Dérivation

Table des matières

$\mathbb{K}$ désigne $\mathbb{R}$ ou $\mathbb{C}$.
Dans tout ce chapitre, les fonctions étudiées sont définies sur un intervalle inclus dans $\mathbb{R}$ d'intérieur non vide (c'est-à-dire contenant au moins deux éléments distincts), que l'on notera le plus souvent I, et sont à valeurs dans un $\mathbb{K}$-espace vectoriel de dimension finie, que l'on notera usuellement E.

1 Dérivabilité

1.1 Interprétations d'une dérivée

Définition. f est dérivable au point a si et seulement si l'application

$$\begin{array}{ccc} I \setminus \{a\} & \longrightarrow & E \\ t & \longmapsto & \dfrac{f(t)-f(a)}{t-a} \end{array}$$

admet une limite lorsque t tend vers a en appartenant à $I \setminus \{a\}$. Dans ce cas, cette limite est appelée la dérivée de f au point a. On la note

$$\boxed{f'(a) = \left[\frac{d}{dt}(f(t))\right]_{t=a} = \lim_{\substack{t\to a \\ t\neq a, t\in I}} \frac{f(t)-f(a)}{t-a} \in \mathbb{K} \quad .}$$

Remarque. La notation f' est due à Lagrange, la notation $\dfrac{d}{dt}(f(t))$ est due à Leibniz. Newton utilise la notation $\dot{f}$, très utilisée en physique.

ATTENTION : ne confondez pas ces deux notations : $(xe^x)'$ n'a pas de sens, pas plus que $\frac{d}{dx}(exp)$. Il convient d'utiliser $\frac{d}{dx}(xe^x)$ et exp'.

Remarque. La notion de dérivée en a est une notion locale au point a.

Remarque. Cette définition est recevable même lorsque I n'est pas un intervalle. Il suffit que a soit un point d'accumulation de I. On se limite cependant tout de suite au cas où I est un intervalle car les théorèmes fondamentaux de la dérivation (cf paragraphes 4 et 5) ne sont pas valables lorsque I est une partie quelconque de $\mathbb{R}$.

Remarque. Dans le chapitre "Dérivation et intégration, une première approche", lorsque $E = \mathbb{R}$, on a vu que la corde du graphe de f entre les abscisses x_0 et x_1 a pour équation : $y - f(x_0) = \dfrac{f(x_1)-f(x_0)}{x_1-x_0} \times (x - x_0)$.

En particulier, la pente de cette droite est égale à $\dfrac{f(x_1)-f(x_0)}{x_1-x_0}$.
Informellement, lorsque f est dérivable en x_0, la corde du graphe de f entre les abscisses x_0 et x_1 tend vers une droite non verticale, de pente $f'(x_0)$, que l'on appelle la tangente au graphe de f en le point de coordonnées $(x_0, f(x_0))$. Cette tangente a donc pour équation : $y - f(x_0) = f'(x_0).(x - x_0)$.
Cela dit que la meilleure approximation de f, au voisinage de x_0, parmi l'ensemble des applications affines, est $x \longmapsto f(x_0) + f'(x_0).(x - x_0)$.

Remarque. La dérivée admet également une interprétation cinématique : Si l'on considère que $f(t)$ représente la position d'un mobile ponctuel à l'instant t, la quantité $\left\|\dfrac{f(t)-f(a)}{t-a}\right\|$ représente la distance parcourue par ce mobile entre les instants t et a, divisé par la durée $|t-a|$ du parcours. Il s'agit donc de la vitesse moyenne du mobile entre les instants a et t.
Ainsi $\|f'(a)\|$ représente la vitesse instantanée du mobile à l'instant a. Et la direction du vecteur $f'(a)$ indique dans quel sens se dirige le mobile lorsque le temps a croît.

1.2 Dérivées à gauche et à droite

Définition. Supposons que $a \neq \sup I$ et posons $J = I \cap [a, +\infty[$. On dit que f est dérivable à droite en a si et seulement si $f_{/J}$ est dérivable en a. Dans ce cas, la dérivée de la restriction est appelée dérivée à droite de f en a, et elle est notée

$$f'_d(a) = \lim_{\substack{t \to a \\ t>a, t\in I}} \frac{f(t)-f(a)}{t-a}.$$

On définirait de même la notion de dérivée à gauche lorsque $a \neq \inf I$.

Théorème. On suppose que a est à l'intérieur de I. Alors
$\boxed{\begin{array}{l} f \text{ est dérivable en } a \text{ si et seulement si} \\ f \text{ est dérivable à droite et à gauche en } a \text{ et si l'on a } f'_d(a) = f'_g(a). \end{array}}$
Dans ce cas, $f'(a) = f'_d(a) = f'_g(a)$.

Exemple. $x \longmapsto |x|$ n'est pas dérivable en 0.

1.3 Dérivées et développements limités

Propriété.
$\boxed{\begin{array}{l} f \text{ est dérivable en } a \text{ si et seulement si il existe } l \in E \text{ tel que} \\ f(t) = f(a) + (t-a)l + \underset{\substack{t \to a \\ t\neq a, t\in I}}{o}(t-a). \text{ Dans ce cas } l = f'(a). \end{array}}$

Démonstration.

$$\begin{aligned} f(t)=f(a)+(t-a)l+\circ(t-a) &\iff \frac{f(t)-f(a)}{t-a}-l=o(1)\\ &\iff \frac{f(t)-f(a)}{t-a}-l \underset{\substack{t\to a\\ t\neq a,t\in I}}{\longrightarrow} 0. \end{aligned}$$ □

Remarque. Lorsque $E=\mathbb{R}$, ceci permet de justifier que, lorsque f est dérivable en a, la tangente en a est la meilleure approximation du graphe de f par une droite non verticale au voisinage de a.
En effet, si D est une droite quelconque, mais non verticale, elle admet une équation de la forme $y=p(x-a)+b$. L'erreur commise en remplaçant le graphe de f par cette droite est l'application $e_D(x)=|f(x)-(p(x-a)+b)|$.
Or $f(x)-(p(x-a)+b)=(f(a)-b)+(x-a)(f'(a)-p)+o(x-a)$. Ainsi, lorsque $b\neq f(a)$, c'est-à-dire lorsque D ne passe pas par le point $(a,f(a))$, $e_D(x)\sim|f(a)-b|$, et lorsque $b=f(a)$, mais que D est différente de la tangente, $e_D(x)\sim|(x-a)(f'(a)-p)|$. C'est infiniment plus grand au voisinage de a que l'erreur commise avec la tangente T, car $e_T(x)=o(x-a)$.

Propriété. Si f est dérivable en a, elle est continue en a.

Démonstration.
Si f est dérivable en a,
$f(t)=f(a)+(t-a)l+\circ(t-a)=f(a)+o(1) \underset{\substack{t\to a\\ t\neq a,t\in I}}{\longrightarrow} f(a)$. □

Remarque. La réciproque est fausse. Par exemple, $x\longmapsto x\sin\frac{1}{x}$ se prolonge continûment sur $\mathbb{R}$, mais ce prolongement n'est pas dérivable en 0.

Remarque. Si f est seulement dérivable à droite et à gauche en a,
alors f est continue en a.

2 Opérations sur les fonctions dérivables

Propriété. Dérivation d'une application à valeurs dans un produit.
Supposons que $E=\prod_{i=1}^{p}E_i$, où pour tout $i\in\mathbb{N}_p$, E_i est un $\mathbb{K}$-espace vectoriel normé.
Pour tout $t\in I$, notons $f(t)=(f_1(t),\dots,f_p(t))$.
Alors f est dérivable en a si et seulement si, pour tout $i\in\mathbb{N}_p$, f_i est dérivable en a, et dans ce cas

$$f'(a)=(f_1'(a),\dots,f_p'(a)).$$

Démonstration.
Il suffit d'utiliser la propriété de limite d'une application à valeurs dans un produit pour l'application $\begin{array}{ccc} I\setminus\{a\} & \longrightarrow & E \\ t & \longmapsto & \dfrac{f(t)-f(a)}{t-a}=\left(\dfrac{f_1(t)-f_1(a)}{t-a},\dots,\dfrac{f_p(t)-f_p(a)}{t-a}\right) \end{array}$. □

Propriété. On suppose que E est de dimension finie $p \in \mathbb{N}$. Il existe alors une base $e = (e_1, \ldots, e_p)$ de E. Cela signifie que, pour tout $x \in E$, il existe un unique $(x_1, \ldots, x_p) \in \mathbb{K}^p$ tel que $x = \sum_{i=1}^{p} x_i e_i$.

Pour tout $t \in I$, notons $f(t) = \sum_{i=1}^{p} f_i(t) e_i$.

Alors f est dérivable en a si et seulement si, pour tout $i \in \{1, \ldots, p\}$, f_i est dérivable en a et dans ce cas

$$f'(a) = \sum_{i=1}^{p} f_i'(a) e_i.$$

Démonstration.

Notons φ l'application de $\mathbb{K}^p$ dans E définie par $\varphi(x_1, \ldots, x_p) = \sum_{i=1}^{p} x_i e_i$.

On vérifie aisément que φ est linéaire et le fait que e est une base signifie que φ est bijective Ainsi, φ est un isomorphisme.

E étant de dimension finie, nous verrons que toutes les normes de E sont équivalentes.

On peut donc choisir d'utiliser la norrme définie par : $\forall x = \sum_{i=1}^{p} x_i e_i \in E$, $\|x\| = \sum_{i=1}^{p} |x_i|$ (on vérifie aisément que c'est bien une norme).

Avec ce choix, pour tout $(x_1, \ldots, x_p) \in \mathbb{K}^p$, $\|\varphi(x_1, \ldots, x_p)\| = \|(x_1, \ldots, x_p)\|_1$, donc φ est une isométrie de $(\mathbb{K}^p, \|.\|_1)$ dans $(E, \|.\|)$.

On en déduit facilement que φ et φ^{-1} sont 1-lipschitziennes, donc φ est un homéomorphisme.

Fort de ces notations, f est dérivable en a si et seulement si il existe $\ell = \sum_{i=1}^{p} \ell_i e_i \in E$ tel que (1) : $\dfrac{f(t) - f(a)}{t - a} \underset{t \to a}{\longrightarrow} \ell$. Alors d'après le théorème de composition des limites,

$$\begin{aligned}
(1) \quad & \iff \varphi^{-1}\Big(\frac{f(t) - f(a)}{t - a}\Big) \underset{t \to a}{\longrightarrow} \varphi^{-1}(\ell) \\
& \iff \Big(\frac{f_1(t) - f_1(a)}{t - a}, \ldots, \frac{f_p(t) - f_p(a)}{t - a}\Big) \underset{t \to a}{\longrightarrow} (\ell_1, \ldots, \ell_p).
\end{aligned}$$

À nouveau, la propriété de limite d'une application à valeurs dans un produit permet de conclure. □

Cas particulier. Si f est une application de I dans $\mathbb{C}$, f est dérivable en a si et seulement si $Im(f)$ et $Re(f)$ sont des applications dérivables en a. Dans ce cas

$$f'(a) = Re(f)'(a) + iIm(f)'(a).$$

Exemple. L'application $\begin{array}{ccc} I & \longrightarrow & \mathcal{M}_n(\mathbb{C}) \\ t & \longmapsto & M(t) = (m_{i,j}(t))_{\substack{1 \le i \le n \\ 1 \le j \le n}} \end{array}$ est dérivable en a si et seulement si pour tout $(i, j) \in \{1, \ldots, n\}^2$ l'application $m_{i,j}$ est dérivable en a, et dans ce cas, $M'(a) = (m'_{i,j}(a))_{\substack{1 \le i \le n \\ 1 \le j \le n}}$.

Propriété. Soient F un second $\mathbb{K}$-espace vectoriel normé et u une application linéaire continue de E dans F.
Si f est dérivable en a, alors $u \circ f$ est dérivable en a et

$$\boxed{(u \circ f)'(a) = u(f'(a)).}$$

Démonstration.
Supposons que f est dérivable en a. Alors
$\dfrac{u(f(t)) - u(f(a))}{t-a} = u\left(\dfrac{f(t)-f(a)}{t-a}\right) \underset{\substack{t \to a \\ t \neq a, t \in I}}{\longrightarrow} u(f'(a)),$
car u est est continue en $f'(a)$. $\square$

Exemple. Si $f\ : I \longrightarrow \mathbb{C}$ est dérivable, $\overline{f}$ est dérivable et $\overline{f}' = \overline{f'}$.

