CHAPITRE 3

DÉRIVABILITÉ, TH. DES ACCROISSEMENTS FINIS, SENS DE VARIATION

3.1. Fonctions dérivables

Définition 3.1 **(de la dérivée comme limite des taux d'accroissement)**
Soient $f : \mathbb{R} \to \mathbb{R}$ une fonction, D son domaine de définition, et $a \in D$. On suppose que D contient un intervalle $]a-\delta, a+\delta[$ (resp. $[a, a+\delta[$, resp. $]a-\delta, a]$) pour un certain $\delta > 0$. Alors, on dit que f est *dérivable* (resp. *dérivable à droite*, resp. *dérivable à gauche*) en a si la condition suivante est vérifiée :

($*_a$) La fonction $g(x) = \dfrac{f(x) - f(a)}{x - a}$ admet en a une limite (resp. une limite à droite, resp. une limite à gauche) ℓ.

Introduisant la fonction $\tau(h) = \dfrac{f(a+h) - f(a)}{h}$, dont le domaine de définition contient $]-\delta, \delta[\setminus\{0\}$ (resp. $]0, \delta[$, resp. $]-\delta, 0[$), ceci équivaut à :[(1)]

($*_0$) La fonction $\tau(h) = \dfrac{f(a+h) - f(a)}{h}$ admet en 0 une limite (resp. une limite à droite, resp. une limite à gauche) ℓ.

Cette limite est alors notée $f'(a)$ (resp. $f'_d(a)$, resp. $f'_g(a)$) et appelée la dérivée (resp. dérivée à droite, resp. dérivée à gauche) de f en a.

Remarque 3.2. — (1) Il se peut que $f'_d(a)$ et $f'_g(a)$ existent et soient distinctes. Dans ce cas, f n'est pas dérivable en a. Par exemple, la fonction $f(x) = |x|$ admet en $a = 0$ une dérivée à droite $f'_d(0)$ égale à 1 et une dérivée à gauche $f'_g(0)$ égale à -1.

(2) Si $f'_d(a)$ et $f'_g(a)$ existent et sont égales, alors f est dérivable en a. Par exemple, la fonction f définie par $f(x) = x$ si $x \leq 0$ et $f(x) = x + x^2$ si $x \geq 0$ admet en $a = 0$ une dérivée à gauche $f'_g(0) = 1$ et aussi une dérivée à droite $f'_d(0) = 1$, donc est dérivable en 0.

La définition précédente est équivalente à la définition suivante :

Définition 3.3 **(de la dérivabilité par l'existence d'un** DL_1**)**
(Q) Soient $f : \mathbb{R} \to \mathbb{R}$ une fonction, D son domaine de définition, et $a \in D$. On suppose qu'il existe un intervalle ouvert $I =]-\delta, \delta[$, avec $\delta > 0$, tel que $a+h \in D$ pour tout $h \in I$.[(2)] Alors f est *dérivable* en a si et seulement si il existe deux réels A, ℓ et une fonction $\varepsilon : I \to \mathbb{R}$ continue et nulle en 0, tels que

$$(**) \qquad \boxed{\forall h \in I, \quad f(a+h) = A + \ell h + h\varepsilon(h).}$$

Dans ce cas, $A = f(a)$ et $\ell = f'(a)$.

[(1)] L'équivalence provient de ce que $\tau(h) = g(a+h)$, i.e. $\tau = g \circ u$ et $g = \tau \circ v$, où u (resp. v) est la fonction $h \mapsto a + h$ (resp. $x \mapsto x - a$).

[(2)] C.-à-d. tel que D contienne l'intervalle ouvert $]a-\delta, a+\delta[$.

Démonstration de l'équivalence des deux définitions. — (1) Supposons f dérivable au sens de la définition 3.1. Alors la fonction $\tau : I \setminus \{0\} \to \mathbb{R}$, $h \mapsto \dfrac{f(a+h)-f(a)}{h}$ se prolonge par continuité en 0 en posant $\tau(0) = f'(a)$, et alors la fonction $\varepsilon(h) = \tau(h) - f'(a)$ est continue et nulle en 0, et l'on a bien $f(a+h) - f(a) = h\tau(h) = hf'(a) + h\varepsilon(h)$ pour tout $h \in I$.

(2) Réciproquement, supposons $(**)$ vérifié, avec ε continue et nulle en 0. Prenant $h = 0$, on obtient $A = f(a)$. Alors pour tout $h \in I \setminus \{0\}$ on a $\tau(h) = \ell + \varepsilon(h)$ et ceci tend vers ℓ quand h tend vers 0. Donc f est dérivable au sens de la définition 3.1, et l'on a de plus $\ell = f'(a)$. □

Remarque 3.4. — En remplaçant la fonction $\varepsilon(h)$ par la fonction $u(h) = \ell + \varepsilon(h)$ on peut récrire la définition 3.3 sous la forme plus condensée suivante : f est dérivable en a si et seulement si il existe une fonction $u : I \to \mathbb{R}$ continue en 0 telle que

$$(**_{\text{bis}}) \qquad \boxed{\forall h \in I, \quad f(a+h) = f(a) + hu(h)}$$

et dans ce cas $u(0) = f'(a)$.

Remarque 3.5. — On a des définitions analogues à 3.3 et 3.4 pour « dérivable à droite (resp. à gauche) », en remplaçant I par $[0, \delta[$ (resp. $]-\delta, 0]$) et la condition « ε continue en 0 » par ε continue à droite (resp. à gauche) en 0.

Corollaire 3.6. — *Si f est dérivable en a alors elle est continue en a.*

Démonstration. — On peut récrire (encore !) la définition 3.3 sous la forme : il existe un intervalle ouvert J centré en a et une application $\varepsilon : J \to \mathbb{R}$, continue et nulle en a, tels que :

$$f(x) = f(a) + (x-a)f'(a) + (x-a)\varepsilon(x).$$

pour tout $x \in J$. Par conséquent, $f(x)$ tend vers $f(a)$ quand x tend vers a, donc f est continue en a. □

***Remarque 3.7* (Interprétation graphique de la dérivabilité)**

Supposons f définie sur un intervalle ouvert I contenant a et considérons dans le plan $\mathbb{R}^2 = \{(t, y) \mid t, y \in \mathbb{R}\}$ son graphe :

$$\Gamma_f = \{(t, f(t)) \mid t \in I\}.$$

Pour tout $x \in I \setminus \{a\}$, les points $(a, f(a))$ et $(x, f(x))$ sont distincts donc définissent une droite D_x, dont la pente est $p_x = \dfrac{f(x)-f(a)}{x-a}$ et dont l'équation est donc $y - f(a) = p_x(x-a)$, i.e. $y = f(a) + p_x(x-a)$. La dérivabilité en a signifie que quand x tend vers a, les pentes p_x tendent vers une limite $\ell = f'(a)$, et donc les droites D_x « tendent » vers la droite d'équation $y = f(a) + f'(a)(x-a)$, qui est appelée la *tangente* à Γ_f au point a.

