Une fonction peut être convexe sans être dérivable : la valeur absolue l'est sur $\mathbb{R}$.

Convexes sur leur domaine : $x \mapsto x^2$, $\exp$, $x \mapsto |x|$, $x \mapsto \frac{1}{x}$ sur $]0, +\infty[$.
Concaves sur leur domaine : $\ln$, $\sqrt{\cdot}$, $\sin$ sur $[0, \pi]$, $x \mapsto -x^2$.

24.4 Points d'inflexion, et les deux pièges

Définition

Soit f définie au voisinage d'un point a . Le point d'abscisse a est un **point d'inflexion** de la courbe de f lorsque f **change de convexité** en a : convexe d'un côté, concave de l'autre. Lorsque f est deux fois dérivable au voisinage de a , cela revient à dire que f'' **change de signe** en a .

C'est bien le **changement de signe** qui définit la notion, et non l'annulation. Les deux pièges qui suivent sont posés à chaque partiel, et ils tombent tous les deux dans cette faille.

Remarque

Piège 1 : $f''(a) = 0$ ne suffit pas. Pour $f(x) = x^4$, on a $f''(x) = 12x^2$, qui s'annule en 0 — mais reste positive de part et d'autre. La fonction est convexe sur $\mathbb{R}$ tout entier, et l'origine n'est **pas** un point d'inflexion. La rédaction correcte est toujours en deux temps : on résout $f''(x) = 0$ pour trouver les **candidats**, puis on étudie le **signe de f'' autour de chaque candidat pour trancher**. Sauter la seconde étape, c'est répondre à côté de la question.

Piège 2 : une inflexion peut exister là où f'' n'existe pas. La fonction $x \mapsto x^{1/3}$, définie sur $\mathbb{R}$ par $x \mapsto \sqrt[3]{x}$, est concave sur $]-\infty, 0]$ et convexe sur $[0, +\infty[$: l'origine est un point d'inflexion, avec une tangente **verticale**. Or $f''(0)$ n'existe pas, et $f'(0)$ non plus. La dérivée seconde est un outil de recherche des inflexions, pas leur définition, et les points où elle n'existe pas doivent être examinés séparément.

La position par rapport à la tangente, traitée à la section 23 par les développements limités, donne le même diagnostic sous un autre angle : un premier terme non nul de degré **impair** au-delà de l'ordre 1 signale une traversée, donc une inflexion. Les deux méthodes doivent donner le même résultat, et c'est un bon contrôle. L'exercice 32 les confronte sur les mêmes fonctions.

24.5 Démontrer une inégalité par convexité

Méthode

Deux positions, deux inégalités. Une fois la convexité (ou la concavité) établie par le signe de f'' , on dispose de deux inégalités gratuites, valables **globalement** sur l'intervalle.

- **Par la tangente**, en un point a bien choisi : $f(x) \geq f(a) + f'(a)(x-a)$ si f est convexe, inégalité renversée si f est concave. C'est la méthode quand l'inégalité à démontrer compare f à une expression **affine**, et l'on choisit a au point où les deux membres sont égaux.
- **Par la corde**, sur un segment $[x_1, x_2]$: f est au-dessous de la corde si f est convexe, au-dessus si elle est concave. C'est la méthode quand l'inégalité n'est valable que **sur un segment** et devient une égalité aux deux bouts.

Le choix entre les deux se lit dans l'énoncé : une inégalité valable partout vient de la tangente, une inégalité valable sur un segment vient de la corde.

Exemple

Démontrer que $\sin x \geq \frac{2x}{\pi}$ **pour tout** $x \in \left[0, \frac{\pi}{2}\right]$. L'inégalité est valable sur un segment et devient une égalité aux deux extrémités, en 0 et en $\frac{\pi}{2}$: c'est la corde.

Concavité. La fonction $\sin$ est deux fois dérivable sur $\mathbb{R}$ et $\sin'' = -\sin$. Sur $\left[0, \frac{\pi}{2}\right]$, on a $\sin x \geq 0$, donc $\sin'' \leq 0$: la fonction $\sin$ est **concave** sur ce segment.

