2 Théorèmes fondamentaux du calcul différentiel

2.1 Théorème des accroissements finis

Pour une fonction d'*une variable réelle* et à *valeurs réelles*, on connaît la *formule des accroissements finis* :

Théorème 20 *Si $f : I \subset \mathbb{R} \to \mathbb{R}$ est dérivable sur l'intervalle I, alors pour tout $(x, y) \in I \times I$ (avec $x < y$), il existe $s \in]x, y[$ tel que*

$$f(y) - f(x) = f'(s)(y - x). \tag{14}$$

Par suite, si $|f'|$ est majorée par une constante $M > 0$ sur I, on a l'*inégalité des accroissements finis* :

$$|f(y) - f(x)| \leq M\,|y - x| \qquad \forall x, y \in I. \tag{15}$$

La *formule* des accroissements finis (14) ne subsiste pas si la fonction f n'est plus à valeurs réelles. Par exemple, la fonction $f : [0, 2\pi] \to \mathbb{R}^2$ définie par

$$f(x) = (\cos x, \sin x)$$

est dérivable, et sa dérivée $f' : x \mapsto (-\sin x, \cos x)$ ne s'annule pas, bien qu'il existe des points (nombreux) x et y distincts où $f(x) = f(y)$. Par contre, l'*inégalité* des accroissements finis (15) reste vraie comme on va le voir.

Définition 13 *Etant donné deux points x et y de $\mathbb{R}^n$, on appelle segment d'extrémités x et y (dans cet ordre), l'ensemble noté $[x, y]$:*

$$[x, y] := \{x + t(y - x) \in \mathbb{R}^n \;;\; t \in [0, 1]\} = \{(1 - t)x + ty \in \mathbb{R}^n \;;\; t \in [0, 1]\}.$$

Un sous-ensemble A de $\mathbb{R}^n$ est dit convexe si pour tout couple (x, y) de points de A, le segment $[x, y]$ est contenu dans A.

Théorème 21 (Inégalité des accroissements finis) *Soient $f : \Omega \subset \mathbb{R}^n \to \mathbb{R}^p$ une fonction différentiable, et a et b deux points de Ω tels que le segment $[a, b]$ soit contenu dans Ω. La fonction f vérifie*

$$\|f(b) - f(a)\| \leq \left(\sup_{z \in [a,b]} |||\mathrm{d}f_z|||_{\mathcal{L}(\mathbb{R}^n, \mathbb{R}^p)} \right) \|b - a\|, \tag{16}$$

où

$$|||\mathrm{d}f_z|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} = \sup_{h\in\mathbb{R}^n, h\neq 0} \frac{\|\mathrm{d}f_z(h)\|}{\|h\|}.$$

En particulier, si Ω est convexe et s'il existe $M \geq 0$ tel que

$$|||\mathrm{d}f_z|||_{\mathcal{L}(\mathbb{R}^n,\mathbb{R}^p)} \leq M \qquad \forall z \in \Omega,$$

alors, la fonction f est lipschitzienne :

$$\|f(x) - f(y)\| \leq M \|x - y\| \qquad \forall x, y \in \Omega.$$

2.2 Théorème d'inversion locale

Définition 14 (Difféomorphisme) *Soient U et V deux ouverts (non vides) de $\mathbb{R}^n$ et $k \in \mathbb{N}^* \cup \{\infty\}$. On dit qu'une fonction $f : U \to V$ est un difféomorphisme de classe C^k (ou un C^k-difféomorphisme) de U sur V si :*

i) *f est bijective de U sur V,*

ii) *f est de classe C^k sur U,*

iii) *f^{-1} est de classe C^k sur V.*

Exemple 26 La fonction trigonométrique tan est un C^1-difféomorphisme de $\left]-\frac{\pi}{2}, \frac{\pi}{2}\right[$ sur $\mathbb{R}$.