***Définitions 3.8* (Fonction dérivée et dérivées successives)**

Soit I un intervalle ouvert $\neq \varnothing$ et soit f une application $I \to \mathbb{R}$.

(i) On dit que f est *dérivable sur I* si elle est dérivable en chaque point $a \in I$. Dans ce cas, on obtient une application $I \to \mathbb{R}$, $x \mapsto f'(x)$, appelée la fonction dérivée de f et notée f'.

(ii) Si la fonction $f' : I \to \mathbb{R}$ est elle-même dérivable, sa dérivée $(f')'$ est notée f'' et appelée la *dérivée seconde* de f, et l'on dit que f est *deux fois dérivable sur I*. Si f'' est encore dérivable, sa dérivée $(f'')'$ est notée f''' et appelée la *dérivée troisième* de f, et l'on dit que f est *trois fois dérivable sur I*. On définit ainsi par récurrence, si elle existe, la dérivée n-ième de f, notée $f^{(n)}$, en posant $f^{(n)} = (f^{(n-1)})'$.[3] On dit alors que f est *n fois dérivable sur I.*

Remarque et définition 3.9. — En ne considérant ci-dessous que des applications $I \to \mathbb{R}$, on a donc des inclusions :

$$\{\text{fonctions continues}\} \supset \{\text{fonctions dérivables}\} \supset \{\text{fonctions deux fois dérivables}\} \supset \cdots \supset \{\text{fonctions } n \text{ fois dérivables}\} \supset \cdots$$

On définit aussi des classes intermédiaires de fonctions en disant, pour chaque $n \in \mathbb{N}^*$, que f est de *classe C^n* si elle est n fois dérivable sur I et si de plus la fonction dérivée $f^{(n)}$ est continue sur I. On convient que « classe C^0 » signifie « continue ». D'après le corollaire 3.6, si f est n fois dérivable alors la dérivée $(n-1)$-ième est continue, donc on a les inclusions suivantes :[4]

$$\left\{\begin{matrix}\text{fonctions}\\ C^0\end{matrix}\right\} \supset \left\{\begin{matrix}\text{fonctions}\\ \text{dérivables}\end{matrix}\right\} \supset \left\{\begin{matrix}\text{fonctions}\\ C^1\end{matrix}\right\} \supset \left\{\begin{matrix}\text{fonctions deux}\\ \text{fois dérivables}\end{matrix}\right\} \supset \left\{\begin{matrix}\text{fonctions}\\ C^2\end{matrix}\right\} \supset \cdots \supset \left\{\begin{matrix}\text{fonctions } n\\ \text{fois dérivables}\end{matrix}\right\} \supset \left\{\begin{matrix}\text{fonctions}\\ C^n\end{matrix}\right\} \supset \cdots$$

Enfin, on dit que f est *de classe C^∞* sur I si elle est n fois dérivable, pour tout $n \in \mathbb{N}$. D'après le corollaire 3.6, ceci implique que chaque $f^{(n)}$ est continue (puisque dérivable), donc que f est de classe C^n pour tout $n \in \mathbb{N}$.

Remarque 3.10. — (1) On verra que la plupart des fonctions usuelles sont de classe C^∞ sur tout intervalle ouvert où elles sont définies. C'est le cas des fonctions polynômes $x \mapsto P(x)$ et des fonctions rationnelles $x \mapsto P(x)/Q(x)$, des fonctions exponentielle et logarithme, sin et cos, etc.

(2) On peut montrer que la fonction f définie par $f(x) = x^2 \sin(1/x)$ si $x \neq 0$ et $f(0) = 0$ est dérivable sur $\mathbb{R}$, mais sa dérivée f' n'est pas continue en 0. (Exercice !)

3.2. Règles de calcul : composition et opérations algébriques

Commençons par le théorème ci-dessous, dont la *démonstration* utilise la 2ème définition de la dérivabilité.

(Q) ***Théorème 3.11* (Dérivée d'une fonction composée).** — *Soient I, J deux intervalles de $\mathbb{R}$ et $f : I \to \mathbb{R}$ et $g : J \to \mathbb{R}$ deux applications, telles que $f(I) \subset J$. Soit $a \in I$. On suppose que :*[5]

(1) *f est dérivable (resp. dérivable à droite, resp. à gauche) en a,*

(2) *g est dérivable en $f(a)$.*

Alors l'application $g \circ f$ est dérivable (resp. dérivable à droite, resp. à gauche) en a, de dérivée

$$\boxed{(g \circ f)'(a) = g'(f(a))\, f'(a)}$$

resp. $(g \circ f)'_d(a) = g'(f(a))\, f'_d(a)$, resp. $(g \circ f)'_g(a) = g'(f(a))\, f'_g(a)$.

Démonstration. — Faisons la démonstration dans le cas où f est dérivable en a. (Le cas où f est seulement dérivable à droite (resp. à gauche) en a est analogue.) D'après la remarque 3.4, la dérivabilité de f en a et de g en $b = f(a)$ peut se récrire de façon plus concise sous la forme suivante : il existe des voisinages U et V de 0 et des applications $u : U \to \mathbb{R}$ et $v : V \to \mathbb{R}$ continues en 0, vérifiant $u(0) = f'(a)$ et $v(0) = g'(b)$ et telles que :

$$\begin{aligned} \forall h \in U, \qquad & f(a+h) = f(a) + hu(h), \\ \forall k \in V, \qquad & g(b+k) = g(b) + kv(k). \end{aligned}$$

[3] Comme celle-ci en obtenue en dérivant n fois f, c'est aussi la dérivée $(n-1)$-ième de f'.