La corde. Les deux extrémités de la courbe sur ce segment sont les points de coordonnées $(0; 0)$ et $\left(\frac{\pi}{2}; 1\right)$. La corde qui les joint a pour pente $\frac{1-0}{\pi/2-0} = \frac{2}{\pi}$ et passe par l'origine : son équation est $y = \frac{2x}{\pi}$.

Conclusion. Une fonction concave est au-dessus de ses cordes, donc pour tout $x \in \left[0, \frac{\pi}{2}\right]$ on a $\sin x \geq \frac{2x}{\pi}$. $\square$

Contrôle. En $x = \frac{\pi}{4}$, l'inégalité annonce $\frac{\sqrt{2}}{2} \geq \frac{1}{2}$, soit $0,707 \geq 0,5$: cohérent. Et hors du segment, en $x = \pi$, elle serait fausse — $0 \geq 2$ — ce qui rappelle que le domaine fait partie du résultat.

Par la tangente, on obtient de la même façon les deux inégalités les plus employées de l'analyse : $e^x \geq 1 + x$ pour tout réel x (tangente en 0 à une fonction convexe) et $\ln x \leq x - 1$ pour tout $x > 0$ (tangente en 1 à une fonction concave). Elles avaient été admises dans le chapitre sur les limites et la continuité ; les voici démontrées en deux lignes chacune.

Remarque

Selon les facs : l'inégalité de Jensen à n points. Certaines maquettes prolongent l'inégalité de convexité à un nombre quelconque de points, d'autres s'arrêtent à deux. L'énoncé est le suivant : si f est convexe sur I , si $x_1, \dots, x_n$ sont dans I et si $\lambda_1, \dots, \lambda_n$ sont des réels positifs de somme 1, alors

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

La démonstration est une récurrence sur n : on regroupe les $n - 1$ derniers points en un seul barycentre, ce qui ramène au cas de deux points, à condition de bien renormaliser les coefficients — c'est le seul passage délicat.

L'application la plus célèbre est l'**inégalité arithmético-géométrique** : en appliquant Jensen à la fonction concave $\ln$, avec $\lambda_i = \frac{1}{n}$, on obtient

$$\sqrt[n]{x_1 x_2 \cdots x_n} \leq \frac{x_1 + x_2 + \cdots + x_n}{n}$$

pour tous réels strictement positifs. La moyenne géométrique ne dépasse jamais la moyenne arithmétique, et c'est la concavité du logarithme qui le dit.

25 25. La méthode de Newton, et les erreurs qui coûtent des points

25.1 La méthode de Newton

Remarque

Selon les facs. La méthode de Newton figure dans l'UE d'analyse de certaines maquettes et dans l'UE d'analyse numérique d'autres, parfois au second semestre. Elle est traitée ici parce qu'elle est la plus belle application des trois outils du chapitre — la tangente, l'inégalité des accroissements finis et la formule de Taylor-Lagrange — et parce qu'elle est, dans la vie réelle, la façon dont votre calculatrice extrait une racine. Que votre université la place ici ou ailleurs ne change rien à ce qu'elle apprend.

Le problème est le suivant : une équation $f(x) = 0$ admet une solution r dans un intervalle, on l'a établi par le théorème des valeurs intermédiaires, et l'on veut maintenant la **calculer**, ou plutôt en obtenir autant de décimales qu'on veut.

Le principe tient en une phrase : là où l'équation $f(x) = 0$ est difficile, l'équation de la tangente est facile, puisqu'elle est affine. On remplace donc la courbe par sa tangente en un point x_n , on résout l'équation affine, et l'on recommence à partir de la solution obtenue.

