I. Calcul différentiel

Qu'ils soient issus des sciences sociales, économiques, physiques ou naturelles, la plupart des problèmes concrets font intervenir plusieurs variables. La question de l'extension du calcul différentiel à une variable à des fonctions de plusieurs variables s'impose donc tout naturellement. Il s'agit pour l'essentiel d'étendre à des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$, la notion de dérivée

$$f'(a) := \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(a+h) - f(a)}{h}$$

que nous connaissons depuis le lycée pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Si cette définition de la dérivée peut être conservée telle quelle pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}^p$, il n'en va plus de même lorsque la variable h n'est plus scalaire mais vectorielle (puisqu'une division par h est alors impossible). Cette difficulté se contourne aisément en observant que la dérivabilité de f en a équivaut à l'existence d'un développement limité d'ordre 1 de f en a, c'est-à-dire à l'existence d'une application linéaire $L : \mathbb{R} \to \mathbb{R}$, $h \mapsto \lambda h$ telle que

$$f(a+h) - f(a) = L(h) + o(h),$$

relation que nous pouvons très bien écrire

$$f(a+h) - f(a) = L(h) + o(|h|),$$

puisque le signe du reste n'a aucune importance. Sous cette forme, il nous est en effet loisible d'étendre la définition à des fonctions de $\mathbb{R}^n$ dans $\mathbb{R}^p$ en remplaçant la valeur absolue par une norme quelconque sur $\mathbb{R}^n$ (rappelons qu'en dimension finie, toutes les normes sont équivalentes) : nous parlerons alors de différentiabilité de f en a. L'idée fondamentale du calcul différentiel est donc l'approximation locale des variations d'une fonction par une application linéaire.

Les applications du calcul différentiel à la géométrie et à l'optimisation (c'est-à-dire à la résolution des problèmes d'extrema) feront l'objet des chapitres II et IV.

Les prérequis de ce cours seront le calcul différentiel à une variable, l'algèbre linéaire et la topologie de $\mathbb{R}^n$.

Sauf mention explicite du contraire, $\mathcal{U}$ désignera un ouvert de $E = \mathbb{R}^n$, a un point de $\mathcal{U}$ et f une application de $\mathcal{U}$ dans $F = \mathbb{R}^p$.

Définition. *On dit que l'application f est différentiable en a s'il existe une application linéaire $L \in \mathcal{L}(E, F)$ telle que*

$$f(a+h) - f(a) = L(h) + o(\|h\|_E),$$

où $o(\|h\|_E)$ désigne une fonction $r(h)$ telle que

$$\lim_{\substack{h \to 0_E \\ h \neq 0_E}} \frac{r(h)}{\|h\|_E} = 0_F,$$

c'est-à-dire de la forme $\|h\|_E \, \varepsilon(h)$, où $\lim\limits_{\substack{h \to 0_E \\ h \neq 0_E}} \varepsilon(h) = 0_F$.

Remarques. 1. En d'autres termes, on dit que f est différentiable en a s'il existe une application linéaire $L \in L(E, F)$ telle que

$$\lim_{\substack{h \to 0_E \\ h \neq 0_E}} \frac{f(a+h) - f(a) - L(h)}{\|h\|_E} = 0_F,$$

ce qui peut s'écrire

$$\lim_{\substack{\|h\|_E \to 0 \\ \|h\| \neq 0}} \frac{\|f(a+h) - f(a) - L(h)\|_F}{\|h\|_E} = 0.$$

2. Sachant qu'en dimension finie, toutes les normes sont équivalentes, il est inutile de préciser ici le choix des normes $\|.\|_E$ et $\|.\|_F$.

3. Si une telle application linéaire $L \in \mathcal{L}(E, F)$ existe, elle est unique (le prouver !). Cette application L est alors appelée *différentielle de f en a* et notée df_a ou $df(a)$. Elle vérifie la relation

$$f(a+h) = f(a) + df_a(h) + o(\|h\|_E).$$

Cette relation qui n'est autre que le *développement limité à l'ordre* 1 *de f en a* peut encore s'écrire

$$f(x) = f(a) + df_a(x-a) + o(\|x-a\|_E)$$

en posant $x = a + h$.

4. Pourquoi se restreindre à un ouvert $\mathcal{U}$ de $\mathbb{R}^n$? Simplement pour qu'étant donné un point quelconque a de $\mathcal{U}$, le point $a+h$ soit encore dans $\mathcal{U}$ pourvu que la variation h soit suffisamment petite en norme.