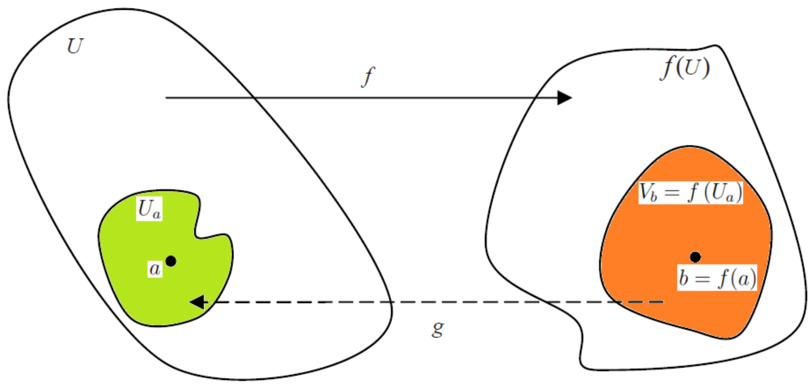
Par contre, la fonction polynômiale $x \mapsto x^3$ n'est pas un C^1-difféomorphisme de $\mathbb{R}$ dans $\mathbb{R}$, bien que ce soit une bijection de classe C^1. En effet, sa réciproque $f^{-1} : y \mapsto \sqrt[3]{y} = y^{1/3}$ n'est pas différentiable en $y = 0$ puisque

$$\frac{f^{-1}(y) - f^{-1}(0)}{y} = \frac{1}{y^{2/3}} \underset{y\to 0^{\pm}}{\longrightarrow} \pm\infty.$$

Etablir qu'une fonction f est un C^k-difféomorphisme à l'aide de la définition n'est pas simple en général, puisque cela nécessite d'avoir une expression de la réciproque de f. Celle-ci n'est pas toujours disponible. Le théorème suivant permet de savoir si une fonction est un C^k-difféomorphisme sans avoir besoin d'expliciter sa réciproque.

Théorème 22 (d'inversion locale) *Soient U un ouvert de $\mathbb{R}^n$ et $f : U \to \mathbb{R}^n$ une fonction de classe C^k, $k \in \mathbb{N}^* \cup \{\infty\}$. Si $a \in U$ est tel que $\det \mathbf{J}(f)_a \neq 0$, alors il existe un voisinage ouvert U_a de a contenu dans U et un voisinage ouvert V_b de $b = f(a)$ dans $\mathbb{R}^n$ tels que la* *restriction de f à U_a soit un C^k-difféomorphisme de U_a sur V_b. De plus, on a pour tout $y \in V_b$*

$$\mathbf{J}\left[\left(f_{|U_a}\right)^{-1}\right]_y = [\mathbf{J}(f)_x]^{-1} \qquad \text{où } x := \left(f_{|U_a}\right)^{-1}(y).$$



Exemple 27 (Coordonnées polaires) La fonction $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$\Phi : (r, \theta) \mapsto (x = r\cos\theta, y = r\sin\theta),$$

est de classe C^∞. Sa matrice jacobienne, au point (r, θ), a pour expression

$$\mathbf{J}(\Phi)_{(r,\theta)} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Son jacobien $\dfrac{D(x,y)}{D(r,\theta)} = \det \mathbf{J}(\Phi)_{(r,\theta)} = r$ est donc non nul en tout point $(r,\theta) \in \mathbb{R}^2$ vérifiant $r \neq 0$. D'après le théorème d'inversion locale, tout point $(r_0, \theta_0) \in \mathbb{R}^2$ vérifiant $r_0 \neq 0$ possède un voisinage ouvert $U_{(r_0,\theta_0)}$ tel que la restriction de Φ à $U_{(r_0,\theta_0)}$ soit un $C^\infty-$difféomorphisme de $U_{(r_0,\theta_0)}$ sur l'ouvert $\Phi\left(U_{(r_0,\theta_0)}\right)$ de $\mathbb{R}^2$.

Le théorème d'inversion locale permet de déterminer si une fonction est un difféomorphisme au voisinage d'un point donné. Il existe une version de ce théorème qui permet d'établir la propriété de C^k-difféomorphisme sur la totalité du domaine U. C'est le théorème d'inversion globale.