Écrivons-le. La tangente à la courbe de f au point d'abscisse x_n a pour équation

$$y = f(x_n) + f'(x_n)(x - x_n).$$

Elle coupe l'axe des abscisses au point où $y = 0$, c'est-à-dire lorsque $f'(x_n)(x - x_n) = -f(x_n)$. Si $f'(x_n) \neq 0$, cette équation a une solution unique, et c'est elle qu'on appelle x_{n+1} :

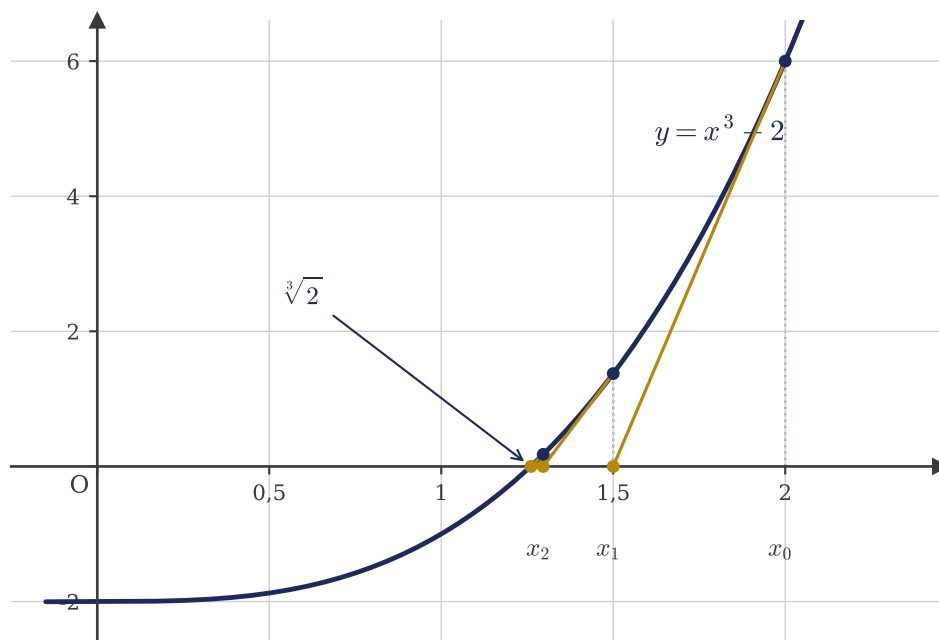
Propriété

Suite de Newton. Soit f dérivable et x_0 un réel donné. La **suite de Newton** associée est définie par

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)},$$

sous réserve que $f'(x_n) \neq 0$ à chaque étape.

La formule n'a donc rien de mystérieux : **elle est l'abscisse du point où la tangente coupe l'axe**, et il faut savoir la retrouver en deux lignes plutôt que la réciter. La figure ci-dessous montre trois itérations sur la fonction $f(x) = x^3 - 2$, dont l'unique racine réelle est $\sqrt[3]{2}$.



25.2 Les hypothèses de convergence

Propriété

Un jeu d'hypothèses suffisant. Soit f de classe $\mathcal{C}^2$ sur un segment $[a, b]$ telle que

1. $f(a)$ et $f(b)$ sont de signes contraires, ce qui assure l'existence d'une racine r dans $]a, b[$ par le théorème des valeurs intermédiaires ;
2. f' ne s'annule pas sur $[a, b]$, ce qui rend f strictement monotone et donc la racine **unique** ;
3. f'' garde un signe constant sur $[a, b]$, ce qui signifie que f y est convexe, ou concave, mais ne change pas ;
4. le point de départ x_0 est choisi **du bon côté**, c'est-à-dire dans celle des deux moitiés de $[a, b]$ où $f(x_0)$ et f'' ont le même signe.

Alors la suite de Newton est bien définie, monotone à partir de x_1 , et elle converge vers r .

L'hypothèse 4 est celle qu'on oublie, et c'est elle qui fait tout. Sa lecture géométrique est limpide : quand la courbe est convexe, elle est **au-dessus de toutes ses tangentes** — c'est la caractérisation démontrée à la section 24 — donc la tangente coupe l'axe **entre** le point courant et la racine, du côté où f est du même signe que f'' . La suite reste alors du même côté de r et progresse vers elle sans jamais la dépasser. Du mauvais côté, ce contrôle n'existe plus, et le premier saut peut envoyer x_1 n'importe où.