5. Nous n'aborderons pas ici le cas où E et F seraient des espaces de dimension infinie. Il conviendrait dans ce cas de préciser les normes choisies et d'imposer à l'application linéaire $L \in L(E, F)$ d'être continue (ce qu'elle est nécessairement dans le cas présent puisque nous sommes en dimension finie).

Exercice. 1. Étudier la différentiabilité d'une fonction $f : \mathcal{U} \to F$ qui est constante.

2. Même question pour une application $f : E \to F$ qui est linéaire.

Notation. Il est à noter que $df_a(h)$ est parfois noté $df(a).h$.

Interprétation

La différentielle de f en a est donc la meilleure approximation linéaire de $h \mapsto f(a+h) - f(a)$ lorsque $h \to 0_E$. L'application

$$T : E \to F, x \mapsto f(a) + df_a(x-a)$$

s'interprétera quant à elle comme la meilleure approximation affine de f lorsque $x \to a$. Nous l'appellerons *application affine tangente à f au point a.*

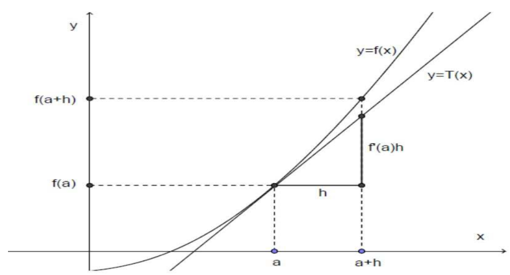
Cas des fonctions réelles d'une variable réelle

Nous avons défini la différentiabilité de f en a de manière à ce qu'elle soit équivalente à la dérivabilité de f en a lorsque f est une fonction de $\mathbb{R}$ dans $\mathbb{R}$. Si $f : \mathbb{R} \to \mathbb{R}$ est une fonction dérivable (et donc différentiable) en $a \in \mathbb{R}$, alors la différentielle de f en a est l'application linéaire

$$df_a : \mathbb{R} \to \mathbb{R},\ h \mapsto f'(a)\, h.$$

Ainsi :

$$\boxed{\begin{bmatrix} f \text{ dérivable en a} \\ \\ \text{et } f'(a) = m \end{bmatrix} \iff \begin{bmatrix} f \text{ différentiable en } a \text{ et } df_a \text{ représentée par la} \\ \\ 1 \times 1 \text{ matrice } (m) \text{ dans la base canonique de } \mathbb{R} \end{bmatrix}}.$$



Le graphe d'une fonction f d'une variable réelle $x \in \mathcal{U}$ est l'ensemble $\Gamma_f := \{(x, y) \in \mathcal{U} \times \mathbb{R} \,|\, y = f(x)\}$. Il s'agit d'une courbe de $\mathbb{R}^2$ si la fonction f est suffisamment sympathique (disons continue). Dire que la fonction f est différentiable (ou dérivable) en $a \in \mathcal{U}$ signifie géométriquement que son graphe Γ_f admet une droite non parallèle à l'axe des y pour tangente au point $(a, f(a))$.

Cas des fonctions réelles de deux variables réelles

Le graphe d'une fonction f de deux variables réelles $(x, y) \in \mathcal{U}$ est l'ensemble $\Gamma_f := \{(x, y, z) \in \mathcal{U} \times \mathbb{R} \,|\, z = f(x, y)\}$. Il s'agit d'une surface de $\mathbb{R}^3$ si la fonction f est suffisamment sympathique (nous ne définirons pas ici ce que nous entendons par « surface »). Dire que la fonction f est différentiable en $(a, b) \in \mathcal{U}$ signifie géométriquement qu'un plan non parallèle à l'axe des z est tangent à son graphe Γ_f au point $(a, b, f(a, b))$. Ce plan tangent aura pour équation $z = m(x - a) + n(y - b) + f(a, b)$ si $(m\ n)$ est la matrice de $df_{(a,b)} : \mathbb{R}^2 \to \mathbb{R}$ dans les bases canoniques de $\mathbb{R}^2$ et $\mathbb{R}$. Nous verrons un peu plus loin quels sont les coefficients m et n de cette matrice.

Différentiabilité et continuité

Proposition. *Si $f : \mathcal{U} \subset E \to F$ est différentiable en a, alors f est continue en a.*

©alcul différentiel par Y. Martinez-Maure

Exercice. Le démontrer.