[4] Pour abréger, on a écrit « f est C^n » au lieu de « f est de classe C^n ».

[5] Pour la question de cours, il suffira d'énoncer le théorème dans le cas où f est dérivable en a.

Comme la fonction $h \mapsto hu(h)$ est continue et nulle en 0, il existe un voisinage U_0 de 0 contenu dans U tel que, pour tout $h \in U_0$, on ait $hu(h) \in V$ et donc :

$$g\big(f(a+h)\big) = g\big(f(a) + hu(h)\big) = g(f(a)) + hu(h)\, v\big(hu(h)\big).$$

La fonction $w : h \mapsto u(h)v\big(hu(h)\big)$ est continue en 0, d'après les théorèmes sur la composée et le produit de fonctions continues, et l'on a $w(0) = f'(a)v(0) = f'(a)g'(b)$. D'après la remarque 3.4 à nouveau, ceci montre que $g \circ f$ est dérivable en a, de dérivée $g'(b)f'(a) = g'(f(a))f'(a)$. □

Remarque 3.12 **(Une identité remarquable).** — La formule suivante est utile à connaître. Pour tout $n \in \mathbb{N}^*$ et $x, a \in \mathbb{R}$, on a :

$$(\star) \qquad x^n - a^n = (x-a)(x^{n-1} + x^{n-2}a + \cdots + xa^{n-2} + a^{n-1}).$$

En effet, en développant le terme de droite on trouve :

$$\begin{aligned} &x^n + x^{n-1}a + \cdots + x^2a^{n-2} + xa^{n-1} \\ &\quad - x^{n-1}a - \cdots - x^2a^{n-2} - xa^{n-1} - a^n \end{aligned}$$

et tous les termes sauf les deux extrêmes se simplifient, i.e. il ne reste que $x^n - a^n$. Notons alors $P(x)$ le 2ème facteur du terme de droite, i.e. $P(x) = \sum_{i=0}^{n-1} x^{n-1-i}a^i$. (Noter qu'il y a n termes dans cette somme, l'indice i variant de 0 à $n-1$.)

Proposition 3.13. — *Pour tout $n \in \mathbb{N}^*$, la fonction $f_n : x \mapsto x^n$ est dérivable sur $\mathbb{R}$, de dérivée la fonction $x \mapsto nx^{n-1}$.*

Démonstration. — Ceci découle de la remarque précédente : pour $x \neq a$, on a

$$\frac{x^n - a^n}{x - a} = P(x)$$

et comme la fonction polynôme $P(x)$ est continue en a, on a $\lim_{x\to a} P(x) = P(a) = na^{n-1}$. Ceci montre que f_n est dérivable en a, de dérivée na^{n-1}. □

(Q) ***Théorème 3.14*** **(Sommes, produits et quotients de fonctions dérivables)** *Soit I un intervalle ouvert et soient f, g deux applications $I \to \mathbb{R}$, dérivables en un point $a \in I$.*

(i) *$f + g$ est dérivable en a, de dérivée $(f+g)'(a) = f'(a) + g'(a)$.*

(ii) *fg est dérivable en a, de dérivée $(fg)'(a) = f'(a)g(a) + f(a)g'(a)$.*

(iii) *Si $g(a) \neq 0$, alors $1/g$ et f/g sont dérivables en a, de dérivées :*

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

Par conséquent, si f, g sont dérivables sur I, alors $f + g$ et fg le sont aussi et l'on a

$$\boxed{(f+g)' = f' + g' \quad \text{et} \quad (fg)' = f'g + fg'.}$$

Si de plus g ne s'annule pas sur I, alors $1/g$ et f/g sont aussi dérivables sur I et l'on a

$$\boxed{\left(\frac{1}{g}\right)' = \frac{-g'}{g^2} \quad \text{et} \quad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}.}$$

Démonstration. — Notons $\tau_f(h) = \dfrac{f(a+h) - f(a)}{h}$ la fonction « taux d'accroissement de f en a », définie sur un intervalle épointé $]-\delta, \delta[\setminus\{0\}$, et définissons de même les taux d'accroissement τ_g, τ_{f+g} et τ_{fg}. On voit immédiatement que $\tau_{f+g} = \tau_f + \tau_g$, donc (i) découle du résultat sur la limite d'une somme (2.11 (i)).

D'autre part, on a :

$$\tau_{fg}(h) = \frac{(fg)(a+h)-(fg)(a)}{h} = \frac{f(a+h)-f(a)}{h}g(a+h) + f(a)\frac{g(a+h)-g(a)}{h}$$

et quand h tend vers 0, le terme de droite tend vers $f'(a)g(a) + f(a)g'(a)$, d'après les résultats sur les limites de produits et de sommes (2.11 (i) et (ii)).

Pour (iii), il suffit de montrer l'assertion concernant $1/g$, car celle pour f/g découlera alors de (ii). Comme $g(a) \neq 0$ et g est continue en a (puisque dérivable en a), alors g reste non nulle sur un intervalle $]a-\delta, a+\delta[$ et donc le taux d'accroissment $\tau_{1/g}$ est défini sur $]-\delta, \delta[\setminus\{0\}$. De plus, on a :

$$\tau_{1/g}(h) = \frac{\dfrac{1}{g(a+h)} - \dfrac{1}{g(a)}}{h} = \frac{g(a)-g(a+h)}{h}\,\frac{1}{g(a)g(a+h)}$$

et quand h tend vers 0 le terme de droite tend vers $-g'(a)\dfrac{1}{g(a)^2}$ (on utilise 2.11 (iii) pour dire que la limite en 0 de $1/g(a)g(a+h)$ est $1/g(a)^2$). Ceci montre que $1/g$ est dérivable en a, de dérivée $-g'(a)/g(a)^2$. □

Remarque 3.15. — On aurait aussi pu démontrer la proposition 3.13 par récurrence sur n, en utilisant la formule de dérivation d'un produit 3.14 (ii).

***Définition et remarque 3.16* (Degré d'un polynôme)**. — On dit qu'un polynôme $Q(x)$ est de *degré* $d \in \mathbb{N}$ s'il s'écrit $Q(x) = b_0 + b_1x + \cdots + b_dx^d$ avec $b_d \neq 0$. (Ainsi, un polynôme de degré 0 est une constante $b_0 \neq 0$, et le polynôme nul 0 n'a pas de degré.)