25.3 La vitesse : le nombre de décimales exactes double

Théorème

Convergence quadratique. Sous les hypothèses ci-dessus, en posant

$$m_1 = \inf_{[a,b]} |f'| > 0, \quad M_2 = \sup_{[a,b]} |f''|, \quad K = \frac{M_2}{2m_1},$$

on a, pour tout n ,

$$|x_{n+1} - r| \leq K |x_n - r|^2.$$

Démonstration, par Taylor-Lagrange. Appliquons la formule de Taylor-Lagrange à l'ordre 1 entre x_n et r : il existe c_n strictement compris entre ces deux points tel que

$$f(r) = f(x_n) + f'(x_n)(r - x_n) + \frac{(r - x_n)^2}{2} f''(c_n).$$

Comme $f(r) = 0$, on divise par $f'(x_n)$, qui n'est pas nul :

$$0 = \frac{f(x_n)}{f'(x_n)} + (r - x_n) + \frac{(r - x_n)^2}{2} \cdot \frac{f''(c_n)}{f'(x_n)}.$$

Or, par définition de la suite, $x_{n+1} - r = x_n - \frac{f(x_n)}{f'(x_n)} - r$, c'est-à-dire l'opposé des deux premiers termes ci-dessus. Il reste donc

$$x_{n+1} - r = \frac{f''(c_n)}{2f'(x_n)} (x_n - r)^2,$$

et en majorant le numérateur par M_2 , en minorant $|f'(x_n)|$ par m_1 , on obtient l'inégalité annoncée. $\square$

Ce que cette inégalité dit vraiment. Posons $e_n = K|x_n - r|$. L'inégalité se réécrit $e_{n+1} \leq e_n^2$. Si l'on a réussi à démarrer avec $e_0 \leq \frac{1}{2}$, une récurrence immédiate donne

$$e_n \leq \left(\frac{1}{2}\right)^{2^n},$$

et l'exposant, lui, double à chaque étape. En clair : **le nombre de décimales exactes double à chaque itération.** C'est ce qui distingue radicalement Newton d'une dichotomie, laquelle gagne une décimale toutes les trois ou quatre étapes.

Exemple

Newton sur $f(x) = x^3 - 2$, à partir de $x_0 = 2$. On a $f'(x) = 3x^2$, donc

$$x_{n+1} = x_n - \frac{x_n^3 - 2}{3x_n^2} = \frac{2}{3}x_n + \frac{2}{3x_n^2}.$$

La racine est $r = \sqrt[3]{2} \simeq 1,2599210$. Les hypothèses sont vérifiées sur $[1, 2]$: $f(1) = -1 < 0 < 6 = f(2)$, $f'(x) = 3x^2 \geq 3 > 0$, et $f''(x) = 6x > 0$; le point $x_0 = 2$ est du bon côté puisque $f(2) > 0$ et $f'' > 0$.

n	x_n	$x_n - \sqrt[3]{2}$
0	2	$7,4 \times 10^{-1}$
1	$\frac{3}{2}$	$2,4 \times 10^{-1}$
2	$\frac{35}{27}$	$3,6 \times 10^{-2}$
3	1,2609322...	$1,0 \times 10^{-3}$
4	1,2599219...	$8,1 \times 10^{-7}$

Lisez la colonne de droite : 10^{-1} , puis 10^{-2} , puis 10^{-3} , puis 10^{-7} . Les exposants sautent, et le facteur de passage est bien le carré du précédent : $0,79 \times (1,0 \times 10^{-3})^2 \simeq 8,1 \times 10^{-7}$, où 0,79 est la valeur de $\frac{f''(r)}{2f'(r)} = \frac{1}{r}$. **Quatre itérations donnent six décimales exactes**, à partir d'un point de départ franchement mauvais.