Propriété de linéarité.
Soit $(\alpha, \beta) \in \mathbb{K}^2$. Si f et g sont dérivables en a, alors $\alpha f + \beta g$ est dérivable en a et

$$(\alpha f + \beta g)'(a) = \alpha f'(a) + \beta g'(a).$$

Définition.
Soient E, F et G trois $\mathbb{K}$-espaces vectoriels et $B : E \times F \longrightarrow G$ une application.
On dit que B est bilinéaire si et seulement si, pour tout $y \in F$, $x \longmapsto B(x, y)$ est linéaire et si, pour tout $x \in E$, $y \longmapsto B(x, y)$ est linéaire.

Remarque. Informellement, la notion d'application bilinéaire est la bonne manière de généraliser le produit entre réels ou complexes : toute application bilinéaire définit en quelque sorte une multiplication entre vecteurs.

Théorème de dérivation d'un produit. Soient E, F et G trois $\mathbb{K}$-espaces vectoriels et $B : E \times F \longrightarrow G$ une application bilinéaire continue. Soient f une application de I dans E et g une application de I dans F. On dispose de l'application

$$B(f,g) : \begin{array}{ccc} I & \longrightarrow & G \\ t & \longmapsto & B(f(t), g(t)) \end{array}.$$

Si f et g sont dérivables en a, alors $B(f, g)$ est dérivable en a et

$$\boxed{B(f,g)'(a) = B(f'(a), g(a)) + B(f(a), g'(a)).}$$

Démonstration.
$$\begin{aligned} \frac{1}{t-a}(B(f(t), g(t)) - B(f(a), g(a))) = & \ \frac{1}{t-a}(B(f(t), g(t)) - B(f(a), g(t)) + \\ & B(f(a), g(t)) - B(f(a), g(a))) \\ = B(\frac{1}{t-a}(f(t)-f(a)), g(t)) \quad & +B(f(a), \frac{1}{t-a}(g(t)-g(a))). \end{aligned}$$
On conclut en passant à la limite et en utilisant la continuité de l'application bilinéaire B. $\square$

Exemple. Si A et B sont deux applications de I dans $\mathcal{M}_n(\mathbb{R})$, dérivables en a, AB est dérivable en a et $(AB)'(a) = A'(a)B(a) + A(a)B'(a)$.
En effet, on verra plus tard que toute application bilinéaire définie sur $E \times F$ où E et F sont de dimensions finies est continue.

Corollaire. Si $f_1, \ldots, f_p$ sont p applications de I dans $\mathbb{C}$ dérivables en a, le produit $\prod_{i=1}^{p} f_i$ est dérivable en a et

$$\left(\prod_{i=1}^{p} f_i\right)'(a) = \sum_{i=1}^{p} \Big[f_i'(a) \prod_{\substack{1 \le j \le p \\ j \ne i}} f_j(a) \Big].$$

Dérivation des fonctions composées.
Soient J un intervalle de $\mathbb{R}$ d'intérieur non vide.
Soient $\varphi : I \longrightarrow J$ et $f : J \longrightarrow E$ deux applications.
Si φ est dérivable en a et f en $\varphi(a)$, alors $f \circ \varphi$ est dérivable en a et

$$\boxed{(f \circ \varphi)'(a) = \varphi'(a) f'(\varphi(a)).}$$

Démonstration.
Lorsque h tend vers 0, $f \circ \varphi(a+h) = f(\varphi(a) + h\varphi'(a) + o(h)) = f(\varphi(a) + k)$, en posant $k = h\varphi'(a) + o(h)$.
$k \underset{h \to 0}{\longrightarrow} 0$, donc par composition de développements limités,

$$\begin{aligned} f \circ \varphi(a+h) &= f(\varphi(a) + k) = f(\varphi(a)) + kf'(\varphi(a)) + o(k) \\ &= f(\varphi(a)) + h\varphi'(a)f'(\varphi(a)) + o(h)f'(\varphi(a)) + o(h\varphi'(a) + o(h)). \end{aligned}$$

Or $o(h)f'(\varphi(a)) = o(h)$ et $o(h\varphi'(a) + o(h)) = o(O(h) + o(h)) = o(O(h)) = o(h)$, donc $f \circ \varphi(a+h) = f(\varphi(a)) + h\varphi'(a)f'(\varphi(a)) + o(h)$, ce qu'il fallait démontrer. □

Remarque. Avec la notation de Leibniz, cette formule devient très naturelle : $\dfrac{d(f \circ \varphi)}{dx} = \dfrac{d(f \circ \varphi)}{d\varphi} \times \dfrac{d\varphi}{dx}$.

Corollaire. Avec les mêmes notations, si φ et f sont dérivables, alors $f \circ \varphi$ est dérivable et on a l'égalité fonctionnnelle suivante.

$$(f \circ \varphi)' = \varphi'(f' \circ \varphi).$$

Corollaire. Soit $n \in \mathbb{N}^*$. Pour tout $i \in \{1, \ldots, n\}$, on suppose que f_i est une application dérivable de I_i dans I_{i+1}, où $I_1, \ldots, I_{n+1}$ sont $n+1$ intervalles d'intérieurs non vides. Alors $f = f_n \circ \cdots \circ f_1$ est dérivable et

$$f' = (f_n' \circ f_{n-1} \circ \cdots \circ f_1) \times (f_{n-1}' \circ f_{n-2} \circ \cdots \circ f_1) \times \cdots \times f_1'.$$

Exemple. Soient $u : I \longrightarrow \mathbb{R}_+^*$ et $v : I \longrightarrow \mathbb{R}$ deux applications dérivables. Alors u^v est dérivable et

$$(u^v)' = (e^{v \ln u})' = \left(v\frac{u'}{u} + v' lnu \right) u^v.$$

Dérivée de l'inverse. Soit $f : I \longrightarrow \mathbb{K}^*$ une application dérivable en a. Alors $\frac{1}{f}$ est dérivable en a et

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

Démonstration.
Soit h non nul tel que $a + h \in I$.
$\frac{1}{h}\left(\frac{1}{f}(a+h) - \frac{1}{f}(a)\right) = \frac{1}{h}\left(\frac{1}{f(a+h)} - \frac{1}{f(a)}\right) = -\frac{f(a+h) - f(a)}{hf(a+h)f(a)} \underset{h\to 0}{\longrightarrow} -\frac{f'(a)}{f(a)^2}$. □

Dérivée logarithmique. Soit $f : I \longrightarrow \mathbb{K}^*$ une application dérivable en a.
La quantité $\frac{f'(a)}{f(a)}$ est appelée la dérivée logarithmique de f en a.
Lorsque $\mathbb{K} = \mathbb{R}$, elle est égale à $(\ln |f|)'(a)$.

Propriété. Soient u et v deux applications dérivables de I dans $\mathbb{K}^*$. On dispose des formules suivantes.

$$\frac{(uv)'}{uv} = \frac{u'}{u} + \frac{v'}{v}, \quad \frac{\left(\frac{u}{v}\right)'}{\left(\frac{u}{v}\right)} = \frac{u'}{u} - \frac{v'}{v}, \quad \forall n \in \mathbb{N}^* \quad \frac{(u^n)'}{u^n} = n\frac{u'}{u},$$

et, si u est à valeurs dans $\mathbb{R}_+^*$,

$$\forall \alpha \in \mathbb{R} \quad \frac{(u^\alpha)'}{u^\alpha} = \alpha\frac{u'}{u}.$$

3 Dérivées d'ordre supérieur

3.1 Définition

Définition récurrente de la dérivée n-ième.
Pour $n = 0$, on convient que $\boxed{f^{(0)} = f}$.
Pour $n \geq 1$, on dit que f est n fois dérivable sur I si et seulement si $f^{(n-1)}$ est définie et dérivable sur I. Dans ce cas, pour tout $t \in I$, on note $f^{(n)}(t) = (f^{(n-1)})'(t) = \frac{d^n}{dt^n}(f(t))$.
On dit que f est de classe D^n.

Notation. $f^{(2)}$ et $f^{(3)}$ sont souvent notées f'' et f'''.

Remarque. L'interprétation cinématique de $f''(a)$ est l'accélération instantanée du mobile ponctuel $f(t)$ à l'instant a.

Propriété. Pour tout $p, q \in \mathbb{N}$, f est $p + q$ fois dérivable sur I si et seulement si $f^{(p)}$ est q fois dérivable sur I, auquel cas, $f^{(p+q)} = [f^{(p)}]^{(q)}$.

Démonstration.
Par récurrence sur $n = p + q$. □

Remarque. On dit que f est n fois dérivable en a si et seulement si il existe une boule ouverte B centrée en a telle que $f|_{B\cap I}$ soit $n-1$ fois dérivable et telle que $[f|_{B\cap I}]^{(n-1)}$ soit dérivable en a.

Exemple.

$$\forall t \in \mathbb{R} \quad \frac{d^n}{dt^n}(sint) = sin(t + n\frac{\pi}{2}) \quad \text{et} \quad \frac{d^n}{dt^n}(cost) = cos(t + n\frac{\pi}{2}),$$

$$\forall a \in \mathbb{C} \quad \forall t \in \mathbb{R} \setminus \{a\} \quad \boxed{\frac{d^n}{dt^n}\left(\frac{1}{t-a}\right) = \frac{(-1)^n n!}{(t-a)^{n+1}}}.$$

Définition. On dit que f est de classe C^n si et seulement si $f^{(n)}$ est une application définie sur I et continue.
On dit que f est de classe C^∞ si et seulement si f est de classe C^n pour tout $n \in \mathbb{N}$.

3.2 Opérations sur les dérivées supérieures

Propriété de linéarité. Soit $n \in \mathbb{N}$.
Si f et g sont de classe D^n, alors pour tout $(\alpha, \beta) \in \mathbb{K}^2$, $\alpha f + \beta g$ est de classe D^n et $[\alpha f + \beta g]^{(n)} = \alpha f^{(n)} + \beta g^{(n)}$.

Formule de Leibniz. Soient E, F et G trois $\mathbb{K}$-espaces vectoriels normés et $B : E \times F \longrightarrow G$ une application bilinéaire continue.
Soient f une application de I dans E et g une application de I dans F. On dispose de l'application

$$B(f,g) : \begin{array}{ccc} I & \longrightarrow & G \\ t & \longmapsto & B(f(t), g(t)) \end{array}.$$

Soit $a \in I$; Si f et g sont dérivables n fois en a, $B(f,g)$ est dérivable n fois en a et

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

Démonstration.
Soit $n \in \mathbb{N}$. Notons $R(n)$ l'assertion suivante : pour toutes applications f de I dans E et g de I dans F, et pour tout $a \in I$, si f et g sont n fois dérivables en a, $B(f,g)$ est n fois dérivable en a et $B(f,g)^{(n)}(a) = \sum_{k=0}^{n} \binom{n}{k} B(f^{(k)}(a), g^{(n-k)}(a))$.

◇ Pour $n = 0$, soient f et g des applications de I dans E et de I dans F respectivement. Elles sont 0 fois dérivables et

$\sum_{k=0}^{n} \binom{n}{k} B(f^{(k)}(a), g^{(n-k)}(a)) = B(f(a), g(a)) = B(f,g)^{(0)}(a)$. Ceci prouve $R(0)$.

$\diamond$ Pour $n \geq 0$, supposons $R(n)$. Soient f et g des applications de I dans E et de I dans F respectivement que l'on suppose $n+1$ fois dérivables en $a \in I$. D'après $R(n)$, sur un intervalle voisinage de a, $B(f,g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} B(f^{(k)}, g^{(n-k)})$.

Pour tout $k \in \{0, \ldots, n\}$, $f^{(k)}$ et $g^{(n-k)}$ sont dérivables en a et B est bilinéaire, donc $B(f^{(k)}, g^{(n-k)})$ est dérivable en a. On peut donc dériver en a l'égalité précédente. On obtient :

$$\begin{aligned}
B(f,g)^{(n+1)}(a) &= \sum_{k=0}^{n} \binom{n}{k} B(f^{(k+1)}(a), g^{(n-k)}(a)) + B(f^{(k)}(a), g^{(n-k+1)}(a)) \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} B(f^{(k)}(a), g^{(n-k+1)}(a)) \\
&\quad + \sum_{k=0}^{n} \binom{n}{k} B(f^{(k)}(a), g^{(n-k+1)}(a)) \\
&= B(f^{(n+1)}(a), g^{(0)}(a)) + B(f^{(0)}(a), g^{(n+1)}(a)) \\
&\quad + \sum_{k=1}^{n} (\binom{n}{k-1} + \binom{n}{k}) B(f^{(k)}(a), g^{(n-k+1)}(a)) \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} B(f^{(k)}(a), g^{(n-k+1)}(a)).
\end{aligned}$$

Ceci prouve $R(n+1)$.