Dans ce cas, on peut montrer que l'équation $Q(x) = 0$ a au plus d solutions dans $\mathbb{R}$, i.e. que Q a au plus d racines dans $\mathbb{R}$ (éventuellement aucune). S'il en existe r et si on les ordonne $x_1 < \cdots < x_r$ alors le domaine de définition de la fonction $x \mapsto 1/Q(x)$ est la réunion des intervalles ouverts $]-\infty, x_1[$, $]x_1, x_2[$, …, $]x_r, +\infty[$.

Corollaire 3.17. — *Soit $P(x) = a_0 + a_1x + \cdots + a_nx^n$ une fonction polynomiale.*

(i) *$P(x)$ est dérivable sur $\mathbb{R}$, de dérivée $P'(x) = a_1 + 2a_2x + \cdots + na_nx^{n-1}$, qui est encore une fonction polynomiale. Par conséquent, P est de classe C^∞.*

(ii) *Soit Q une fonction polynomiale qui n'est pas la fonction nulle. Alors la fonction rationnelle $F : x \mapsto \dfrac{P(x)}{Q(x)}$ est dérivable en tout point de son domaine de définition $D_F = \{x \in \mathbb{R} \mid Q(x) \neq 0\}$ et pour tout $x \in D_F$ on a :*

$$F'(x) = \frac{P'(x)Q(x) - P(x)Q'(x)}{Q(x)^2},$$

qui est encore une fonction rationnelle dont le domaine de définition est D_F. Par conséquent, F est de classe C^∞ sur D_F.

Corollaire 3.18. — *Soit $n \in \mathbb{N}^*$.*

(i) *Soit f une fonction dérivable sur un intervalle ouvert I. Alors l'application $f^n : I \to \mathbb{R}$, $x \mapsto f(x)^n$ est dérivable, de dérivée $x \mapsto nf^{n-1}(x)f'(x)$.*

(ii) *La fonction $x \mapsto x^{-n} = \dfrac{1}{x^n} = \left(\dfrac{1}{x}\right)^n$ est dérivable sur $\mathbb{R}^*$, de dérivée $x \mapsto -nx^{-n-1}$.*

Démonstration. — (i) découle de la proposition 3.13 et du théorème 3.11 appliqué à f et à $g(x) = x^n$.

Prouvons (ii), en prenant pour I l'un des intervalles $]-\infty, 0[$ ou $]0, +\infty[$ et pour f la fonction $x \mapsto 1/x$. D'après le théorème 3.14 (iii), la fonction $x \mapsto 1/x$ est dérivable sur I, de dérivée $x \mapsto -1/x^2$. Par conséquent, d'après (i) la dérivée de $x \mapsto (1/x)^n$ est :

$$x \mapsto n\left(\frac{1}{x}\right)^{n-1}\left(\frac{-1}{x^2}\right) = \frac{-n}{x^{n+1}} = -nx^{-n-1}.$$

$\square$

3.3. Théorèmes de Rolle et des accroissements finis

***Définition 3.19* (Extremum local).** — Soient I un intervalle, f une application $I \to \mathbb{R}$, et c un point *intérieur* de I. On dit que f admet en c un *maximum local* (resp. un *minimum local*) s'il existe un intervalle ouvert non vide $J =]c-\delta, c+\delta[$ contenu dans I tel que pour tout $x \in J$ on ait $f(x) \leq f(c)$ (resp. $f(x) \geq f(c)$). Dans les deux cas, on dit que f admet en c un *extremum local*.

Exemple 3.20. — Soient $I = [-\pi, \pi]$ et $f : I \to \mathbb{R}$ l'application définie par $f(x) = \sin(x)$ si $x \leq 0$ et $f(x) = 2\sin(x/2)$ si $x \geq 0$. Comme $f(-\pi/2) = -1$ est $\leq f(x)$ pour tout $x \in]-\pi, 0[$, alors f admet en $a = -\pi/2$ un minimum local. Mais ce n'est pas un minimum global sur I puisque $f(\pi/4) = -2$.

***Proposition 3.21* (Condition nécessaire pour un extremum local)**

(Q) *Soient I un intervalle, f une application $I \to \mathbb{R}$ et c un point intérieur de I en lequel f admet un extremum local. Si f est dérivable en c, alors $f'(c) = 0$.*

Démonstration. — Supposons que f admette en c un maximum local. (La démonstration est analogue dans le cas d'un minimum local.) Par hypothèse, il existe un intervalle ouvert non vide $J =]c-\delta, c+\delta[$ contenu dans I tel que pour tout $x \in J$ on ait $f(x) \leq f(c)$. Alors pour tout $x \in]c-\delta, c[$ (resp. $y \in]c, c+\delta[$) on a :

$$\frac{f(x) - f(c)}{x - c} \geq 0 \qquad \text{resp.} \qquad \frac{f(y) - f(c)}{y - c} \leq 0.$$

Par conséquent, d'après le théorème de passage à la limite pour les inégalités larges, on a à la fois $f'(c) \geq 0$ et $f'(c) \leq 0$, d'où $f'(c) = 0$. $\square$

Remarque 3.22. — La condition $f'(c) = 0$ est *nécessaire* pour que c soit un extremum local, mais elle n'est *pas suffisante*. Par exemple, la fonction $f(x) = x^3$ est dérivablesur $\mathbb{R}$, sa dérivée $f'(x) = 3x^2$ s'annule en $c = 0$, mais $c = 0$ n'est pas un extremum local de f.

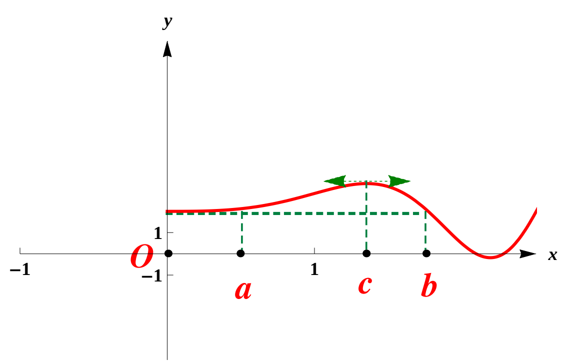

FIG. 1. Une illustration graphique du théorème de Rolle.