Théorème 23 (d'inversion globale) *Soient U un ouvert de $\mathbb{R}^n$, $k \in \mathbb{N}^* \cup \{\infty\}$ et $f : U \to \mathbb{R}^n$ une fonction de classe C^k. Si f est **injective**, alors les assertions suivantes sont équivalentes :*

i) *Le jacobien de f ne s'annule pas sur U ;*

ii) *$f(U)$ est un ouvert de $\mathbb{R}^n$ et f réalise un C^k-difféomorphisme de U sur $f(U)$. De plus, on a pour tout $x \in U$:*

$$[\mathbf{J}(f)_x]^{-1} = \mathbf{J}\left(f^{-1}\right)_{f(x)}.$$

L'intérêt du théorème d'inversion globale est qu'il dispense de déterminer la bijection réciproque f^{-1} de f et d'établir qu'elle est de classe C^k sur $f(U)$.

Exemple 28 (Coordonnées polaires) Reprenons l'exemple de la fonction $\Phi : \mathbb{R}^2 \to \mathbb{R}^2$ définie par

$$\Phi : (r, \theta) \mapsto (r\cos\theta, r\sin\theta).$$

Observons que pour tout $(r, \theta) \in \mathbb{R}^2$, on a $\Phi(r, \theta) = f(r, \theta + 2\pi)$ ce qui prouve que la fonction Φ n'est pas injective, ce n'est donc pas un C^1-difféomorphisme de sur $\mathbb{R}^2$ sur $\mathbb{R}^2$.
En remarquant que l'application Φ est injective sur l'ouvert $U =]0, +\infty[\, \times\,]0, 2\pi[$ et que sur cet ouvert, le jacobien de Φ ne s'annule pas, on établit à l'aide du théorème d'inversion globale que Φ est un $C^\infty-$difféomorphisme de U sur son image $\Phi(U) = \mathbb{R}^2 \setminus \{(x, 0) \in \mathbb{R}^2 \ ; \ x \geq 0\}$ de façon plus rapide qu'en ayant recours à la définition 14.

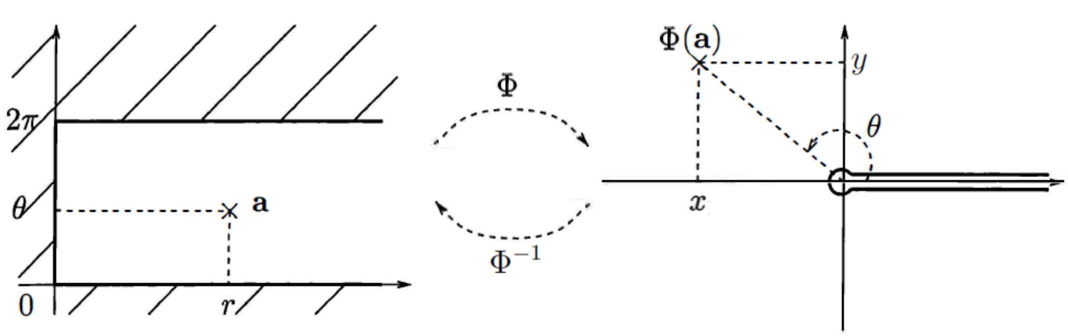


Illustration du changement de coordonnées en polaire

Remarque 14 D'un point de vue pratique, pour montrer qu'une fonction f définie sur un ouvert U de $\mathbb{R}^n$ et à valeurs dans un ouvert V de $\mathbb{R}^n$ est un C^k-difféomorphisme, on vérifie que :

- la fonction f est de classe C^k sur U,
- le jacobien de f ne s'annule pas sur U,
- la fonction f est injective sur U,
- l'image de U par f est l'ouvert V.

2.3 Théorème des fonctions implicites

2.3.1 Fonction de deux variables

Pour une fonction f de classe C^1 de deux variables, on considère l'équation $f(x, y) = 0$ et on cherche à comprendre si on peut tirer y comme fonction de x, i.e. si cette équation est en un certain sens équivalente à $y = \varphi(x)$ où φ est une fonction d'une variable. On dit alors que $f(x, y) = 0$ définit y comme fonction implicite de x. Le théorème des fonctions implicites donne un résultat général allant dans ce sens.
Ainsi, l'idée de ce paragraphe est que, si deux variables réelles x et y sont liées par une relation différentiable suffisamment régulière, alors dans une certaine mesure, l'une est fonction de l'autre, mais sans que l'on puisse donner nécessairement une expression explicite cette fonction.