Ainsi, d'après le principe de récurrence, pour tout $n \in \mathbb{N}$, on a montré $R(n)$. $\square$

Corollaire. Pour tout $n \in \mathbb{N} \cup \{\infty\}$, le produit de deux applications D^n (resp : C^n) est D^n (resp : C^n).

Théorème de composition. Soient J un intervalle de $\mathbb{R}$ d'intérieur non vide.
Soient $\varphi : I \longrightarrow J$ et $f : J \longrightarrow E$ deux applications. Soit $n \in \mathbb{N} \cup \{\infty\}$.
Si φ et f sont D^n (resp : C^n) alors $f \circ \varphi$ est D^n (resp : C^n).

Démonstration.

Soit $n \in \mathbb{N}$. Notons $R(n)$ l'assertion suivante : pour toute application f de J dans E, si φ et f sont C^n, $f \circ \varphi$ est C^n.

$\diamond$ Pour $n = 0$, on sait qu'une composée d'applications continues est continue.

$\diamond$ Pour $n \geq 0$, supposons $R(n)$. Soit f une application de J dans E. Supposons que φ et f sont C^{n+1}.

En particulier, φ et f sont dérivables, donc $f \circ \varphi$ est dérivable et $(f \circ \varphi)' = \varphi' \times (f' \circ \varphi)$. f' et φ sont C^n, donc d'après $R(n)$, $f' \circ \varphi$ est C^n. De plus, φ' est C^n et on sait qu'un produit d'applications de classe C^n est de classe C^n. Ainsi, l'égalité précédente montre que $(f \circ \varphi)'$ est C^n, donc que $f \circ \varphi$ est C^{n+1}. Ceci prouve $R(n+1)$.

Ainsi, d'après le principe de récurrence, pour tout $n \in \mathbb{N}$, on a montré $R(n)$. $\square$

Remarque. Pour calculer la dérivée n-ième d'une application, on peut s'inspirer des méthodes permettant d'intégrer cette application.

Exercice. Calculez la dérivée n-ième de $\cos^5(t)$.

Résolution. Soit $t \in \mathbb{R}$.
$\cos^5 t = \dfrac{1}{2^5}(e^{it} + e^{-it})^5 = \dfrac{1}{2^5}(e^{5it} + 5e^{3it} + 10e^{it} + 10e^{-it} + 5e^{-3it} + e^{-5it})$,
donc $\cos^5 t = \dfrac{1}{2^4}(\cos(5t) + 5\cos(3t) + 10\cos(t))$ et
$\dfrac{d^n(\cos^5 t)}{dt^n} = \dfrac{1}{2^4}(5^n \cos(5t + n\dfrac{\pi}{2}) + 5 \times 3^n \cos(3t + n\dfrac{\pi}{2}) + 10\cos(t + n\dfrac{\pi}{2}))$.

4 L'égalité des accroissements finis

Dans ce paragraphe, sauf précision du contraire, toutes les applications utilisées sont définies sur I et **sont à valeurs dans $\mathbb{R}$.**

4.1 Extremum et point critique

Définition. On dit que f admet un maximum local en a si et seulement s'il existe un voisinage V de a tel que
$$\forall t \in V \cap I \quad f(t) \leq f(a).$$
On dit que f présente en a un maximum local strict si et seulement s'il existe un voisinage V de a tel que
$$\forall t \in V \cap I \setminus \{a\} \quad f(t) < f(a).$$
On définit de même les notions de minimum local, strict ou non.

Définition. Supposons que f est dérivable en $a \in \overset{\circ}{I}$.

On dit que a est un point critique de f si et seulement si $f'(a) = 0$.

Théorème.
Si f est dérivable en $a \in \overset{\circ}{I}$ et si f admet un extremum local en a, alors $f'(a) = 0$.

Ainsi, les extremums locaux de f sur $\overset{\circ}{I}$ sont des points critiques de f.

Démonstration.
Supposons par exemple que a est un maximum local de f.
$a \in I \setminus \{\sup(I), \inf(I)\}$, donc il existe $\varepsilon > 0$ tel que $]a - \varepsilon, a + \varepsilon[\subset I$.
Pour tout $x \in]a, a + \varepsilon[$, $x - a > 0$ et $f(x) - f(a) \leq 0$, donc $\dfrac{f(x) - f(a)}{x - a} \leq 0$, mais $f'(a) = \lim\limits_{\substack{x \to a \\ x \in]a, a+\varepsilon[}} \dfrac{f(x) - f(a)}{x - a}$, donc $f'(a) \leq 0$.

De même, pour tout $x \in]a-\varepsilon, a[$, $\dfrac{f(x)-f(a)}{x-a} \geq 0$ et on en déduit que $f'(a) \geq 0$. Ainsi, $f'(a) = 0$. □

Remarque. La réciproque est fausse, ainsi qu'on le voit en considérant l'application $x \longmapsto x^3$: 0 est un point critique, mais ce n'est pas un extremum de cette fonction.

4.2 Le lemme de Rolle

Lemme de Rolle. Soient $(a,b) \in \mathbb{R}^2$ avec $a < b$ et $f : [a,b] \longrightarrow \mathbb{R}$ une application continue sur $[a,b]$ et dérivable sur l'ouvert $]a,b[$.
Si $f(a) = f(b)$, il existe $c \in]a,b[$ tel que $f'(c) = 0$.

Démonstration.
f est continue sur le compact $[a,b]$, donc elle est bornée et elle atteint ses bornes. Ainsi, il existe $(m,M) \in [a,b]^2$ tel que, pour tout $x \in [a,b]$, $f(m) \leq f(x) \leq f(M)$.
⋄ Si $\{m,M\} \subset \{a,b\}$, comme $f(a) = f(b)$, $f(m) = f(M)$ et f est une application constante. Ainsi, pour tout $x \in]a,b[$, $f'(x) = 0$, donc il existe $c \in]a,b[$ tel que $f'(c) = 0$.
⋄ Sinon, supposons par exemple que $m \notin \{a,b\}$. Ainsi, $m \in]a,b[$ et m est un minimum global, donc local, de f. D'après la propriété précédente, $f'(m) = 0$. □

Remarque. Géométriquement, le lemme de Rolle affirme que le graphe de f présente au moins un point à tangente horizontale.

Remarque. Cinématiquement, le lemme de Rolle affirme qu'un mobile ponctuel se déplaçant sur la droite des réels qui revient sur sa position initiale admet au moins une fois une vitesse instantanée nulle.
C'est faux pour une application à valeur dans $\mathbb{C}$. Par exemple, avec $\begin{array}{rccc} f : & [0,2\pi] & \longrightarrow & \mathbb{C} \\ & \theta & \longmapsto & e^{i\theta} \end{array}$,
f est dérivable, $f(0) = f(2\pi)$, mais pour tout $\theta \in [0,2\pi]$, $|f'(\theta)| = |ie^{i\theta}| = 1 \neq 0$.

Un exemple fondamental : Soit P un polynôme à coefficients réels.
Si a et b sont deux racines réelles de P, avec $a < b$, alors son polynôme dérivé P' possède une racine $c \in]a,b[$.
Supposons que P est de degré $n \in \mathbb{N}^*$ et qu'il possède exactement n racines réelles distinctes $x_1 < x_2 < \cdots < x_n$ (on dit que P est simplement scindé dans $\mathbb{R}$). Alors P' est également simplement scindé sur $\mathbb{R}$, car ce qui précède permet de montrer qu'il possède $n-1$ racines distinctes, or P' est de degré $n-1$.
C'est faux dans $\mathbb{C}$. Par exemple $P(x) = x^n - 1 = \displaystyle\prod_{k=0}^{n-1}(x - e^{2i\frac{k\pi}{n}})$ est scindé dans $\mathbb{C}$, mais $P'(x) = nx^{n-1}$ n'est pas scindé.

Théorème de Rolle généralisé ***(Hors programme).***
Soit $(a,b) \in \mathbb{R} \cup \{-\infty, +\infty\}$ avec $a < b$. Si f est dérivable sur $]a,b[$ et si $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to b} f(x) \in \mathbb{R} \cup \{-\infty, +\infty\}$, alors il existe $c \in]a,b[$ tel que $f'(c) = 0$.

Démonstration.
On sait qu'une application continue $f :]a, b[\longrightarrow \mathbb{R}$ est injective si et seulement si elle est strictement monotone. Mais ici, f n'est pas strictement monotone : par exemple, si f était strictement croissante, on aurait en prenant $(a', b') \in]a, b[^2$ avec $a' < b'$, $\lim_{x \to a} f(x) \leq f(a') < f(b') \leq \lim_{x \to b} f(x)$, ce qui est faux. Donc f n'est pas injective. Ainsi, il existe $(a', b') \in]a, b[^2$ avec $a' < b'$ tel que $f(a') = f(b')$. D'après le lemme de Rolle, il existe $c \in]a', b'[$ tel que $f'(c) = 0$. $\Box$

4.3 L'égalité des accroissements finis

Théorème des accroissements finis (TAF). Soit $(a, b) \in \mathbb{R}^2$ avec $a \neq b$. On notera $I = [min(a, b), max(a, b)]$. Soit $f : I \longrightarrow \mathbb{R}$ une application continue sur I et dérivable sur $]min(a, b), max(a, b)[$.
Alors $\boxed{\text{il existe } c \text{ dans }]min(a, b), max(a, b)[\text{ tel que } f(b) - f(a) = (b - a)f'(c).}$

Démonstration.
Posons, pour tout $x \in [a, b]$, $g(x) = f(x) - f(a) - \dfrac{f(b) - f(a)}{b - a}(x - a)$. g est continue sur $[a, b]$ et dérivable sur $]a, b[$ et $g(a) = 0 = g(b)$, donc d'après le lemme de Rolle, il existe $c \in]a, b[$ tel que $0 = g'(c) = f'(c) - \dfrac{f(b) - f(a)}{b - a}$. $\Box$

Interprétation géométrique : au tableau.

Interprétation cinématique : Un mobile ponctuel se déplaçant sur une droite possède au moins une fois une vitesse instantanée égale à sa vitesse moyenne.

> **Exercice.** Montrer que, pour tout $x > 0$, $\dfrac{1}{1 + x} < \ln(1 + x) - \ln x < \dfrac{1}{x}$.
> **Solution :** Soit $x > 0$. D'après le TAF appliqué à ln entre x et $x + 1$, il existe $c_x \in]x, x + 1[$ tel que $\ln(1 + x) - \ln x = (\ln)'(c_x) = \frac{1}{c_x}$, ce qui permet de conclure.
> On pouvait aussi utiliser que $\ln(1 + x) - \ln x = \displaystyle\int_x^{x+1} \frac{dt}{t}$.

Exemple. Pour tout $x, y \in \mathbb{R}$, le TAF montre que $|\sin x - \sin y| \leq |x - y|$, donc en particulier avec $y = 0$, $|\sin x| \leq |x|$.

4.4 Théorème de la limite de la dérivée

TLD : On suppose que f est continue sur I, dérivable (resp : de classe C^1) sur $I \setminus \{a\}$ et qu'il existe $l \in \mathbb{R}$ tel que $f'(x) \underset{\substack{x \to a \\ x \in I \setminus \{a\}}}{\longrightarrow} l$.
Alors f est dérivable (resp : de classe C^1) sur I, avec $f'(a) = l$.