***Théorème 3.23* (de Rolle).** — *Soient $a < b$ dans $\mathbb{R}$ et soit f une fonction $\mathbb{R} \to \mathbb{R}$ continue sur l'intervalle fermé $[a, b]$, dérivable sur l'intervalle ouvert $]a, b[$ et telle que $f(a) = f(b)$. Alors il existe $c \in]a, b[$ tel que $f'(c) = 0$.*

Démonstration. — D'après le théorème des bornes atteintes, $f([a, b])$ est un intervalle fermé borné $[m, M]$. Si $m = M$ alors f est constante sur $[a, b]$ donc sa dérivée est nulle en tout point $c \in]a, b[$. On peut donc supposer $m < M$ et donc au moins l'un des deux, disons M, est distinct de $f(a) = f(b)$. Il existe $c \in [a, b]$ tel que $f(c) = M$ et nécessairement $c \in]a, b[$ puisque $f(c) = M$ est distinct de $f(a) = f(b)$. Alors f admet en c un maximal global (a fortiori local) donc $f'(c) = 0$ d'après la proposition précédente. $\square$

(Q) ***Théorème 3.24*** **(des accroissements finis).** — *Soient $a < b$ dans $\mathbb{R}$ et soit f une fonction $\mathbb{R} \to \mathbb{R}$ continue sur l'intervalle fermé $[a, b]$ et dérivable sur l'intervalle ouvert $]a, b[$. Alors il existe $c \in]a, b[$ tel que $f'(c) = \dfrac{f(b) - f(a)}{b - a}$.*

Démonstration. — On se ramène au théorème de Rolle en soustrayant à f la fonction linéaire qui vaut 0 en a et $f(b) - f(a)$ en b, c.-à-d. on considère la fonction auxiliaire

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

Alors g est, comme f, continue sur $[a, b]$ et dérivable sur $]a, b[$, et l'on a de plus $g(a) = f(a) = g(b)$. D'après le théorème de Rolle, il existe donc $c \in]a, b[$ tel que $g'(c) = 0$. Comme on a $g'(x) = f'(x) - \dfrac{f(b) - f(a)}{b - a}$ pour tout $x \in]a, b[$, on obtient donc $f'(c) = \dfrac{f(b) - f(a)}{b - a}$. $\square$

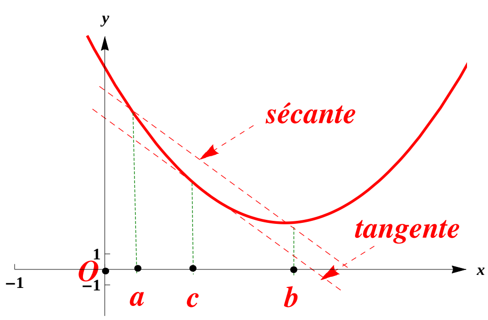

FIG. 2. Une illustration graphique du théorème des accroissements finis.

Remarque 3.25. — [6] Le théorème des accroissements finis est très important. Il permet, entre autre, de déterminer si une fonction dérivable f est croissante (ou décroissante) en étudiant le signe de f' (voir la section suivante).

Exercice 3.26. — Soit $P(x) = a_0 + a_1 x + \cdots + a_n x^n$ une fonction polynomiale de degré n. En utilisant le théorème de Rolle, montrer par récurrence sur n que l'équation $P(x) = 0$ a au plus n solutions dans $\mathbb{R}$.

3.4. Étude des variations d'une fonction

Définition 3.27 **(Fonctions croissantes, décroissantes, strictement monotones)**

Soit f une fonction $\mathbb{R} \to \mathbb{R}$ et soit $D = D_f$ son domaine de définition.

(i) On dit que f est croissante (resp. strictement croissante) si pour tout $x < y$ dans D, on a $f(x) \leq f(y)$ (resp. $f(x) < f(y)$).

(ii) On dit que f est décroissante (resp. strictement décroissante) si pour tout $x < y$ dans D, on a $f(x) \geq f(y)$ (resp. $f(x) > f(y)$).

[6] La figure ci-dessus et la précédente sont tirées du livre *Mathématiques L1* par C. David et S. Mustapha, à paraître (Dunod, printemps 2014) et nous ont été aimablement communiquées par C. David.

(iii) Enfin on dit que f est *strictement monotone* si elle est ou bien strictement croissante, ou bien strictement croissante.

Remarque 3.28. — Une fonction constante est à la fois croissante et décroissante (et réciproquement !).

***Théorème 3.29* (Fonctions dérivables croissantes ou décroissantes)**
Soient I un intervalle de $\mathbb{R}$ et f une application $I \to \mathbb{R}$ continue sur I et dérivable en tout point intérieur de I.

(i) *f est croissante (resp. décroissante) sur I si et seulement si $f'(x) \geq 0$ (resp. ≤ 0) pour tout point intérieur x de I.*

(i bis) *f est constante si et seulement si $f'(x) = 0$ pour tout point intérieur x de I.*

(ii) *SI pour tout point intérieur x de I on a $f'(x) > 0$ (resp. $f'(x) < 0$), ALORS f est strictement croissante (resp. strictement décroissante).*

Démonstration. — (i) Faisons la démonstration pour « croissante ». (Le cas décroissant est analogue.) L'implication $\Rightarrow$ est facile : supposons f croissante, soit a un point intérieur de I et soit V un intervalle ouvert centré en a et contenu dans I. Comme f est croissante, pour tout $x \in V \setminus \{a\}$ on a

$$\frac{f(x) - f(a)}{x - a} \geq 0$$

et donc $f'(a) \geq 0$ d'après le théorème de passage à la limite pour les inégalités larges.

Montrons l'implication $\Leftarrow$: soient $a < b$ dans I. D'après le théorème des accroissements finis, il existe c dans $]a, b[$, qui est donc un point intérieur de I, tel que $f(b) - f(a) = (b - a)f'(c)$, et ceci est ≥ 0 par hypothèse, donc $f(b) \geq f(a)$. Ceci montre que f est croissante.

(i bis) découle immédiatement de (i), ou bien se démontre directement, de façon analogue à (i).