Théorème 24 (Cas de deux variables) *Soient Ω un ouvert de $\mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction de classe C^1. Soit (a, b) un point de Ω tel que*

$$f(a, b) = 0 \quad \text{et} \quad \frac{\partial f}{\partial y}(a, b) \neq 0.$$

Alors, il existe un intervalle ouvert I contenant a, un intervalle ouvert J contenant b et une fonction $\varphi : I \to J$ de classe C^1 tels que :

i) $I \times J \subset \Omega$,

ii) *Pour $(x, y) \in I \times J$, on a l'équivalence :*

$$f(x, y) = 0 \iff y = \varphi(x).$$

Autrement dit, pour tout $x \in I$, $y = \varphi(x)$ est l'unique solution de $f(x, y) = 0$ dans l'intervalle J.

En particulier, on a $f(x, \varphi(x)) = 0$ pour tout $x \in I$.

De plus, si f est de classe C^k sur Ω alors φ est de classe C^k sur I.

On a un résultat particulièrement intéressant : bien que l'on ne connaisse pas φ, on est à même de calculer sa dérivée quand le théorème s'applique. Pour cela, il suffit de dériver l'identité $f(x, \varphi(x)) = 0$ dans I, ce qui donne

$$\frac{\partial f}{\partial x}(x, \varphi(x)) + \frac{\partial f}{\partial y}(x, \varphi(x))\varphi'(x) = 0, \quad \forall x \in I$$

ce qui permet de calculer $\varphi'(x)$ en tout point x de I où $\dfrac{\partial f}{\partial y} \neq 0$:

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))}.$$

Exemple 29 En considérant la fonction $f : (x, y) \mapsto x^2 + y^2 - 1$, on voit que $\frac{\partial f}{\partial y}(x, y) = 2y$ si bien que $\frac{\partial f}{\partial y}(1, 0) = \frac{\partial f}{\partial y}(-1, 0) = 0$, et donc le théorème ci-dessus est applicable en tout point $(x, y) \in \mathbb{R}^2$ où $x^2 + y^2 = 1$ sauf aux points $(1, 0)$ et $(-1, 0)$.

2.3.2 Fonction de trois variables

Le théorème des fonctions implicites se généralise aux fonctions de plus de deux variables.

Théorème 25 (Cas de trois variables) *Soient Ω un ouvert de $\mathbb{R}^3$ et $f : \Omega \to \mathbb{R}$ une fonction de classe C^1. Soit (a, b, c) un point de Ω tel que*

$$f(a, b, c) = 0 \quad \text{et} \quad \frac{\partial f}{\partial z}(a, b, c) \neq 0.$$

Alors, il existe un ouvert $U \subset \mathbb{R}^2$ contenant (a, b), un intervalle ouvert J contenant c et une fonction $\varphi : U \to J$ de classe C^1 tels que :

i) *$U \times J \subset \Omega$,*

ii) *Pour $(x, y, z) \in U \times J$, on a l'équivalence :*

$$f(x, y, z) = 0 \iff z = \varphi(x, y).$$

Autrement dit, pour tout $(x, y) \in U$, $z = \varphi(x, y)$ est l'unique solution de $f(x, y, z) = 0$ dans l'intervalle J.

En particulier, on a $f(x, y, \varphi(x, y)) = 0$ pour tout $(x, y) \in U$.

De plus, si f est de classe C^k sur Ω alors φ est de classe C^k sur U.

Dans ce cas, on peut aussi calculer les dérivées partielles de *la fonction implicite* φ quand le théorème s'applique. Il suffit de dériver par rapport à x puis par rapport à y l'identité $f(x, y, \varphi(x, y)) = 0$ dans U, pour obtenir les relations