La réciproque est naturellement fausse même pour des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ (une fonction $f : \mathbb{R} \to \mathbb{R}$ peut être continue sur $\mathbb{R}$ et dérivable en aucun point de $\mathbb{R}$: voir exercice d'approfondissement n°1).

Dérivée suivant un vecteur

Définition *Soit $v \in E$. On dit que la fonction f est dérivable en a suivant le vecteur v si l'application d'une variable réelle $t \mapsto f(a+tv)$ est dérivable en 0, c'est-à-dire si*

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a+tv) - f(a)}{t} \quad \textit{existe.}$$

Lorsqu'elle existe, cette limite est notée $(\partial_v f)(a)$ et appelée dérivée de f en a suivant le vecteur v (ou dérivée de f en a dans la direction v si $\|v\|_E = 1$).

Proposition. *Si f est différentiable en a, alors f est dérivable en a suivant tout vecteur $v \in E$ et*

$$\boxed{(\partial_v f)(a) = df_a(v)}.$$

Nous verrons plus bas que la dérivabilité suivant tout vecteur n'assure nullement la différentiabilité.

Dérivées partielles

Définition. *Notons $(a_1, \ldots, a_n)$ les coordonnées du point $a \in \mathcal{U}$ dans la base canonique de $\mathbb{R}^n$. Soit $i \in \{1, \ldots, n\}$. On dit que la fonction $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^p, (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n)$ est dérivable en a par rapport à sa $i^{\text{ème}}$ variable x_i si $x_i \mapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$ est dérivable en a_i, c'est-à-dire si la limite*

$$\lim_{\substack{x_i \to a_i \\ x_i \neq a_i}} \frac{f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n)}{x_i - a_i} \quad \textit{existe.}$$

Lorsqu'elle existe, cette limite est notée $\frac{\partial f}{\partial x_i}(a)$ ou $f'_{x_i}(a)$ ou encore $\partial_i f(a)$, et elle appelée $i^{\text{ème}}$ dérivée partielle de f en a.

Remarque. Les dérivées partielles de f (si elles existent) ne sont autres que ses dérivées suivant les vecteurs de la base canonique de $\mathbb{R}^n$. En effet, si $(e_1, \ldots, e_n)$ désigne cette base canonique, alors

$$\begin{aligned}(\partial_{e_i} f)(a) &= \lim_{\substack{t\to 0\\ t\neq 0}} \tfrac{f(a_1,\ldots,a_{i-1},a_i+t,a_{i+1},\ldots,a_n)-f(a_1,\ldots,a_n)}{t} \\ &= \lim_{\substack{x_i\to a_i\\ x_i\neq a_i}} \tfrac{f(a_1,\ldots,a_{i-1},x_i,a_{i+1},\ldots,a_n)-f(a_1,\ldots,a_n)}{x_i-a_i} \\ &= \tfrac{\partial f}{\partial x_i}(a)\,.\end{aligned}$$

La proposition précédente admet le corollaire suivant :

Corollaire. *Si $f : \mathcal{U} \subset \mathbb{R}^n \to F$ est différentiable en a, alors f est dérivable en a par rapport à chacune de ses variables, et :*

$$\boxed{\forall h = (h_1, \ldots, h_n) \in \mathbb{R}^n, \quad df_a(h) = \sum_{i=1}^{n} h_i \, \tfrac{\partial f}{\partial x_i}(a)}.$$

Démonstration. En vertu de la proposition précédente, la fonction f est dérivable en a suivant tout vecteur (et donc, en particulier, suivant ceux de la base canonique) et : $\forall h = \sum_{i=1}^{n} h_i e_i \in \mathbb{R}^n$,

$$df_a(h) = \sum_{i=1}^{n} h_i \, df_a(e_i) = \sum_{i=1}^{n} h_i \, (\partial_{e_i} f)(a) = \sum_{i=1}^{n} h_i \, \frac{\partial f}{\partial x_i}(a)\,.$$

□

Remarque. Attention, l'existence des n dérivées partielles $\frac{\partial f}{\partial x_i}(a)$ n'assure pas la différentiabilité de f en a, ni même sa continuité. Pire, la fonction f peut très bien être dérivable en a suivant tout vecteur v sans être continue en ce point comme le prouve l'exemple suivant.