Enfin, (iii) se démontre comme (ii) : soient $a < b$ dans I. D'après le théorème des accroissements finis, il existe c dans $]a, b[$, qui est donc un point intérieur de I, tel que $f(b) - f(a) = (b - a)f'(c)$, et ceci est > 0 par hypothèse, donc $f(b) > f(a)$. Ceci montre que f est strictement croissante. $\square$

Remarque 3.30. — Dans le point (iii), la condition $f'(x) > 0$ pour tout point intérieur x est une condition *suffisante* pour que f soit strictement croissante, mais elle n'est *pas nécessaire*. Par exemple, la fonction $f(x) = x^3$ est dérivable et strictement croissante sur $\mathbb{R}$, mais sa dérivée $f'(x) = 3x^2$ s'annule pour $x = 0$. En fait, il n'est pas difficile de donner une condition nécessaire et suffisante. Pour simplifier, on l'énonce seulement pour « strictement croissante » (le cas « strictement décroissante » est analogue).

Proposition 3.31. — *Soient I un intervalle de $\mathbb{R}$ et f une application $I \to \mathbb{R}$ continue sur I et dérivable en tout point intérieur de I. Les conditions suivantes sont équivalentes :*

(i) *f est strictement croissante sur I.*

(ii) *On a $f'(x) \geq 0$ pour tout point intérieur x de I et de plus, pour tout $a < b$ dans I, il existe $c \in]a, b[$ tel que $f'(c) > 0$.*

Démonstration. — Supposons (i) vérifié. Alors d'après le théorème précédent, on a $f'(x) \geq 0$ pour tout point intérieur x de I. De plus, pour tout $a < b$ dans I, le théorème des accroissements finis assure qu'il existe $c \in]a, b[$ tel que $f'(c) = (f(b) - f(a))/(b - a)$, et ceci est > 0 puisque f est strictement croissante.

Réciproquement, supposons (ii) vérifié et soient $a < b$ dans I. D'après le théorème précédent, on sait déjà que f est croissante, d'où $f(a) \leq f(b)$. Montrons que $f(a) < f(b)$ en raisonnant par l'absurde. Si on avait $f(a) = f(b)$ alors pour tout $x \in [a, b]$ on aurait $f(a) \leq f(x) \leq f(b) = f(a)$, donc f serait constante sur $[a, b]$ et l'on aurait $f'(x) = 0$ pour tout $x \in]a, b[$, contredisant l'hypothèse (ii). $\square$

3.5. Fonctions réciproques

***Définition 3.32* (Application réciproque).** — Soient I un intervalle non vide de $\mathbb{R}$ et $f : I \to \mathbb{R}$ une application continue et strictement monotone. D'après le théorème des valeurs intermédiaires, $f(I)$ est un intervalle J de $\mathbb{R}$. D'autre part, comme f est strictement monotone, elle est *injective* : pour tout $x \neq x'$ dans I on a $f(x) \neq f(x')$.

On peut donc définir une application $g : J \to I$ comme suit. Soit $y \in J = f(I)$, alors il existe au moins un $x \in I$ tel que $f(x) = y$, et comme f est injective cet x est en fait unique, et l'on pose $x = g(y)$. On a donc les formules :

$$\boxed{\forall x \in I, \quad g(f(x)) = x, \qquad\qquad \forall y \in J, \quad f(g(y)) = y}$$

et l'application $g : J \to I$ ainsi construite est appelée *l'application réciproque* de f et est notée f^{-1}.

Remarque 3.33. — Les notions précédentes sont des notions générales. Soient X, Y deux ensembles et $f : X \to Y$ une application.

(i) On dit que f est *injective* si pour tout $x \neq x'$ dans I on a $f(x) \neq f(x')$. Ceci équivaut à dire que tout $y \in Y$ admet *au plus* un antécédent (il en admet un s'il appartient au sous-ensemble $f(X)$ de Y, et zéro sinon).

(ii) On dit que f est *surjective* si tout $y \in Y$ admet *au moins* un antécédent, c.-à-d. si $f(X) = Y$.

(iii) On dit que f est *bijective* si elle est à la fois injective et surjective, c.-à-d. si pour tout $y \in Y$ il existe un unique $x \in X$ tel que $f(x) = y$. Dans ce cas, en posant $x = g(y)$ on obtient une application $g : Y \to X$ qui vérifie :

$$(**) \qquad f(g(y)) = y \quad \text{et} \quad g(f(x)) = x, \qquad \forall y \in Y,\ x \in X.$$

Ces conditions entraînent que g aussi est surjective (car tout x est l'image par g de $f(x)$) et injective (car si $g(y) = g(y')$ alors $y = f(g(y)) = f(g(y')) = y'$), donc bijective, et que l'application réciproque de g est f. On dit donc que f et g sont des *bijections réciproques* l'une de l'autre.

(iv) Si f est simplement supposée injective, elle définit une application $X \to f(X)$ qui est injective et surjective, donc bijective, et on peut définir l'application inverse $g : f(X) \to X$.

Remarque 3.34. — (1) Il existe des applications $f : \mathbb{R} \to \mathbb{R}$ strictement monotones mais non continues. C'est le cas par exemple pour f définie par $f(x) = x$ pour $x < 0$ et $f(x) = 1 + x$ pour $x \geq 0$.

(2) Soit I un intervalle non vide de $\mathbb{R}$. On peut montrer que toute application continue $f : I \to \mathbb{R}$ qui est injective est nécessairement strictement monotone, mais nous n'aurons pas besoin de cela. Il est plus intéressant de savoir que toute application continue et strictement monotone $f : I \to \mathbb{R}$ est une bijection de I sur l'intervalle $J = f(I)$ donc possède une application réciproque $f^{-1} : J \to I$.

Le théorème suivant permet alors de construire un certain nombre de fonctions continues (ou dérivables).

(Q) ***Théorème 3.35* (de l'application réciproque).** — *Soient I un intervalle non vide et non réduit à un point, $f : I \to \mathbb{R}$ une application continue et strictement monotone, J l'intervalle $f(I)$ et $g : J \to I$ l'application réciproque de f.*

(i) *g est strictement monotone, de même sens que f, et est continue sur J.*

(ii) *Soit $a \in I$. Si f est dérivable en a et si $f'(a) \neq 0$,* [7] *alors g est dérivable en $b = f(a)$ et* $\boxed{g'(b) = 1/f'(a).}$ *Donc, avec la notation $g = f^{-1}$ on a la formule :*

$$\boxed{\left(f^{-1}\right)'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}}$$

[7] Pour simplifier la question de cours, on a donné l'énoncé sous forme condensée, en écrivant ici « dérivable ». Si a est une extrémité de I, il faut modifier de façon évidente (dérivable à droite ou à gauche) l'hypothèse et la conclusion, cf. la démonstration.