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y, \varphi(x, y)) + \frac{\partial f}{\partial z}(x, y, \varphi(x, y))\frac{\partial \varphi}{\partial x}(x, y) &= 0, \\ \frac{\partial f}{\partial y}(x, y, \varphi(x, y)) + \frac{\partial f}{\partial z}(x, y, \varphi(x, y))\frac{\partial \varphi}{\partial y}(x, y) &= 0, \end{aligned}$$

valables pour tout $(x,y) \in U$, d'où l'on tire

$$\frac{\partial \varphi}{\partial x}(x,y) = -\frac{\frac{\partial f}{\partial x}(x,y,\varphi(x,y))}{\frac{\partial f}{\partial z}(x,y,\varphi(x,y))},$$

$$\frac{\partial \varphi}{\partial y}(x,y) = -\frac{\frac{\partial f}{\partial y}(x,y,\varphi(x,y))}{\frac{\partial f}{\partial z}(x,y,\varphi(x,y))},$$

en tout point (x,y) de U où $\dfrac{\partial f}{\partial z}$ ne s'annule pas.

2.4 Formules de Taylor

2.4.1 Formules de Taylor d'ordre 1

A l'ordre 1, les formules de Taylor se réduisent au théorème de la valeur moyenne.

Théorème 26 *Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^1. Soient $a \in \Omega$ et $h \in \mathbb{R}^n$ avec $h \neq 0$ tels que le segment $[a, a+h]$ soit contenu dans Ω. Alors*

i) *Théorème de la valeur moyenne : Il existe un nombre réel $s \in]0,1[$, dépendant de a et h, tel que*

$$f(a+h) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a+sh)h_i = f(a) + h\,\nabla f(a+sh).$$

ii) *Théorème de la valeur moyenne intégrale :*

$$f(a+h) = f(a) + \sum_{i=1}^{n} \left(\int_0^1 \frac{\partial f}{\partial x_i}(a+th)dt \right) h_i = f(a) + \int_0^1 h\,\nabla f(a+th)dt.$$

2.4.2 Formules de Taylor de second ordre

Théorème 27 *Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^2. Soient $a \in \Omega$ et $h \in \mathbb{R}^n$ avec $h \neq 0$ tels que le segment $[a, a+h]$ soit contenu dans Ω. Alors*

i) *Formule de Taylor avec reste de Lagrange : Il existe un nombre réel $s \in]0,1[$, dépendant de a et h, tel que*

$$\begin{aligned} f(a+h) &= f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)h_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(a+sh)h_i h_j \\ &= f(a) + h\,\nabla f(a) + \frac{1}{2} h\,\mathbf{H}_f(a+sh)\,h^{\mathrm{t}}. \end{aligned}$$

où $\mathbf{H}_f(a+sh)$ est la matrice hessienne de f au point $a+sh$ et h^{t} est le vecteur transposé de $h = (h_1, ..., h_n)$.

ii) *Formule de Taylor avec reste intégral :*

$$\begin{aligned} f(a+h) &= f(a) + h\,\nabla f(a) + \int_0^1 (1-t)\left[h\,\mathbf{H}_f(a+th)\,h^{\mathrm{t}}\right] dt \\ &= f(a) + h\,\nabla f(a) + \sum_{i,j=1}^{n}\left(\int_0^1 (1-t)\frac{\partial^2 f}{\partial x_i \partial x_j}(a+th)dt\right) h_i h_j. \end{aligned}$$

iii) *Formule de Taylor avec reste de Young :*

$$\begin{aligned} f(a+h) &= f(a) + h\,\nabla f(a) + \frac{1}{2} h\,\mathbf{H}_f(a)\,h^{\mathrm{t}} + o\left(\|h\|^2\right) \\ &= f(a) + h\,\nabla f(a) + \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 f}{\partial x_i \partial x_j}(a) h_i h_j + o\left(\|h\|^2\right) \qquad \textit{lorsque } h \to 0. \end{aligned}$$

Exemple 30 Si $n = 2$ et f est de classe C^2 au voisinage de $a \in \Omega$, on obtient pour $h = (h_1, h_2)$:

$$\begin{aligned} f(a+h) &= f(a) + h_1 \frac{\partial f}{\partial x}(a) + h_2 \frac{\partial f}{\partial y}(a) \\ &\quad + \frac{1}{2}\left(h_1^2 \frac{\partial^2 f}{\partial x^2}(a) + 2h_1 h_2 \frac{\partial^2 f}{\partial x \partial y}(a) + h_2^2 \frac{\partial^2 f}{\partial y^2}(a)\right) + o\left(\|h\|^2\right). \end{aligned}$$