Démonstration.
L'égalité des accroissements finis garantit, pour tout $x \in I\setminus\{a\}$, l'existence de $c_x \in]a, x[$ tel que $\dfrac{f(x)-f(a)}{x-a} = f'(c_x)$, mais $c_x \underset{\substack{x\to a\\ x\in I\setminus\{a\}}}{\longrightarrow} a$, donc par composition des limites, $\dfrac{f(x)-f(a)}{x-a} = f'(c_x) \underset{\substack{x\to a\\ x\in I\setminus\{a\}}}{\longrightarrow} l$. □

Remarque. Supposons que f est continue sur I, dérivable sur $I \setminus \{a\}$ et que $f'(x) \underset{\substack{x\to a\\ x\in I\setminus\{a\}}}{\longrightarrow} +\infty$. Alors, en adaptant la démonstration, on obtient que $\dfrac{f(x)-f(a)}{x-a} \underset{\substack{x\to a\\ x\in I\setminus\{a\}}}{\longrightarrow} +\infty$, ce qui prouve que f n'est pas dérivable en a et que son graphe présente en a une tangente verticale.

Remarque. En utilisant les parties réelle et imaginaire, ce théorème est encore valable pour une fonction à valeurs dans $\mathbb{C}$.

Exemple. Pour tout $x > 0$, posons $f(x) = x^2 \ln x$.
D'après les théorèmes usuels, f est de classe C^∞ sur $\mathbb{R}_+^*$.
D'après les c.c, $f(x) \underset{x\to 0^+}{\longrightarrow} 0$, donc en posant $f(0) = 0$, f est continue sur $\mathbb{R}_+$.
Pour tout $x > 0$, $f'(x) = 2x\ln x + x \underset{x\to 0^+}{\longrightarrow} 0$, donc d'après le TLD, f est de classe C^1 sur $\mathbb{R}_+$ et $f'(0) = 0$.
Pour $x > 0$, $f''(x) = 2\ln x + 3 \underset{x\to 0^+}{\longrightarrow} -\infty$, donc d'après la remarque précédente, f' n'est pas dérivable en 0. On l'obtient ici directement en évaluant le taux de variation en 0 : $\dfrac{f'(x)-f'(0)}{x-0} = 2\ln x + 1 \underset{x\to 0^+}{\longrightarrow} -\infty$.

TLD : Généralisation aux dérivées d'ordre supérieur. Soient $k \in \mathbb{N} \cup \{\infty\}$.
On suppose que f est continue sur I, à valeurs dans $\mathbb{K}$, que f est de classe C^k sur $I \setminus \{a\}$ et que, pour tout $h \in [1, k] \cap \mathbb{N}$, il existe $l_h \in \mathbb{R}$ tel que $f^{(h)}(x) \underset{\substack{x\to a\\ x\in I\setminus\{a\}}}{\longrightarrow} l_h$.

Alors f est de classe C^k sur I.

Démonstration.
Montrons par récurrence sur h que pour tout $h \in [1, k] \cap \mathbb{N}$, f est de classe C^h sur I.
Pour $h = 1$, cela résulte du théorème précédent.
Pour $h \in [1, k-1]\cap\mathbb{N}$, supposons que f est de classe C^h. Alors $f^{(h)}$ vérifie les hypothèses du théorème précédent, donc $f^{(h)}$ est de classe C^1 sur I, ce qui prouve que f est de classe C^{h+1} sur I. □

Exemple. Considérons l'application $f : \mathbb{R} \longrightarrow \mathbb{R}$ définie par les relations suivantes : $f(0) = 0$ et si $t \neq 0$, $f(t) = e^{-\frac{1}{t^2}}$.
D'après les théorèmes usuels, f est une application de classe C^∞ sur $\mathbb{R}^*$.
On démontre par récurrence sur n que, pour tout $n \in \mathbb{N}$, il existe un polynôme P_n à coefficients réels tel que, pour tout $x \in \mathbb{R}^*$, $f^{(n)}(x) = \dfrac{P_n(x)}{x^{3n}} e^{-\frac{1}{x^2}}$.
Soit $n \in \mathbb{N}$. Il existe donc $a \in \mathbb{R}^*$ et $k \in \mathbb{Z}$ tel que

$|f^{(n)}(x)| \sim a|x|^k e^{-\frac{1}{x^2}} = ae^{-\frac{1}{x^2}+k\ln|x|} \underset{x\to 0^+}{\longrightarrow} 0$ d'après les c.c, donc d'après le TLD, f est une application de classe C^∞ sur $\mathbb{R}$ et pour tout $n \in \mathbb{N}$, $f^{(n)}(0) = 0$.

5 Formules de Taylor

Notation. $\mathbb{K}$ désigne toujours $\mathbb{R}$ ou $\mathbb{C}$.
On fixe $(a, b) \in \mathbb{R}^2$ avec $a \neq b$ et on note $[a, b]$ (respectivement : $]a, b[$) l'intervalle $[\min(a, b), \max(a, b)]$ (respectivement : $]\min(a, b), \max(a, b)[$) , même lorsque $a > b$.

5.1 L'égalité de Taylor-Lagrange (hors programme)

Théorème. Soient $n \in \mathbb{N}$ et $f : [a, b] \longrightarrow \mathbb{R}$.
On suppose que f est de classe C^n sur $[a, b]$ et $n + 1$ fois dérivable sur $]a, b[$.
Alors il existe $c \in]a, b[$ tel que $f(b) = f(a) + \displaystyle\sum_{k=1}^{n} \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)!} f^{(n+1)}(c)$.

Démonstration.
Pour tout $t \in [a, b]$, posons $\varphi(t) = f(b) - f(t) - \displaystyle\sum_{k=1}^{n} \frac{(b-t)^k}{k!} f^{(k)}(t) - \frac{(b-t)^{n+1}}{(n+1)!}\alpha$, où α est choisi tel que $\varphi(a) = 0$ (c'est possible car $\dfrac{(b-a)^{n+1}}{(n+1)!} \neq 0$).
φ est continue sur $[a, b]$, dérivable sur $]a, b[$ et $\varphi(a) = 0 = \varphi(b)$, donc d'après le lemme de Rolle, il existe $c \in]a, b[$ tel que

$$\begin{aligned} 0 &= \varphi'(c) \\ &= -f'(c) - \sum_{k=1}^{n} \Big(\frac{(b-c)^k}{k!} f^{(k+1)}(c) - \frac{(b-c)^{k-1}}{(k-1)!} f^{(k)}(c) \Big) + \frac{(b-c)^n}{n!}\alpha \\ &= \frac{(b-c)^n}{n!}(-f^{(n+1)}(c) + \alpha). \end{aligned}$$

Ainsi, $\alpha = f^{(n+1)}(c)$ et la relation $\varphi(a) = 0$ dans laquelle on remplace α par $f^{(n+1)}(c)$ donne la formule de Taylor-Lagrange. □

Exercice. Montrer que $\forall x \in \mathbb{R}_+^* \setminus \{1\}$ $\ln x < x - 1$.

Résolution. Soit $x \in \mathbb{R}_+^* \setminus \{1\}$. Appliquons la formule de Taylor-Lagrange à l'ordre 1, pour la fonction ln sur l'intervalle $[\min(x, 1), \max(x, 1)]$.
Il existe $c_x \in]\min(x, 1), \max(x, 1)[$ tel que
$\ln x = \ln 1 + (x-1)\ln' 1 + \dfrac{(x-1)^2}{2} \ln'' c_x$.
Ainsi $\ln x = x - 1 + \dfrac{(x-1)^2}{2} \times \dfrac{-1}{c_x^2} < x - 1$.

Exemple. Appliquons l'inégalité de Taylor-Lagrange à la fonction cos entre 0 et $x \in \mathbb{R}$, à l'ordre 2 : il existe $c \in]0, x[$ tel que $\cos x = \cos 0 - x\sin 0 - \frac{x^2}{2}\cos c$, donc $|\cos x - 1| \leq \frac{x^2}{2}$.

5.2 L'inégalité des accroissements finis (IAF)

Remarque. On a vu que le lemme de Rolle est faux pour des applications à valeurs dans $\mathbb{C}$, donc c'est a fortiori le cas pour le TAF. Cependant :

Théorème. Inégalité des accroissements finis (IAF)
Soit $f\ :[a,b]\longrightarrow\mathbb{K}$ une application de classe C^1 sur $[a,b]$.
Alors $|f(b)-f(a)|\le\lambda|b-a|$, où $\lambda=\sup\limits_{x\in[a,b]}|f'(x)|$.

Démonstration.
$|f(b)-f(a)|=\left|\int_a^b f'(x)dx\right|\le\int_{\min(a,b)}^{\max(a,b)}|f'(x)|dx\le\int_{\min(a,b)}^{\max(a',b')}\lambda dx\le\lambda|b-a|.$ □

Interprétation cinématique. Si l'intensité de la vitesse instantanée d'un mobile est inférieure à λ pour tout instant compris entre a et b, la distance parcourue par le mobile entre les temps a et b est inférieure à $\lambda|b-a|$.

Corollaire. Soient $k\in\mathbb{R}_+$ et $f\ :I\longrightarrow\mathbb{K}$ une application de classe C^1 sur I.
Alors f est k-lipschitzienne si et seulement si pour tout $x\in[a,b]$, $|f'(x)|\le k$.

Démonstration.
Si f est k-lipschitzienne, les taux d'accroissement sont en modules inférieurs à k. Or, $f'(x)$ est la limite lorsque t tend vers 0 du taux d'accroissement entre x et $x+t$, donc $|f'(x)|\le k$.
Réciproque. C'est clair d'après l'inégalité des accroissements finis. □

5.3 Formules de Taylor

5.3.1 TRI et inégalité de TL

Théorème. Formule de Taylor avec reste intégral.
Soient $k\in\mathbb{N}$ et $f\ :[a,b]\longrightarrow\mathbb{K}$ une application C^{k+1} sur $[a,b]$. Alors

$$\boxed{f(b)=f(a)+\sum_{h=1}^{k}\frac{(b-a)^h}{h!}f^{(h)}(a)+\int_a^b\frac{(b-t)^k}{k!}f^{(k+1)}(t)dt.}$$

On note parfois $T_k(b)=f(a)+\sum\limits_{h=1}^{k}\frac{(b-a)^h}{h!}f^{(h)}(a)$, appelé somme de Taylor (ou développement de Taylor) à l'ordre k et $R_k(b)=\int_a^b\frac{(b-t)^k}{k!}f^{(k+1)}(t)dt$, appelé le reste intégral d'ordre k.

Démonstration.
Déjà fait : récurrence + intégration par parties. □

Remarque. On peut déduire l'égalité de Taylor-Lagrange de la formule de Taylor avec reste intégral en utilisant le résultat suivant (hors programme) :

Formule de la moyenne : Soit f et g deux applications continues de $[a,b]$ avec $a<b$. On suppose que $g(t)\geq 0$ pour tout $t\in[a,b]$.
Alors il existe $c\in[a,b]$ tel que $\displaystyle\int_a^b f(t)g(t)\ dt = f(c)\int_a^b g(t)\ dt$.
[*Démonstration :* f étant continue sur le compact $[a,b]$, elle est bornée et elle atteint ses bornes, donc il existe $\alpha,\beta\in[a,b]$ tels que, pour tout $t\in[a,b]$, $f(\alpha)\leq f(t)\leq f(\beta)$. Or si $t\in[a,b]$, $g(t)\geq 0$, donc $f(\alpha)g(t)\leq f(t)g(t)\leq f(\beta)g(t)$, puis en intégrant, $\displaystyle f(\alpha)\int_a^b g(t)\ dt\leq\int_a^b f(t)g(t)\ dt\leq f(\beta)\int_a^b g(t)\ dt$.
Si $\displaystyle\int_a^b g(t)=0$, alors cet encadrement implique que $\displaystyle\int_a^b f(t)g(t)\ dt=0=f(c)\int_a^b g(t)\ dt$ quel que soit $c\in[a,b]$.
Si $\displaystyle\int_a^b g(t)\ dt\neq 0$, alors cette intégrale est strictement positive, donc l'encadrement peut s'écrire $f(\alpha)\leq\dfrac{\int_a^b f(t)g(t)\ dt}{\int_a^b g(t)\ dt}\leq f(\beta)$ et le théorème des valeurs intermédiaires permet de conclure.]
La formule de la moyenne reste valable lorsque $g(t)$ est négatif pour tout $t\in[a,b]$, donc sous les hypothèses et notations de la formule de Taylor avec reste intégral, il existe $c\in[a,b]$ tel que $\displaystyle R_k(b)=f^{(k+1)}(c)\int_a^b\frac{(b-t)^k}{k!}\ dt=\frac{(b-a)^{k+1}}{(k+1)!}f^{(k+1)}(c)$. On retrouve bien l'égalité de Taylor-Lagrange, sous l'hypothèse plus forte que f est de classe C^{k+1}.