25.4 Les trois modes d'échec

C'est le point que presque aucun cours ne dit, et qui explique pourquoi les hypothèses ci-dessus ne sont pas décoratives. **Sans elles, Newton échoue, et il échoue de trois façons distinctes.**

Premier mode : la divergence. Prenons $f = \text{Arctan}$, dont l'unique racine est 0, et partons de $x_0 = 2$. Comme $\text{Arctan}'(x) = \frac{1}{1+x^2}$, la relation s'écrit $x_{n+1} = x_n - (1+x_n^2) \text{Arctan}(x_n)$. On obtient successivement $x_1 \simeq -3,54$, puis $x_2 \simeq 13,95$, puis $x_3 \simeq -279,3$: la suite s'éloigne de plus en plus vite. La raison est visible sur la courbe : loin de 0, l'Arctan est presque plate, donc sa tangente coupe l'axe très loin. Le même calcul depuis $x_0 = 1$ converge très bien. **Le point de départ n'est pas un détail.**

Deuxième mode : le cycle. Prenons $f(x) = x^3 - 2x + 2$ et $x_0 = 0$. Alors $f'(x) = 3x^2 - 2$, et

$$x_1 = 0 - \frac{2}{-2} = 1, \quad x_2 = 1 - \frac{1}{1} = 0.$$

La suite vaut 0, 1, 0, 1, 0, ... indéfiniment. Elle ne diverge pas, elle ne converge pas : **elle tourne**. Aucun test du type « les valeurs restent bornées » ne détecte cette situation, et c'est ce qui la rend traître.

Troisième mode : la tangente horizontale. Si $f'(x_n) = 0$ pour un certain n , la tangente est horizontale, elle ne coupe pas l'axe des abscisses, et la suite **n'est plus définie**. Le cas le plus simple est $f(x) = x^2 - 2$ avec $x_0 = 0$: $f'(0) = 0$, et le calcul s'arrête à la première étape. En pratique, il suffit que $f'(x_n)$ soit très petit pour que x_{n+1} parte très loin, ce qui ramène au premier mode.

Un quatrième comportement, moins spectaculaire mais tout aussi gênant, est la **convergence vers la mauvaise racine** : quand f en possède plusieurs, un mauvais départ peut faire converger la suite vers une racine qui n'est pas celle qu'on cherchait. L'exercice 33 fait construire ces quatre situations, et surtout vérifier, dans chaque cas, **laquelle des quatre hypothèses est en défaut** : c'est ce va-et-vient entre l'échec et l'hypothèse manquante qui fait comprendre à quoi servent les théorèmes.

25.5 Les erreurs qui coûtent des points au partiel

Ce qui suit n'est pas un catalogue de fautes exotiques : ce sont les sept que les correcteurs voient à chaque session, et chacune est le symptôme d'une hypothèse sautée.

1. « $f'(a) = 0$ donc f admet un extremum en a ». C'est la réciproque du lemme de Fermat, et elle est fautive : $x \mapsto x^3$ a une dérivée nulle en 0 sans y avoir d'extremum. Le lemme dit qu'un extremum **intérieur** d'une fonction dérivable annule la dérivée, et rien d'autre. La conclusion correcte se rédige avec le **signe de f' de part et d'autre**, ou avec le premier terme non nul du développement limité, qui donne la réponse immédiatement.

2. Appliquer Rolle sans vérifier la continuité aux bornes. Le théorème demande la continuité sur le **segment fermé** $[a, b]$ et la dérivabilité sur l'**ouvert** $]a, b[$. Cette dissymétrie n'est pas un ornement : elle est ce qui permet d'appliquer Rolle à $x \mapsto \sqrt{1-x^2}$ sur $[-1, 1]$, où la dérivée n'existe pas aux deux bouts. Une copie qui écrit « f est dérivable sur $[a, b]$ » à propos d'une telle fonction énonce quelque chose de faux, et perd le point d'hypothèse.