2.4.3 Formule de Taylor d'ordre 3

Théorème 28 (Formule de Taylor-Lagrange d'ordre 3) *Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^3. Soient $a \in \Omega$ et $h \in \mathbb{R}^n$ avec $h \neq 0$ tels que le segment $[a, a+h]$ soit contenu dans Ω. Alors,*

$$f(a+h) = f(a) + h\,\nabla f(a) + \frac{1}{2} h\,\mathbf{H}_f(a)\,h^{\mathrm{t}} + \frac{1}{3!}\sum_{i,j,l=1}^{n}\frac{\partial^3 f}{\partial x_i \partial x_j \partial x_l}(a+sh) h_i h_j h_l,$$

pour un certain $s \in]0,1[$, dépendant de a et h.

Appendice : Formules de Taylor pour des fonctions d'une variable réelle

Théorème A Soient I un intervalle non vide de $\mathbb{R}$, $F : I \to \mathbb{R}$ une fonction de classe C^k ($k \in \mathbb{N}^*$), et a un point intérieur de I. Alors,

(i) Formule de Taylor avec reste de Lagrange : Pour tout $h \in \mathbb{R}^*$ tel que $a+h \in I$, il existe un nombre réel $s \in]0,1[$ tel que

$$F(a+h) = F(a) + \sum_{j=1}^{k-1} \frac{h^j}{j!} F^{(j)}(a) + \frac{h^k}{k!} F^{(k)}(a+sh). \tag{RL}$$

(Notons que s dépend de a et de h).

(ii) Formule de Taylor avec reste intégral (de Laplace) : Pour tout $h \in \mathbb{R}^*$ tel que $a+h \in I$, on a

$$F(a+h) = F(a) + \sum_{j=1}^{k-1} \frac{h^j}{j!} F^{(j)}(a) + \frac{h^k}{(k-1)!} \int_0^1 (1-t)^{k-1} F^{(k)}(a+th)\, dt. \tag{RI}$$

Remarque A. Le reste intégrale admet une autre expression. Plus précisément, on a l'égalité

$$\frac{h^k}{(k-1)!} \int_0^1 (1-t)^{k-1} F^{(k)}(a+th)\, dt = \int_a^{a+h} \frac{(a+h-t)^{k-1}}{(k-1)!} F^{(k)}(t)\, dt$$

qui découle d'un changement de variable $t \mapsto a+th$.

Théorème B (Formule de Taylor-Young à l'ordre k) Soient I un intervalle non vide de $\mathbb{R}$, $F : I \to \mathbb{R}$ une fonction k fois dérivable ($k \in \mathbb{N}^*$) sur I, et a un point intérieur de I. Alors, il existe une fonction $\varepsilon : I \to \mathbb{R}$ vérifiant $\lim_{t \to a} \varepsilon(t) = 0$ telle que pour tout $t \in I$:

$$F(t) = F(a) + \sum_{j=1}^{k} \frac{(t-a)^j}{j!} F^{(j)}(a) + (t-a)^k \varepsilon(t),$$

ou de façon équivalente, par changement de variable de t en $a+h$,

$$F(a+h) = F(a) + \sum_{j=1}^{k} \frac{h^j}{j!} F^{(j)}(a) + h^k\, \tilde{\varepsilon}(h), \qquad \text{avec } \lim_{h \to 0} \tilde{\varepsilon}(h) = 0.$$

Remarque B. Dire que $\lim_{h \to 0} \tilde{\varepsilon}(h) = 0$ signifie que le terme complémentaire $h^k\, \tilde{\varepsilon}(h)$ est négligeable devant h^k quand h tend vers 0. On peut donc écrire la formule de Taylor-Young sous la forme :

$$F(a+h) = F(a) + \sum_{j=1}^{k} \frac{h^j}{j!} F^{(j)}(a) + o\left(h^k\right). \tag{RY}$$

Remarque C. Les hypothèses du théorème de Taylor-Young sont plus faibles que celles du théorème A, mais on n'a pas d'expression précise pour le reste (on sait seulement qu'il est négligeable devant h^k au voisinage de 0).