Démonstration. — Faisons la démonstration lorsque f est strictement croissante. (Le cas strictement décroissant est analogue). Montrons d'abord que g est strictement croissante. Soient $y < y'$ dans J. On ne peut avoir $g(y) \geq g(y')$ car sinon, comme f est croissante, $y = f(g(y))$ serait $\geq f(g(y')) = y'$. Donc on a $g(y) < g(y')$, ce qui montre que g est strictement croissante.

Montrons maintenant que g est continue. Soient $b \in J$ et $a = g(b) \in I$. Si a est un point intérieur de I, il existe $\varepsilon > 0$ tel que $]a-\varepsilon, a+\varepsilon[$ soit contenu dans I et l'on a

$$(*) \qquad f(a-\varepsilon) < f(a) = b < f(a+\varepsilon).$$

Alors $V =]f(a-\varepsilon), f(a+\varepsilon)[$ est un voisinage de b contenu dans J, et $g(V) \subset]a-\varepsilon, a+\varepsilon[$ puisque g est croissante. Ceci montre que g est continue en b.

Si a est le plus petit élément de I alors, comme f est croissante, $b = f(a)$ est le plus petit élément de J ; d'autre part, il existe $\varepsilon > 0$ tel que $[a, a+\varepsilon[$ soit contenu dans I et l'on a $f(a) = b < f(a+\varepsilon)$. Posant $\delta = f(a+\varepsilon) - f(a)$, on a $\delta > 0$ et $[b, f(a+\varepsilon)[=]b-\delta, b+\delta[\cap J$ est envoyé par g dans $[a, a+\varepsilon[$, ce qui montre que g est continue à droite en b. Enfin, si a est le plus grand élément de I alors b est le plus grand élément de J et l'on montre de même que g est continue à gauche en b. Ceci prouve (i).

Prouvons (ii). Si a est un point intérieur de I, alors $b = f(a)$ est un point intérieur de J (ceci résulte de $(*)$ dans la preuve de (i)). La fonction « taux d'accroissement de g en b », notée τ_g, est donc définie sur un intervalle ouvert épointé $V^* =]b-\delta, b+\delta[\setminus \{b\}$ par :

$$\tau_g(y) = \frac{g(y) - a}{y - b}.$$

Soit τ_f la fonction « taux d'accroissement de f en a », i.e. $\tau_f(x) = \dfrac{f(x) - f(a)}{x - a}$. Comme g est continue en y, alors $\lim_{y \to b} g(y) = g(b) = a$, et d'après le théorème sur la composition de limites (2.25),

$$\tau_f\big(g(y)\big) = \frac{y - b}{g(y) - a}$$

tend vers $f'(a)$ quand y tend vers b. Comme $f'(a) \neq 0$, alors $\tau_g(y) = \dfrac{1}{\tau_f(g(y))}$ tend vers $\dfrac{1}{f'(a)}$ quand y tend vers b. Donc g est dérivable en b, de dérivée $g'(b) = 1/f'(a)$. De plus, si a est le plus petit (resp. plus grand) élément de I et si $f'_d(a) \neq 0$ (resp. $f'_g(a) \neq 0$), alors $b = f(a)$ est le plus petit (resp. plus grand) élément de I, et l'on montre comme plus haut que $g'_d(b) = 1/f'_d(a)$ (resp. $g'_g(b) = 1/f'_g(a)$). La démonstration est analogue lorsque f est strictement décroissante, mais on notera que dans ce cas, si a est le plus petit (resp. plus grand) élément de I et si $f'_d(a) \neq 0$ (resp. $f'_g(a) \neq 0$), alors $b = f(a)$ est le plus *grand* (resp. plus *petit*) élément de I et l'on a $g'_g(b) = 1/f'_d(a)$ (resp. $g'_d(b) = 1/f'_g(a)$). $\square$

Remarque 3.36. — Ayant vu, une fois pour toutes, que f^{-1} est dérivable en $b = f(a)$ si f est dérivable en a et $f'(a) \neq 0$, on peut retrouver facilement la formule donnant $(f^{-1})'(b)$ de la manière suivante. On écrit que $(f^{-1} \circ f)(x) = x$ pour tout $x \in I$ d'où, en utilisant la formule pour la dérivée d'une fonction composée :

$$(f^{-1})'(f(a)) \times f'(a) = 1 \qquad \text{d'où} \qquad (f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Ceci montre d'ailleurs que si $f'(a) = 0$ alors la fonction f^{-1} n'est *pas dérivable* en $f(a)$.

***Remarque 3.37* (Interprétation graphique)**. — Dans le plan $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}$, notons σ la symétrie par rapport à la 1ère diagonale (la droite d'équation $y = x$), i.e. $\sigma((x,y)) = (y,x)$ pour tout $(x,y) \in \mathbb{R}^2$. Si D est une droite d'équation $Y = \alpha(X-a)+\beta$, c.-à-d. si $D = \{(x,y) \mid y = \alpha(x-a)+\beta\}$, et si sa pente α est $\neq 0$ alors

$$\sigma(D) = \{(y,x) \mid x = \alpha^{-1}(y-\beta)+a\}$$

est une droite de pente α^{-1}. Par contre, si $\alpha = 0$ i.e. si D est la droite horizontale $Y = \beta$, alors $\sigma(D)$ est la droite verticale d'équation $X = \beta$.

Sous les hypothèses du théorème 3.35, le graphe de f est

$$\Gamma_f = \{(x,y) \mid x \in I,\ y = f(x) \in J\}$$

et son image par σ n'est autre que le graphe de f^{-1} :

$$\sigma(\Gamma_f) = \{(y,x) \mid y \in J,\ x = f^{-1}(y) \in I\} = \Gamma_{f^{-1}}.$$

On voit ainsi, de façon géométrique, que si Γ_f a pour tangente au point (a,b) la droite D, alors $\Gamma_{f^{-1}} = \sigma(\Gamma_f)$ a pour tangente au point (b,a) la droite $\sigma(D)$. On obtient ainsi que si la pente $f'(a)$ de D est $\neq 0$, alors $\sigma(D)$ a pour pente $1/f'(a)$ donc f^{-1} est dérivable en b de dérivée $(f^{-1})'(b) = 1/f'(a)$, tandis que si $f'(a) = 0$ alors la tangente $\sigma(D)$ à $\Gamma_{f^{-1}}$ au point (b,a) est verticale, donc f^{-1} n'est pas dérivable en b.