Théorème. Inégalite de Taylor-Lagrange.
Soient $k\in\mathbb{N}$ et $f\ :[a,b]\longrightarrow\mathbb{K}$ une application C^{k+1} sur $[a,b]$. Alors

$$\|f(b)-f(a)-\sum_{h=1}^{k}\frac{(b-a)^h}{h!}f^{(h)}(a)\|\leq\lambda\frac{|b-a|^{k+1}}{(k+1)!},\ \text{où } \lambda=\sup_{x\in[a,b]}|f^{(k+1)}(x)|.$$

Démonstration.
Il suffit de majorer correctement le reste $R_k(b)$ de la formule de Taylor avec reste intégral : Supposons que $a<b$ (le cas $a>b$ est laissé en exercice).
$\displaystyle|R_k(b)|\leq\int_a^b\frac{(b-t)^k}{k!}|f^{(k+1)}(t)|\,dt\leq\lambda\int_a^b\frac{(b-t)^k}{k!}\ dt=\lambda\Big[-\frac{(b-t)^{k+1}}{(k+1)!}\Big]_a^b=\lambda\frac{|b-a|^{k+1}}{(k+1)!}$.
□

Exemple. Pour tout $t\in\mathbb{R}$, $|\sin t-t+\dfrac{t^3}{6}|\leq\dfrac{|t|^5}{120}$.

5.3.2 Primitivation d'un développement limité

Lemme. Soit $k\in\mathbb{N}$. Au voisinage de a, $\displaystyle\int_a^x o((t-a)^k)dt=o((x-a)^{k+1})$.

Démonstration.
Soit η une application à valeurs dans $\mathbb{K}$, définie sur un intervalle I contenant a, continue et telle qu'au voisinage de a, $\eta(x) = o((x-a)^k)$.
Soit $\varepsilon > 0$. Il existe $\alpha > 0$ tel que pour tout $t \in I$, $(|t-a| \leq \alpha \Longrightarrow \|\eta(t)\| \leq \varepsilon |t-a|^k)$.
Soit $x \in I$ tel que $|x-a| \leq \alpha$.
$|\int_a^x \eta(t)dt| \leq \int_{\min(a,x)}^{\max(a,x)} \varepsilon |t-a|^k dt \leq \int_{\min(a,x)}^{\max(a,x)} \varepsilon |x-a|^k dt \leq \varepsilon |x-a|^{k+1}$. □

Théorème. Primitivation d'un développement limité.
Soient $a \in I$ et $f : I \longrightarrow \mathbb{K}$ une application de classe C^1. Soit $k \in \mathbb{N}$.
On suppose que f' admet au voisinage de a un développement limité de la forme suivante :

$$f'(x) = \sum_{h=0}^{k} \alpha_h (x-a)^h + o((x-a)^k).$$

Alors f admet au voisinage de a le développement limité suivant :

$$f(x) = f(a) + \sum_{h=0}^{k} \frac{\alpha_h}{h+1}(x-a)^{h+1} + o((x-a)^{k+1}).$$

Démonstration.
$f(x) = f(a) + \int_a^x f'(t)dt$, puis on applique le lemme. □

Remarque. Ainsi, lorsque l'on sait que f' possède un développement limité, ses coefficients sont obtenus par dérivation du polynôme intervenant dans le développement limité de f.

Exemple. Pour tout $k \in \mathbb{N}$,
$\arctan x = \displaystyle\int_0^x \frac{dt}{1+t^2} = \int_0^x \left(\sum_{h=0}^{k} (-t^2)^h + o(t^{2k}) \right) dt = \sum_{h=0}^{k} (-1)^h \frac{x^{2h+1}}{2h+1} + o(x^{2k+1}).$

5.3.3 Formule de TY

Formule de Taylor-Young. Soient $a \in I$, $k \in \mathbb{N}^*$ et $f : I \longrightarrow \mathbb{K}$ une application k fois dérivable sur un intervalle voisinage de a. Alors f admet au voisinage de a un développement limité à l'ordre k. Plus précisément,

$$f(x) = f(a) + \sum_{h=1}^{k} \frac{(x-a)^h}{h!} f^{(h)}(a) + o((x-a)^k).$$

Démonstration.
Soit $k \in \mathbb{N}^*$. Notons $R(k)$ l'assertion suivante : pour tout $f : I \longrightarrow \mathbb{K}$ k fois dérivable sur un intervalle voisinage de a, $f(x) = f(a) + \displaystyle\sum_{h=1}^{k} \frac{(x-a)^h}{h!} f^{(h)}(a) + o((x-a)^k)$.

$\diamond$ Pour $k = 1$, on sait que $f : I \longrightarrow \mathbb{K}$ est dérivable si et seulement s'il existe $l \in E$ tel que $f(x) = f(a) + (x-a)l + o(x-a)$ et que dans ce cas, $l = f'(a)$.
$\diamond$ Pour $k \geq 1$, on suppose $R(k)$. Soit $f : I \longrightarrow \mathbb{K}$ une application $k+1$ fois dérivable sur un voisinage de a. D'après $R(k)$ appliquée à f',
$$f'(x) = f'(a) + \sum_{h=1}^{k} \frac{(x-a)^h}{h!} f^{(h+1)}(a) + o((x-a)^k).$$
f est deux fois dérivable sur un voisinage a, donc elle est de classe C^1 sur un voisinage de a. On peut donc appliquer le théorème d'intégration d'un développement limité à l'égalité précédente, et l'on obtient $R(k+1)$. $\square$

Remarque. On admettra que la formule de TY reste valable si l'on suppose seulement que f est k fois dérivable en a.

Remarque. La formule de Taylor-Young est encore valable lorsque $a = \min(I)$ (en supposant que ce minimum existe) ou bien lorsque $a = \max(I)$, en utilisant $f_d^{(k)}(a)$ ou bien $f_g^{(k)}(a)$. Ce fut utile pour étudier des raccordements de solutions d'équations différentielles d'ordre 2 (cf cours sur les équations différentielles page 7).

Remarque. Taylor-Young donne une information au sujet du comportement de f qui est locale au point a. Au contraire, Taylor-Lagrange donne une information au sujet du comportement de f qui est globale sur un intervalle. Ces deux formules ne sont donc pas de même nature malgré la similitude de leur énoncé.

Remarque. Il existe des applications admettant un développement limité à l'ordre 2 au voisinage d'un point mais qui ne sont pas 2 fois dérivables en ce point.
Par exemple si l'on pose $f(x) = x^3 \sin \frac{1}{x}$ pour $x \neq 0$ et $f(0) = 0$,
alors $f(x) = O(x^3) = o(x^2)$ au voisinage de 0, ce qui constitue un développement limité de f à l'ordre 2.
Pour $x \neq 0$, $f'(x) = 3x^2 \sin \frac{1}{x} - x \cos \frac{1}{x}$ et d'après le théorème de la limite de la dérivée, f est dérivable sur $\mathbb{R}$ avec $f'(0) = 0$, mais $\dfrac{f'(x)}{x} = 3x \sin \dfrac{1}{x} - \cos \dfrac{1}{x}$, donc f n'est pas deux fois dérivable en 0.

Exemple. $\dfrac{1}{(1-x)} = 1 + \displaystyle\int_0^x \frac{dt}{(1-t)^2}$, or d'après la formule de TY,
il existe $c_0, \ldots, c_n \in \mathbb{R}$ tel que, au voisinage de 0, $\dfrac{1}{(1-t)^2} = c_0 + c_1 t + \cdots + c_n t^n + o(t^n)$.
D'après le théorème de primitivation d'un développement limité,
$\dfrac{1}{1-x} = 1 + c_0 x + \dfrac{c_1}{2} x^2 + \cdots + \dfrac{c_n}{n+1} x^{n+1} + o(x^{n+1})$, mais on sait que
$\dfrac{1}{1-x} = 1 + x + x^2 + \cdots + x^{n+1} + o(x^{n+1})$, donc d'après l'unicité du développement limité, $\dfrac{1}{(1-x)^2} = 1 + 2x + 3x^2 + \cdots + (n+1)x^n + o(x^n)$.
Cette formule n'est cependant qu'un cas particulier du développement limité connu de $(1+t)^\alpha$, lorsque $\alpha = -2$.

Remarque. La formule de Taylor-Young est pratique pour calculer les dérivées successives d'une fonction en un point, sans calculer ses dérivées en tout point.

Exercice. Calculer les dérivées à tout ordre en 0 de $\begin{array}{rccc} f: &]-1,1[& \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \dfrac{1}{1-x^3} \end{array}$.

Solution. Soit $k \in \mathbb{N}$. Au voisinage de 0, $\dfrac{1}{1-x^3} = \displaystyle\sum_{h=0}^{k} x^{3h} + o(x^{3k})$.

De plus f est de classe C^∞, donc on peut appliquer la formule de Taylor-Young.

Ainsi $\dfrac{1}{1-x^3} = \displaystyle\sum_{h=0}^{3k} \frac{f^{(h)}(0)}{h!} x^h + o(x^{3k})$.

D'après l'unicité du développement limité,
pour tout $h \in \mathbb{N}$, $f^{(h)}(0) = 0$ si h n'est pas un multiple de 3 et $f^{(3h)}(0) = (3h)!$.

Propriété. (Hors programme ?) Soit $f : I \longrightarrow \mathbb{R}$ une application deux fois dérivable en un point a de $\overset{\circ}{I}$. On suppose que $f'(a) = 0$ et que $f''(a) > 0$.
Alors a est un minimum local strict : il existe un voisinage V de a tel que pour tout $t \in V \cap I \setminus \{a\}$, $f(t) > f(a)$.

Démonstration.
D'après la formule de Taylor-Young, au voisinage de a,
$f(x) = f(a) + \dfrac{(x-a)^2}{2} f''(a) + o((x-a)^2)$, donc $f(x) - f(a) \sim \dfrac{(x-a)^2}{2} f''(a) > 0$. □

6 Monotonie et dérivabilité

Jusqu'à la fin de ce chapitre, $\mathbb{K} = \mathbb{R}$, donc sauf précision du contraire, toutes les applications utilisées sont maintenant définies sur I et sont à valeurs dans $\mathbb{R}$.

6.1 Sens de variation

Théorème. f est constante si et seulement si $f' = 0$, elle est croissante si et seulement si $f' \geq 0$ et elle est décroissante si et seulement si $f' \leq 0$.

Démonstration.
◇ Supposons que f est croissante. Soit $a \in I$ avec $a \neq Sup(I)$ et $a \neq Inf(I)$. Il existe $\varepsilon > 0$ tel que $[a-\varepsilon, a+\varepsilon[\subset I$. Pour tout $x \in]a, a+\varepsilon[$, $x - a > 0$ et $f(x) - f(a) \geq 0$, donc $\dfrac{f(x)-f(a)}{x-a} \geq 0$, or $f'(a) = \displaystyle\lim_{\substack{x \to a \\ x \in]a, a+\varepsilon[}} \frac{f(x)-f(a)}{x-a}$, donc $f'(a) \geq 0$.
On aurait pu également raisonner avec l'intervalle $]a-\varepsilon, a[$, donc ce résultat est aussi valable lorsque a est une borne de l'intervalle I. Ainsi, pour tout $x \in I$, $f'(x) \geq 0$.
◇ Réciproquement, supposons que $f' \geq 0$. Soit $(a, b) \in I^2$ tel que $a < b$. D'après la formule des accroissements finis, il existe $c \in]a, b[$ tel que $f(b) - f(a) = (b-a)f'(c)$, donc $f(b) - f(a) \geq 0$, puis $f(b) \geq f(a)$. Ainsi f est croissante.

◇ f est décroissante si et seulement si $-f$ est croissante, donc si et seulement si $(-f)' \geq 0$, c'est-à-dire si et seulement si $f' \leq 0$.
◇ f est constante si et seulement si f est croissante et décroissante, donc si et seulement si f' est identiquement nulle. □

Remarque. Soit $f : I \longrightarrow \mathbb{R}$ une application dérivable et croissante. Si elle n'est pas strictement croissante, il existe $(a, b) \in I^2$ avec $a < b$ tel que $f(a) = f(b)$. Alors f est constante sur l'intervalle $[a, b]$. Par contraposée, lorsque f n'est constante sur aucun sous-intervalle de I contenant plus de deux réels, f est strictement croissante.
La réciproque étant évidente, on obtient que, si f est croissante, alors f est strictement croissante si et seulement si l'intérieur de $f'^{-1}(\{0\})$ est vide.
En particulier, si $f'(x) > 0$ pour tout $x \in I$ sauf pour un nombre fini d'éléments de I, f est strictement croissante.