3. « $f' > 0$ sur $\mathcal{D}$ donc f est croissante sur $\mathcal{D}$ » quand $\mathcal{D}$ n'est pas un intervalle. La caractérisation des variations par le signe de la dérivée, à la section 15, se démontre par l'égalité des accroissements finis, laquelle exige de pouvoir relier deux points **en restant dans le domaine**. Sur $\mathbb{R}^*$, la fonction inverse a une dérivée strictement négative partout et n'est pourtant pas décroissante sur $\mathbb{R}^*$, puisque $-1 < 1$ et $f(-1) = -1 < 1 = f(1)$. La rédaction correcte énonce le résultat **sur chacun des intervalles**, jamais sur leur réunion.

4. Dériver un développement limité. Traité à la section 21, et rappelé ici parce que c'est la faute la plus mécanique : on primitive un DL, on ne le dérive jamais. Le témoin $x \mapsto x^2 \sin(1/x)$ montre que l'opération peut produire une égalité tout simplement fautive.

5. « $f''(a) = 0$ donc le point d'abscisse a est un point d'inflexion ». Faux, et $x \mapsto x^4$ en 0 le montre. L'annulation fournit un **candidat** ; c'est le changement de signe de f'' qui conclut. Symétriquement, ne pas oublier d'examiner les points où f'' n'existe pas, qui peuvent être des inflexions.

6. Invoquer Taylor-Young à un ordre que la régularité ne permet pas. Écrire un développement à l'ordre 4 pour une fonction dont l'énoncé ne garantit que la classe $\mathcal{C}^2$ est une faute d'hypothèse, au même titre qu'appliquer Rolle à une fonction discontinue. La bonne rédaction annonce d'abord la régularité, et l'ordre ensuite : « f étant de classe $\mathcal{C}^4$ au voisinage de 0, la formule de Taylor-Young à l'ordre 4 donne... ».

7. Utiliser la règle de L'Hôpital sans vérifier la forme indéterminée. La règle ne s'applique qu'à une forme $\frac{0}{0}$ ou $\frac{\infty}{\infty}$, avec un dénominateur dérivé qui ne s'annule pas au voisinage épointé du point. Appliquée à $\frac{x+1}{x^2+1}$ en 0, dont la limite est 1, elle donnerait $\frac{1}{2x} \rightarrow \infty$: le résultat est faux, et la seule chose qui manquait était la vérification initiale. **Chaque application successive exige une nouvelle vérification.**

Une huitième, transversale et plus grave que les sept autres : **conclure sans avoir vérifié le domaine**. Une inégalité démontrée par convexité sur un segment ne vaut que sur ce segment, un développement limité en 0 ne dit rien en $x = 5$, une asymptote en $+\infty$ ne présume de rien en $-\infty$. L'exercice 34 propose six copies à corriger, contenant chacune une de ces erreurs, et demande de dire non pas « c'est faux » mais **quelle hypothèse manque** : c'est le seul niveau de diagnostic qui serve à quelque chose en partiel.

25.6 Quel outil pour quelle question

La question posée	L'outil
Montrer une inégalité entre deux fonctions	étude de $f - g$ par le signe de la dérivée (section 15), ou convexité (section 24)
Majorer un écart, chiffrer une erreur	inégalité des accroissements finis (section 13), ou inégalité de Taylor-Lagrange (section 19)
Calculer une limite indéterminée	développement limité (section 23), ou L'Hôpital si une dérivation suffit
Obtenir un équivalent, une vitesse	premier terme non nul du développement limité (section 23)
Situer la courbe par rapport à sa tangente	premier terme non nul de degré ≥ 2 (section 23), ou convexité (section 24)
Trouver une asymptote et sa position	développement en $\frac{1}{x}$ à l'infini (section 23)
Compter les racines d'une équation	variations plus théorème de la bijection, ou Rolle entre deux racines (section 11)
Approcher numériquement une solution	méthode de Newton (section 25)

Ce tableau est le résumé de tout le chapitre, et il se lit dans les deux sens. De gauche à droite, il indique l'outil devant une question. De droite à gauche, il rappelle ce que chaque théorème sert à faire, ce qui est la meilleure façon de vérifier qu'on l'a compris : un théorème dont on ne sait pas à quelle question il répond n'est pas encore su.