Remarque 3.38. — Pour tout $n \in \mathbb{N}^*$, l'application $f_n : x \mapsto x^n$ est dérivable sur $\mathbb{R}$, de dérivée $x \mapsto nx^{n-1}$, sa limite en $+\infty$ est $+\infty$ et l'on a $f(0) = 0$. C'est une fonction impaire (resp. paire) si n est impair (resp. pair).

(1) Si n est *impair*, f_n est strictement croissante et $\lim_{x\to-\infty} x^n = -\infty$. Donc $f_n(\mathbb{R}) = \mathbb{R}$ et f_n est une bijection strictement croissante de $\mathbb{R}$ sur $\mathbb{R}$. L'application réciproque $\mathbb{R} \to \mathbb{R}$ est notée $x \mapsto x^{1/n} = \sqrt[n]{x}$.

(2) Si n est *pair*, on a $\lim_{x\to-\infty} x^n = +\infty$, f_n est strictement décroissante sur $\mathbb{R}_-$ et strictement croissante sur $\mathbb{R}_+$. On a $f(\mathbb{R}) = \mathbb{R}_+$ et comme $f(-x) = f(x)$ pour tout x, f_n n'est pas injective mais l'on a $f_n(\mathbb{R}_+) = \mathbb{R}_+$ et f_n induit une bijection strictement croissante de $\mathbb{R}_+$ sur $\mathbb{R}_+$, qu'on notera p_n. L'application réciproque $\mathbb{R}_+ \to \mathbb{R}_+$ est aussi notée $x \mapsto x^{1/n} = \sqrt[n]{x}$.

On déduit alors du théorème 3.35 (et de 3.14 (iii)) le corollaire suivant.

Corollaire 3.39. — *Soit $n \in \mathbb{N}^*$. Notons provisoirement g_n la fonction $x \mapsto x^{1/n} = \sqrt[n]{x}$.*

(i) *Si n est* impair, *g_n est définie et continue sur $\mathbb{R}$ et dérivable sur $\mathbb{R}^*$ de dérivée*

$$x \mapsto \frac{1}{n}\, x^{\frac{1}{n}-1} = \frac{1}{nx^{\frac{n-1}{n}}}$$

Par contre, elle n'est pas dérivable en 0 si $n > 1$. De plus, la fonction $x \mapsto x^{-1/n} = 1/\sqrt[n]{x}$ est définie et dérivable sur $\mathbb{R}^*$, de dérivée

$$x \mapsto \frac{-1}{n}\, x^{\frac{-1}{n}-1}$$

(ii) *Si n est* pair, *g_n est définie et continue sur $\mathbb{R}_+$ et dérivable sur $\mathbb{R}_+^*$ de dérivée*

$$x \mapsto \frac{1}{n}\, x^{\frac{1}{n}-1} = \frac{1}{nx^{\frac{n-1}{n}}}$$

Par contre, elle n'est pas dérivable à droite en 0 si $n > 1$. De plus, la fonction $x \mapsto x^{-1/n} = 1/\sqrt[n]{x}$ est définie et dérivable sur $\mathbb{R}_+^*$, de dérivée

$$x \mapsto \frac{-1}{n}\, x^{\frac{-1}{n}-1}$$

Démonstration. — Notons D_n le domaine de définition de g_n ($\mathbb{R}$ si n est impair et $\mathbb{R}_+$ si n est pair). D'après le théorème 3.35, g_n est dérivable en tout $x \in D_n^* = D_n \setminus \{0\}$, de dérivée :

$$g_n'(x) = \frac{1}{f_n'(g_n(x))} = \frac{1}{n g_n(x)^{n-1}} = \frac{1}{n x^{\frac{n-1}{n}}} = \frac{1}{n}\, x^{\frac{1-n}{n}} = \frac{1}{n}\, x^{\frac{1}{n}-1}\,.$$

De plus, d'après 3.14, la fonction $x \mapsto x^{-1/n} = 1/g_n(x)$ est définie et dérivable sur D_n^*, de dérivée

$$\frac{-g_n'(x)}{g_n(x)^2} = \frac{-1}{n}\, x^{\frac{1}{n}-1} x^{\frac{-2}{n}} = \frac{-1}{n}\, x^{\frac{-1}{n}-1}\,.$$

□

Plus généralement, on a la proposition ci-dessous. Rappelons d'abord que la fonction exp (qu'on étudiera dans le prochain chapitre) est dérivable sur $\mathbb{R}$, de dérivée $\exp' = \exp$, et est une bijection strictement croissante de $\mathbb{R}$ sur $\mathbb{R}_+^*$. Son application réciproque, la fonction logarithme népérien $\ln : \mathbb{R}_+^* \to \mathbb{R}$ vérifie $\ln'(x) = 1/x$ pour tout $x \in \mathbb{R}_+^*$.

Définition et proposition 3.40. — *On pose* $\boxed{x^a = \exp(a\ln(x))}$ *pour tout* $a \in \mathbb{R}$ *et tout* $x \in \mathbb{R}_+^*$.

(i) *Pour a fixé, la fonction $x \mapsto x^a$ est dérivable sur $\mathbb{R}_+^*$, de dérivée $x \mapsto a x^{a-1}$.*

(ii) *Pour tout $a, b \in \mathbb{R}$ et $x \in \mathbb{R}_+^*$, on a $x^a x^b = x^{a+b}$.*

Démonstration. — (i) D'après le théorème sur la dérivabilité des fonctions composées, la fonction $x \mapsto x^a$ est dérivable sur $\mathbb{R}_+^*$, de dérivée $x \mapsto a\exp(a\ln(x))/x$ et comme $1/x = \exp(-\ln(x))$ et $\exp(y)\exp(z) = \exp(y+z)$, alors $a\exp(a\ln(x))/x$ égale

$$a\exp\big(a\ln(x) - \ln(x)\big) = a\exp\big((a-1)\ln(x)\big) = a x^{a-1}.$$

(ii) On a $x^a x^b = \exp(a\ln(x))\exp(b\ln(x)) = \exp\big((a+b)\ln(x)\big) = x^{a+b}$. □