6.2 Difféomorphismes

Théorème. Supposons que f est dérivable et strictement monotone.
Soit $t \in I$. L'application $f^{-1} : f(I) \longrightarrow I$ est dérivable en $f(t)$ si et seulement si $f'(t) \neq 0$, et dans ce cas

$$(f^{-1})'(f(t)) = \frac{1}{f'(t)}.$$

Ainsi, lorsque $[\forall t \in I,\ \ f'(t) \neq 0]$, f^{-1} est dérivable sur $f(I)$ et

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

Démonstration.
Supposons d'abord que $f'(t) \neq 0$.
Soit $y \in f(I)$ avec $y \neq f(t)$. Posons $x = f^{-1}(y)$.
$\dfrac{f^{-1}(y) - f^{-1}(f(t))}{y - f(t)} = \dfrac{x - t}{f(x) - f(t)} = \dfrac{1}{\frac{f(x)-f(t)}{x-t}}$.
f^{-1} étant continue en $f(t)$, $x = f^{-1}(y) \underset{y \to f(t)}{\longrightarrow} t$, et f étant dérivable en t,
$\dfrac{f(x) - f(t)}{x - t} \underset{\substack{x \to t \\ t \neq x}}{\longrightarrow} f'(t)$, donc par composition des limites,
$\dfrac{f^{-1}(y) - f^{-1}(f(t))}{y - f(t)} \underset{y \to f(t)}{\longrightarrow} \dfrac{1}{f'(t)}$.
◇ Supposons maintenant que $f'(t) = 0$. Le calcul précédent montre alors que
$\left|\dfrac{f^{-1}(y) - f^{-1}(f(t))}{y - f(t)}\right| \underset{y \to f(t)}{\longrightarrow} +\infty$, donc le graphe de f^{-1} présente en $f(t)$ une tangente verticale. □

Interprétation géométrique.

Définition. Soient I et J deux intervalles inclus dans $\mathbb{R}$ d'intérieurs non vides, et $f : I \longrightarrow J$ une application. Soit $n \in \mathbb{N}^*$. On dit que f est un C^n-difféomorphisme si et seulement si f est bijective, de classe C^n et si f^{-1} est aussi de classe C^n.

Théorème. $f : I \longrightarrow \mathbb{R}$ est un C^n-difféomorphisme de I dans $f(I)$ si et seulement si f est de classe C^n et si $[\forall t \in I,\ \ f'(t) \neq 0]$.

Démonstration.
Supposons que f est C^n avec $f'(t) \neq 0$ pour tout $t \in I$.
f' est continue et ne s'annule jamais, donc d'après le théorème des valeurs intermédiaires, elle est strictement positive ou strictement négative. Ainsi f est strictement monotone, donc f est injective. On peut donc parler de son application réciproque f^{-1} de $f(I)$ dans I.
On sait alors que f^{-1} est dérivable avec $(f^{-1})' = \dfrac{1}{f' \circ f^{-1}}$. Cette formule permet de montrer par récurrence sur k que f^{-1} est C^k, pour tout $k \in \{0, \ldots, n\}$. □

7 Suites récurrentes d'ordre 1

Notations On suppose que $f : I \longrightarrow I$ est une application continue et monotone.
Soit $x_0 \in I$. On souhaite étudier la suite (x_n) définie par la condition

$$\forall n \in \mathbb{N} \quad x_{n+1} = f(x_n).$$

Pour $n \in \mathbb{N}$, f^n désignera $f \underbrace{\circ \cdots \circ}_{n \text{ fois}} f$.

Représentation graphique.

Remarque. En pratique, le plus souvent, on donne x_0 et f, mais l'intervalle I tel que $x_0 \in I$, $f(I) \subset I$ et $f_{/I}$ est monotone et continue, n'est pas donné. Il faudra donc commencer par rechercher un tel intervalle. Un bon moyen est de dresser le tableau de variations de f.

Propriété. Si $x_n \underset{n \to +\infty}{\longrightarrow} l$ et si $l \in I$, alors l est un point fixe de f (ie : $f(l) = l$).

Ainsi les "limites éventuelles" de (x_n) sont les points fixes de f, et les bornes de I qui n'appartiennent pas à I.

Propriété. Si f est croissante, alors (x_n) est monotone.
Plus précisément, (x_n) est croissante si et seulement si $f(x_0) - x_0 \geq 0$,
et (x_n) est décroissante si et seulement si $f(x_0) - x_0 \leq 0$.

Démonstration.
Supposons que $x_1 = f(x_0) \leq x_0$. Soit $n \in \mathbb{N}$. f^n est croissante, donc
$x_{n+1} = f^n(x_1) \leq f^n(x_0) = x_n$. Ainsi la suite (x_n) est décroissante. □

Propriété. On suppose que f est croissante. Soit $l \in I$ un point fixe de f.
Si $x_0 \leq l$, alors $\forall n \in \mathbb{N} \quad x_n \leq l$.
Si $x_0 \geq l$, alors $\forall n \in \mathbb{N} \quad x_n \geq l$.

Démonstration.
Exercice, en s'inspirant de la démonstration précédente. □

Remarque. La combinaison de ces deux propriétés permet souvent de montrer que (x_n) est monotone et bornée, donc qu'elle converge. Pour appliquer ces propriétés, il faut étudier le signe et les zéros de $x \longmapsto f(x) - x$, ce que l'on peut faire dès le début en même temps que la détermination du tableau de variations de f.

Propriété.

On suppose que f est décroissante. Alors $f \circ f$ est croissante. Les deux suites (x_{2n}) et (x_{2n+1}) sont monotones et de sens contraires.

Remarque. Les deux propriétés qui suivent montrent qu'il y a essentiellement 4 comportements différents au voisinage d'un point fixe l selon que f est croissante ou décroissante et que $|f'(l)| > 1$ ou $|f'(l)| < 1$.

Propriété. Soit f une application de classe C^1 de I dans $\mathbb{R}$.
Soit $\ell \in \overset{\circ}{I}$ tel que $|f'(\ell)| < 1$. Alors il existe $\varepsilon > 0$ tel que l'intervalle $J =]\ell - \varepsilon, \ell + \varepsilon[$ est inclus dans I, stable par f (c'est-à-dire $f(J) \subset J$) et tel que, pour tout $x_0 \in J$, en posant $x_{n+1} = f(x_n)$, $x_n \underset{n \to +\infty}{\longrightarrow} \ell$.

Démonstration.
$\ell \in \overset{\circ}{I}$, donc il existe $\alpha > 0$ tel que $]\ell - \alpha, \ell + \alpha[\subset I$.

Posons $k = \frac{1}{2}(1 + |f'(\ell)|)$. $|f'(\ell) < k < 1$ et f' est continue en ℓ, donc il existe $\varepsilon \in]0, \alpha]$ tel que pour tout $x \in J =]\ell - \varepsilon, \ell + \varepsilon[$, $|f'(x)| \leq k$.
Soit $y \in J$. D'après l'IAF, $|f(y) - \ell| = |f(y) - f(\ell)| \leq k|y - \ell| < \varepsilon$, donc $f(y) \in J$. Ainsi, J est stable par f.
Soit $x_0 \in J$. Alors en posant $x_{n+1} = f(x_n)$, par récurrence, on montre que, pour tout $n \in \mathbb{N}$, $x_n \in J$.
Alors, pour tout $n \in \mathbb{N}$, $|x_{n+1} - \ell| = |f(x_n) - f(\ell)| \leq k|x_n - \ell|$, donc par récurrence, $|x_n - \ell| \leq k^n|x_0 - \ell| \underset{n \to +\infty}{\longrightarrow} 0$. □

Remarque.
On peut faire le lien avec le théorème du point fixe vu en TD (exercice 18.15).

Propriété. Soit f une application de classe C^1 de I dans $\mathbb{R}$.
Soit $\ell \in \overset{\circ}{I}$ tel que $|f'(\ell)| > 1$. Alors il existe $\varepsilon > 0$ tel que,
pour tout $x_0 \in J =]\ell - \varepsilon, \ell + \varepsilon[\setminus \{\ell\}$, en posant $x_{n+1} = f(x_n)$,
il existe $n \in \mathbb{N}$ tel que $x_n \notin J$.

Démonstration.
Posons $k = \frac{1}{2}(1 + |f'(\ell)|)$. Ainsi, $1 < k < |f'(\ell)|$. Comme précédemment, on montre qu'il existe $\varepsilon > 0$ tel que $J =]\ell - \varepsilon, \ell + \varepsilon[$ est inclus dans I et tel que, pour tout $x \in J$, $|f'(x)| \geq k$.
Soit $x_0 \in J$. Supposons que pour tout $n \in \mathbb{N}$, $x_n \in J$.

Soit $n \in \mathbb{N}$. D'après l'égalité des accroissements finis, il existe $c_n \in]x_{n+1}, \ell[$ tel que $|x_{n+1} - \ell| = |f(x_n) - f(\ell)| = |f'(c)||x_n - \ell|$, donc $|x_{n+1} - \ell| \geq k|x_n - \ell|$, puis par récurrence, $|x_n - \ell| \geq k^n|x_0 - \ell| \underset{n\to+\infty}{\longrightarrow} +\infty$. Ainsi, $|x_n - \ell| \underset{n\to+\infty}{\longrightarrow} +\infty$ et $|x_n - \ell| \leq \varepsilon$. C'est impossible. □

Conclusion. Pour étudier une suite (x_n) vérifiant une relation de récurrence de la forme $x_{n+1} = f(x_n)$, on peut s'inspirer du plan suivant.

- ⋄ Etudiez f (tableau de variations).
- ⋄ Lorsque le graphe de f est simple, conjecturer graphiquement le comportement de la suite (x_n).
- ⋄ Trouvez un intervalle I tel que $f(I) \subset I$ et $x_0 \in I$ et f est monotone et continue sur I.
- ⋄ Recherchez les "limites éventuelles".
- ⋄ Si f est croissante sur I, étudiez les signes de $f(x_0) - x_0$ et de $x_0 - l$ (où l est un point fixe), puis conclure.
- ⋄ Si f est décroissante sur I, se ramener au cas précédent en considérant $f \circ f$.

8 Fonctions convexes

8.1 Sous-espaces affines

Notation. Dans ce paragraphe, E désigne un $\mathbb{K}$-espace vectoriel quelconque.

Définition. Soient $\mathcal{E}$ un $\mathbb{K}$-espace affine de direction E et $\mathcal{F}$ une partie de $\mathcal{E}$.
$\mathcal{F}$ est un ***sous-espace affine*** de $\mathcal{E}$ si et seulement si il existe $A \in \mathcal{E}$ et un sous-espace vectoriel F de E tel que $\boxed{\mathcal{F} = A + F = \{A + x \ / \ x \in F\}}$.
Dans ce cas, $F = \{\overrightarrow{MN} \ / \ M, N \in \mathcal{F}\}$: on dit que F est la direction du sous-espace affine $\mathcal{F}$. De plus, pour tout $B \in \mathcal{F}$, $\mathcal{F} = B + F$.

Démonstration.
Exercice. □

Exemples. Un singleton est un sous-espace affine dirigé par $\{0\}$.
Une droite affine de $\mathcal{E}$ est de la forme $\mathcal{D} = A + \mathbb{K}x$, où $A \in \mathcal{E}$ et $x \in E \setminus \{0\}$: on dit que x est un vecteur directeur de $\mathcal{D}$.

Propriété. Soit E et F deux $\mathbb{K}$-espaces vectoriels et $f \in L(E, F)$. Soit $y \in F$. L'ensemble des solutions de l'équation linéaire $(E)\ :\ f(x) = y$ en l'inconnue $x \in E$, est ou bien vide, ou bien un sous-espace affine de E

Démonstration.
On a vu que, si x_0 est une solution particulière de (E), alors $\mathcal{S}_E = x_0 + \mathrm{Ker}(f)$. □

Définition.
Deux sous-espaces affines sont parallèles si et seulement si ils ont la même direction.

Propriété. Soient $\mathcal{E}$ un $\mathbb{K}$-espace affine de direction E et $(\mathcal{E}_i)_{i\in I}$ une famille de sous-espaces affines de $\mathcal{E}$. Pour $i \in I$, on note E_i la direction de $\mathcal{E}_i$.
$\bigcap_{i\in I} \mathcal{E}_i$ est ou bien $\emptyset$, ou bien un sous-espace affine de $\mathcal{E}$ de direction $\bigcap_{i\in I} E_i$.

Démonstration.
Supposons que $\bigcap_{i\in I} \mathcal{E}_i$ est non vide : il existe $A \in \bigcap_{i\in I} \mathcal{E}_i$. Alors, pour tout $i \in I$,
$\mathcal{E}_i = A+E_i$, donc pour tout $M \in \mathcal{E}$, $M \in \mathcal{E}_i \Longleftrightarrow [\exists x \in E_i,\ \ M = A+x] \Longleftrightarrow \overrightarrow{AM} \in E_i$, puis $M \in \bigcap_{i\in I} \mathcal{E}_i \Longleftrightarrow \forall i \in I,\ \ \overrightarrow{AM} \in E_i \Longleftrightarrow \overrightarrow{AM} \in \bigcap_{i\in I} E_i$,
ce qui prouve que $\bigcap_{i\in I} \mathcal{E}_i = A + \bigcap_{i\in I} E_i$. □

Exemple. L'intersection de deux droites affines est soit vide, soit un singleton, soit une droite affine.

Définition. Soit $\mathcal{E}$ un $\mathbb{K}$-espace affine de direction E. Un repère de $\mathcal{E}$ est un couple $R = (O, b)$, où O est un point de $\mathcal{E}$, appelé l'origine du repère et où b est une base de E. Si $M \in \mathcal{E}$, les coordonnées de M dans le repère R sont les coordonnées du vecteur $\overrightarrow{OM}$ dans la base b.

Définition. Soient $\mathcal{E}$ un $\mathbb{K}$-espace affine et $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$ dont la direction est notée F.
On dit que $\mathcal{F}$ est de dimension finie si et seulement si F est de dimension finie et, dans ce cas, la dimension de F sera aussi appelée la dimension de $\mathcal{F}$.

8.2 Barycentres et convexité

Notation. On fixe un $\mathbb{K}$-espace vectoriel E, un espace affine $\mathcal{E}$ de direction E, p points $A_1, \ldots, A_p$ de $\mathcal{E}$ et p scalaires $\lambda_1, \ldots, \lambda_p$ dans $\mathbb{K}$.

Définition. On appelle fonction vectorielle de Leibniz l'application $\varphi\ :\ \mathcal{E} \longrightarrow E$ définie par $\varphi(M) = \sum_{i=1}^{p} \lambda_i \overrightarrow{A_iM}$.

Propriété. Lorsque $\sum_{i=1}^{p} \lambda_i = 0$, φ est constante, et lorsque $\sum_{i=1}^{p} \lambda_i \neq 0$, φ est bijective.
L'unique point G tel que $\varphi(G) = 0$ s'appelle alors le barycentre des $(A_i, \lambda_i)_{1\leq i\leq p}$.
On a donc $\sum_{i=1}^{p} \lambda_i \overrightarrow{GA_i} = 0$.

On en déduit que, pour tout $M \in \mathcal{E}$, $\overrightarrow{MG} = \dfrac{1}{\sum_{i=1}^{p} \lambda_i} \sum_{i=1}^{p} \lambda_i \overrightarrow{MA_i}$.

Ceci justifie la notation $G \triangleq \dfrac{\lambda_1 A_1 + \cdots \lambda_p A_p}{\lambda_1 + \cdots + \lambda_p}$.

Démonstration.
Soit G un point quelconque de E. D'après la relation de Chasles, pour tout $M \in \mathcal{E}$, $\varphi(M) = \sum_{i=1}^{p} \lambda_i(\overrightarrow{A_iG} + \overrightarrow{GM}) = \varphi(G) + \Big(\sum_{i=1}^{p} \lambda_i\Big)\overrightarrow{GM}$.

Ainsi, lorsque $\sum_{i=1}^{p} \lambda_i = 0$, φ est constante. Supposons maintenant que $\sum_{i=1}^{p} \lambda_i \neq 0$. Posons $\Lambda = \sum_{i=1}^{p} \lambda_i$. Alors d'après le calcul précédent, pour tout $x \in E$, pour tout $M \in \mathcal{E}$, $\varphi(M) = x \iff \overrightarrow{GM} = \frac{1}{\Lambda}(x - \varphi(G)) \iff M = G + \frac{1}{\Lambda}(x - \varphi(G))$, ce qui prouve que φ est bijective.

Si G désigne maintenant l'unique antécédent du vecteur nul, on a bien $0 = \sum_{i=1}^{p} \lambda_i \overrightarrow{GA_i}$, puis d'après la relation de Chasles, pour tout $M \in \mathcal{E}$, $0 = \Lambda\overrightarrow{GM} + \sum_{i=1}^{p} \lambda_i \overrightarrow{MA_i}$, ce qui fournit la dernière relation de l'énoncé. $\square$

Définition. Lorsque, pour tout $i \in \mathbb{N}_p$, $\lambda_i = 1$, G s'appelle l'isobarycentre des points $A_1, \ldots, A_p$.

Propriété. ***Homogénéïté du barycentre :*** Si l'on remplace chaque λ_i par $\alpha\lambda_i$ où $\alpha \in \mathbb{K} \setminus \{0\}$, G n'est pas modifié.

Démonstration.
D'après la propriété précédente, $\sum_{i=1}^{p} \lambda_i \overrightarrow{GA_i} = 0$, donc on a aussi $\sum_{i=1}^{p} \alpha\lambda_i \overrightarrow{GA_i} = 0$, ce qui permet de conclure par définition du barycentre. $\square$

Propriété. ***Associativité du barycentre :*** Soit $k \in \mathbb{N}_p$. Notons G' le barycentre des $(A_i, \lambda_i)_{1 \leq i \leq k}$ (on suppose que $\lambda' = \sum_{i=1}^{k} \lambda_i \neq 0$) et G'' le barycentre des $(A_i, \lambda_i)_{k+1 \leq i \leq p}$ (on suppose que $\lambda'' = \sum_{i=k+1}^{p} \lambda_i \neq 0$). Alors G est le barycentre de $((G', \lambda'), (G'', \lambda''))$.

Démonstration.
$0 = \sum_{i=1}^{p} \lambda_i \overrightarrow{GA_i} = \sum_{i=1}^{k} \lambda_i \overrightarrow{GA_i} + \sum_{i=k+1}^{p} \lambda_i \overrightarrow{GA_i} = \Big(\sum_{i=1}^{k} \lambda_i\Big)\overrightarrow{GG'} + \Big(\sum_{i=k+1}^{p} \lambda_i\Big)\overrightarrow{GG''}$. $\square$

Exemple. L'isobarycentre d'un triangle A, B, C est le barycentre de $(A', 2)$ et de $(C, 1)$, où A' est le milieu de (A, B), donc il est situé sur la médiane $(A'C)$. De même, il est situé sur les deux autres médianes, ce qui prouve qu'elles se rencontrent en un unique point, égal au centre de gravité du triangle.

Propriété. Soit $\mathcal{F}$ un sous-espace affine de $\mathcal{E}$.
Si pour tout $i \in \mathbb{N}_p$, $A_i \in \mathcal{F}$, alors $G \in \mathcal{F}$.

Exemple. Si A et B sont deux points distincts de $\mathcal{E}$, la droite (AB) est égale à l'ensemble des barycentres de A et B.
Si A, B et C sont trois points non alignés de $\mathcal{E}$, l'ensemble des barycentres de A, B et C est l'unique plan affine contenant ces trois points.

Démonstration.
Notons F la direction de $\mathcal{F}$. Il existe $B \in \mathcal{E}$ tel que $\mathcal{F} = B + F$.
$\overrightarrow{BG} = \dfrac{1}{\displaystyle\sum_{i=1}^{p} \lambda_i} \displaystyle\sum_{i=1}^{p} \lambda_i \overrightarrow{BA_i} \in F$. □

Définition. Une partie $\mathcal{C}$ de $\mathcal{E}$ est convexe si et seulement si elle vérifie l'une des propriétés équivalentes suivantes :

1. Pour tout $(A_1, A_2) \in \mathcal{C}^2$, $[A_1, A_2] \subset \mathcal{C}$, où $[A_1, A_2]$ est le segment d'extrémités A_1 et A_2, c'est-à-dire l'ensemble des barycentres de $((A_1, t), (A_2, 1-t))$, lorsque t décrit $[0, 1]$.
2. Pour tout $(A_1, A_2) \in \mathcal{C}^2$, pour tout $(\lambda_1, \lambda_2) \in \mathbb{R}_+^2 \setminus \{0\}$, le barycentre de $((A_1, \lambda_1), (A_2, \lambda_2))$ est dans $\mathcal{C}$.
3. Pour tout $p \in \mathbb{N}^*$, pour tout $(A_i)_{1 \leq i \leq p} \in \mathcal{C}^p$, pour tout $(\lambda_i)_{1 \leq i \leq p} \in \mathbb{R}_+^p \setminus \{0\}$, le barycentre de $(A_i, \lambda_i)_{1 \leq i \leq p}$ est dans $\mathcal{C}$.

Une partie est donc convexe si et seulement si elle est stable *par barycentration à coefficients positifs*. **Exemple.** Les sous-espaces affines sont des convexes.

On a vu que les parties convexes de $\mathbb{R}$ sont exactement les intervalles.

Propriété. Une intersection de parties convexes est convexe.

Définition. Soit B une partie de $\mathcal{E}$.
L'enveloppe convexe de B est le plus petit convexe de $\mathcal{E}$ contenant B.

8.3 Inégalités de convexité

Définition. f est une application convexe si et seulement si

$$\forall (x, y) \in I^2 \quad \forall \alpha \in [0, 1] \quad f(\alpha x + (1-\alpha) y) \leq \alpha f(x) + (1-\alpha) f(y).$$

f est concave si et seulement si $-f$ est convexe.

Interprétation géométrique. f est convexe si et seulement si tout sous-arc de son graphe est situé sous sa corde.

Démonstration.
Fixons $x, y \in I$. Le sous-arc du graphe de f pour u variant entre x et y est $\{(u, f(u)) \ / \ u \in [x, y]\} = \{((1-t)x + ty, f((1-t)x + ty)) \ / \ t \in [0,1]\}$ et la sous-corde correspondante est égale au segment joignant les points $(x, f(x))$ et $(y, f(y))$, c'est-à-dire l'ensemble des $\{((1-t)x + ty, (1-t)f(x) + tf(y)) \ / \ t \in [0,1]\}$, donc le sous-arc est situé sous sa corde si et seulement si pour tout $t \in [0,1]$,
$f((1-t)x + ty) \leq (1-t)f(x) + tf(y)$. □

Remarque. On peut également définir la stricte convexité et la stricte concavité, en remplaçant l'inégalité large par une inégalité stricte lorsque $\alpha \in]0,1[$.

Exemple.
- Une fonction affine, de la forme $x \longmapsto ax + b$, est à la fois concave et convexe. On montrera ci-dessous que ce sont les seules applications à la fois concaves et convexes.
- L'application valeur absolue est convexe car, d'après l'inégalité triangulaire, $|\alpha x + (1-\alpha)y| \leq \alpha|x| + (1-\alpha)|y|$.
- Une somme d'un nombre fini d'applications convexes est convexe. C'est aussi le cas d'une combinaison linéaire à coefficients positifs de fonctions convexes.

Définition. Soit $x_0 \in \overset{\circ}{I}$. On dit que c'est un point d'inflexion de f si et seulement si il existe $\varepsilon > 0$ tel que $f|_{I \cap]x_0 - \varepsilon, x_0]}$ est convexe (resp : concave) et $f|_{I \cap [x_0, x_0 + \varepsilon[}$ est concave (resp : convexe), c'est-à-dire si et seulement si le graphe de f change de concavité au voisinage de x_0.

Propriété. L'épigraphe de f désigne la partie de $\mathbb{R}^2$ située au dessus du graphe de f, c'est-à-dire $\{(x, y) \in \mathbb{R}^2 / x \in I \text{ et } y \geq f(x)\}$.
f est convexe si et seulement si son épigraphe est une partie convexe de $\mathbb{R}^2$.

Démonstration.
Notons E l'épigraphe de f.
$\diamond$ Supposons que l'épigraphe est convexe. Soit $x, y \in I$ et $\alpha \in [0,1]$.
Les points $(x, f(x))$ et $(y, f(y))$ sont dans E et E est convexe, donc $\alpha(x, f(x)) + (1-\alpha)(y, f(y)) \in E$, c'est-à-dire $\alpha f(x) + (1-\alpha)f(y) \geq f(\alpha x + (1-\alpha)y)$.
Ainsi, f est convexe.
$\diamond$ Réciproquement, supposons que f est convexe. Soit $(x, y), (x', y') \in E$ et $\alpha \in [0,1]$.

$\alpha y+(1-a)y' \geq \alpha f(x)+(1-\alpha)f(x') \geq f(\alpha x+(1-\alpha)x')$ car f est convexe, donc $(\alpha x+(1-\alpha)x', \alpha y+(1-\alpha)y') \in E$, ce qui prouve que E est convexe. □

Propriété. ***Inégalité de Jensen.***

f est convexe si et seulement si
$\forall n \in \mathbb{N}^* \ \forall (x_1, \ldots, x_n) \in I^n \ \forall (\lambda_1, \ldots, \lambda_n) \in \mathbb{R}_+^n$,
$$\sum_{i=1}^n \lambda_i = 1 \Longrightarrow f\Big(\sum_{i=1}^n \lambda_i x_i\Big) \leq \sum_{i=1}^n \lambda_i f(x_i).$$

Démonstration.
f étant convexe, son épigraphe est une partie convexe de $\mathbb{R}^2$, donc ce dernier est stable par barycentration à coefficients positifs. □

Exercice. Soit $(x_1, \ldots, x_n) \in \mathbb{R}_+^n$. La moyenne arithmétique est $\displaystyle\frac{1}{n}\sum_{i=1}^n x_i$ et la moyenne géométrique est $\displaystyle\prod_{i=1}^n x_i^{\frac{1}{n}}$. Montrer que la moyenne géométrique est inférieure à la moyenne arithmétique.

Résolution. S'il existe $i \in \mathbb{N}_n$ tel que $x_i = 0$, l'inégalité est simple à établir.
Sinon, pour tout $i \in \mathbb{N}_n$, $x_i > 0$ et l'inégalité à établir est équivalente à
$\displaystyle\sum_{i=1}^n \frac{1}{n}\ln(x_i) \leq \ln\Big(\frac{1}{n}\sum_{i=1}^n x_i\Big)$, laquelle est un cas particulier de l'inégalité de Jensen pour la fonction ln qui est concave car $\ln''(x) = -\dfrac{1}{x^2}$.

8.4 Croissance des pentes

Propriété. pour tout $x_0 \in I$, on définit l'application "pente" :
$$p_{x_0} : \begin{array}{ccc} I \setminus \{x_0\} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{f(x)-f(x_0)}{x-x_0} \end{array}$$
: $p_{x_0}(x)$ est la pente de la sécante d'extrémités $(x_0, f(x_0))$ et $(x, f(x))$.
Les propriétés suivantes sont équivalentes :

1. f est convexe sur I.
2. Pour tout $a, b, c \in I$ avec $a < b < c$, $p_a(b) \leq p_a(c)$.
3. Pour tout $a, b, c \in I$ avec $a < b < c$, $p_b(a) \leq p_b(c)$.
4. Pour tout $a, b, c \in I$ avec $a < b < c$, $p_c(a) \leq p_c(b)$.

En résumé, f est convexe si et seulement si pour tout $x_0 \in I$ l'application p_{x_0} est croissante.

Démonstration.
Soit $x, y \in I$ avec $x < y$. Soit $\lambda \in]0,1[$ et $z \in]x, y[$.
Alors $z = \lambda x + (1-\lambda)y \Longleftrightarrow \lambda(x-y) = z-y \Longleftrightarrow \lambda = \dfrac{z-y}{x-y}$.

Ainsi les applications $\begin{array}{ccc}]0,1[& \longrightarrow &]x,y[\\ \lambda & \longmapsto & \lambda x+(1-\lambda)y \end{array}$ et $\begin{array}{ccc}]x,y[& \longrightarrow &]0,1[\\ z & \longmapsto & \dfrac{z-y}{x-y} \end{array}$, qui sont bien définies, sont des bijections réciproques l'une de l'autre.
Or f est convexe si et seulement si pour tout $x,y\in I$ avec $x<y$, pour tout $\lambda\in]0,1[$, $f(\lambda x+(1-\lambda)y)\leq \lambda f(x)+(1-\lambda)f(y)$, donc f est convexe si et seulement si pour tout $x,y\in I$ avec $x<y$, pour tout $z\in]x,y[$, (1) : $f(z)\leq \dfrac{z-y}{x-y}f(x)+\dfrac{x-z}{x-y}f(y)$.
Soit $x,y,z\in I$ tels que $x<z<y$:

$$\begin{aligned}\frac{f(x)-f(z)}{x-z}\leq\frac{f(y)-f(z)}{y-z} &\Longleftrightarrow f(z)\Big(\frac{1}{y-z}-\frac{1}{x-z}\Big)\leq -\frac{f(y)}{z-y}-\frac{f(x)}{x-z}\\ &\Longleftrightarrow f(z)\frac{x-y}{(x-z)(y-z)}\leq\frac{f(y)}{y-z}+\frac{f(x)}{z-x}\\ &\Longleftrightarrow (1),\end{aligned}$$

car $\dfrac{x-y}{(x-z)(y-z)}>0$. Ceci prouve que f est convexe si et seulement si pour tout $x,y,z\in I$ tels que $x<z<y$, $p_z(x)\leq p_z(y)$, donc on a prouvé "1." $\Longleftrightarrow$ "3.".
De même, si $x<z<y$,

$$\begin{aligned}p_x(z)\leq p_x(y) &\Longleftrightarrow \frac{f(z)-f(x)}{z-x}\leq\frac{f(y)-f(x)}{y-x}\\ &\Longleftrightarrow f(z)\frac{1}{z-x}\leq f(y)\frac{1}{y-x}+f(x)\Big(\frac{1}{z-x}-\frac{1}{y-x}\Big)\\ &\Longleftrightarrow f(z)\leq f(y)\frac{z-x}{y-x}+f(x)\frac{y-z}{y-x}\\ &\Longleftrightarrow (1),\end{aligned}$$

ce qui prouve que "1." $\Longleftrightarrow$ "2.".
La preuve de "1." $\Longleftrightarrow$ "4." est laissée en exercice. □

Remarque. Supposons que f est à la fois convexe et concave. Fixons $x,y\in I$ avec $x<y$. Alors d'après (1), pour tout $z\in]x,y[$, $f(z)=\dfrac{z-y}{x-y}f(x)+\dfrac{x-z}{x-y}f(y)$, donc $f(z)$ est un polynôme en z de degré inférieur à 1. On en déduit que f est deux fois dérivable sur I et que $f''=0$. Ainsi, il existe $a,b\in\mathbb{R}$ tel que $f=(x\longmapsto ax+b)$.

Propriété. (Hors programme) Si f est convexe sur I, elle est dérivable à droite et à gauche en tout point de $\overset{\circ}{I}$. En particulier, elle est continue sur $I\setminus\{Inf(I),Sup(I)\}$.

Démonstration.
Soit x_0 un point de $I\setminus\{Inf(I),Sup(I)\}$. p_{x_0} est une application monotone, donc elle admet une limite à droite et une limite à gauche en x_0. Ainsi, f est dérivable à droite et à gauche en x_0, donc f est continue en x_0, pour tout $x_0\in I\setminus\{Inf(I),Sup(I)\}$. □

8.5 Fonctions convexes dérivables

Propriété. On suppose que f est dérivable.
Alors f est convexe si et seulement si f' est croissante.

Démonstration.
Supposons d'abord que f' est croissante. Soit $(a,b,c) \in I^3$ tel que $a < b < c$.
$p_b(a) = \dfrac{f(b)-f(a)}{b-a} = f'(\alpha)$ où $\alpha \in]a,b[$, et $p_b(c) = f'(\beta)$ où $\beta \in]b,c[$.
f' est croissante, donc $p_b(a) \leq p_b(c)$. Ceci montre que f est convexe.

Réciproquement, supposons que f est convexe. Soit $(a,b) \in I^2$ avec $a < b$.
Soit $(x,y) \in I^2$ tel que $a < x < y < b$. On a $p_x(a) \leq p_x(y) = p_y(x) \leq p_y(b)$, donc $\dfrac{f(x)-f(a)}{x-a} \leq \dfrac{f(y)-f(b)}{y-b}$.
On fait tendre d'abord x vers a, puis y vers b et on obtient : $f'(a) \leq f'(b)$. □

Propriété. On suppose que f est dérivable. Alors f est convexe si et seulement si le graphe de f est situé au dessus de chacune de ses tangentes.

Démonstration.
Soit $x_0 \in I$. La tangente en x_0 au graphe de f a pour équation cartésienne :
$y = f(x_0) + f'(x_0)(x - x_0)$, donc le graphe de f est situé au dessus de chacune de ses tangentes si et seulement si pour tout $x, x_0 \in I$, $f(x) \geq f(x_0) + f'(x_0)(x - x_0)$, c'est-à-dire si et seulement si (T) : pour tout $x, x_0 \in I$, $\begin{cases} \text{pour } x < x_0, & p_{x_0}(x) \leq f'(x_0) \\ \text{pour } x > x_0, & p_{x_0}(x) \geq f'(x_0) \end{cases}$

◇ Supposons que f est convexe. Soit $x, x_0 \in I$.
Supposons que $x < x_0$. D'après l'égalité des accroissements finis, il existe $\alpha \in]x, x_0[$ tel que $p_{x_0}(x) = f'(\alpha)$, or f est dérivable et convexe, donc f' est croissante. De plus $\alpha < x_0$, donc $p_{x_0}(x) = f'(\alpha) \leq f'(x_0)$.
De même, on montre que si $x > x_0$, alors $p_{x_0}(x) \geq f'(x_0)$.
◇ Réciproquement, supposons (T). Soit $x, x_0, y \in I$ tels que $x < x_0 < y$. Alors $p_{x_0}(x) \leq f'(x_0) \leq p_{x_0}(y)$, donc d'après la définition "2.", f est convexe. □

Propriété. Si f est deux fois dérivable sur I,
$\boxed{f \text{ est convexe si et seulement si } \forall x \in I \ \ f''(x) \geq 0.}$

Exercice. Montrez que

$$\forall t \in [0, \frac{\pi}{2}] \ \ sin(t) \geq \frac{2t}{\pi}.$$

Résolution. L'application $\begin{array}{rccc} \sin : & [0, \frac{\pi}{2}] & \longrightarrow & \mathbb{R} \\ & t & \longmapsto & \sin t \end{array}$ est deux fois dérivable et $\forall t \in [0, \frac{\pi}{2}]$ $\sin''(t) = -\sin t \leq 0$. Ainsi sin est concave sur $[0, \frac{\pi}{2}]$, donc son graphe est au-dessus de la corde d'extrémités $(0, \sin 0) = 0$ et $(\frac{\pi}{2}, \sin\frac{\pi}{2}) = (\frac{\pi}{2}, 1)$, dont une équation cartésienne est $y = \frac{2}{\pi}x$. On en déduit l'inégalité ci-dessus.

Exercice. Montrer que, pour tout $x \in]-1, +\infty[$, $\ln(1+x) \leq x$.
Solution : Posons $f(x) = \ln(1+x)$. f est définie et de classe C^∞ sur $]-1, +\infty[$.
De plus $f''(x) = -\dfrac{1}{(1+x)^2} \leq 0$, donc f est concave. En particulier, le graphe de f est situé au dessous de sa tangente en 0, d'équation $y = x$.

Exercice. Montrer que pour tout $x \geq 0$, $\sin x \leq x$.

Propriété. On suppose que f est deux fois dérivable sur $\overset{\circ}{I}$ et que $x_0 \in \overset{\circ}{I}$.
Si f'' change de signe au voisinage de x_0, alors x_0 est un point d'inflexion de f.

Exemple. arctan admet un point d'inflexion en l'origine car $(\arctan)''(x) = \dfrac{-2x}{(1+x^2)^2}$.