Exercice. Soit f la fonction réelle définie sur $\mathbb{R}^2$ par

$$f(0,0) = 0 \quad \text{et} \quad f(x,y) = \frac{x^2 y}{x^4 + y^2} \quad \text{pour tout } (x,y) \neq (0,0)\,.$$

---

1. La fonction f est-elle dérivable en $(0,0)$ suivant tout vecteur $v=(x,y)\in\mathbb{R}^2$?

2. La fonction f est-elle différentiable (resp. continue) en $(0,0)$?

Critère pratique de différentiabilité - Fonctions de classe C^1

Théorème. *Si $f:\mathcal{U}\subset\mathbb{R}^n\to\mathbb{R}^p,(x_1,\ldots,x_n)\mapsto f(x_1,\ldots,x_n)$ est dérivable par rapport à chacune de ses variables au voisinage de a et si toutes ses dérivées partielles $\frac{\partial f}{\partial x_i}$ sont continues en a, alors f est différentiable en a.*

Définition. *La fonction $f:\mathcal{U}\subset\mathbb{R}^n\to\mathbb{R}^p$ est dite de classe C^1 sur $\mathcal{U}$ si elle est différentiable en chaque point de $\mathcal{U}$ et si sa différentielle, c'est-à-dire $df:\mathcal{U}\subset\mathbb{R}^n\to\mathcal{L}(E,F)$, $x\mapsto df_x$, est continue sur $\mathcal{U}$.*

Proposition. *$f:\mathcal{U}\subset\mathbb{R}^n\to\mathbb{R}^p$, $(x_1,\ldots,x_n)\mapsto f(x_1,\ldots,x_n)$ est de classe C^1 sur $\mathcal{U}$ si, et seulement si, ses n dérivées partielles $\frac{\partial f}{\partial x_i}$ sont définies et continues sur $\mathcal{U}$.*

Exercice. Démontrer que les applications polynomiales $P:\mathbb{R}^n\to\mathbb{R}$, $(x_1,\ldots,x_n)\mapsto P(x_1,\ldots,x_n)$ sont différentiables sur $\mathbb{R}^n$, c'est-à-dire en tout point de $\mathbb{R}^n$.

Notation. Désignons par $(dx_1,\ldots,dx_n)$ la base duale de la base canonique de $\mathbb{R}^n$. Ainsi :

$$\forall i\in\{1,\ldots,n\}, \forall h=(h_1,\ldots,h_n)\in\mathbb{R}^n,\quad dx_i(h)=h_i.$$

Si $f:\mathcal{U}\subset\mathbb{R}^n\to\mathbb{R}$ est une fonction différentiable en $a\in\mathcal{U}$, la différentielle de f en a peut ainsi s'écrire :

$$\boxed{df_a=\sum_{i=1}^{n}\frac{\partial f}{\partial x_i}(a)\ dx_i}.$$

Exercice. Démontrer que la fonction

$$f:\mathcal{U}=\mathbb{R}^*\times\mathbb{R},(x,y)\mapsto Arc\tan\frac{y}{x}$$

est différentiable sur $\mathcal{U}$ et donner l'expression de $df_{(x,y)}$ pour tout $(x,y)\in\mathcal{U}$.

Composantes - Matrice jacobienne

Soit $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^p$ une application. Pour tout $x \in \mathbb{R}^n$, notons $(f_1(x), \ldots, f_p(x))$ les coordonnées de $f(x)$ dans la base canonique de $\mathbb{R}^p$. Les applications $f_i : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ sont appelées *(applications) composantes* de f et on note $f = (f_1, \ldots, f_p)$.

Proposition. *L'application $f = (f_1, \ldots, f_p) : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^p$ est différentiable en a si, et seulement si, toutes ses composantes f_i le sont, et on a alors :*

$$\forall h \in \mathbb{R}^n, \quad df_a(h) = \left((df_1)_a(h), \ldots, (df_p)_a(h)\right).$$

Exercice. Le démontrer.

Exercice. Démontrer que $f : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (x + y, xy)$ est différentiable sur $\mathbb{R}^2$ et déterminer sa différentielle.

Reprenons les notations de la proposition précédente et supposons que f est différentiable en a. Dans les bases canoniques de $\mathbb{R}^n$ et $\mathbb{R}^p$, l'application $df_a \in \mathcal{L}(E, F)$ esr représentée par la matrice suivante :

$$\boxed{J_f(a) := \left(\frac{\partial f_i}{\partial x_j}(a)\right)_{\substack{1 \leq i \leq p \\ 1 \leq j \leq n}} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) \cdots \frac{\partial f_1}{\partial x_n}(a) \\ \vdots \quad \vdots \quad \vdots \\ \frac{\partial f_p}{\partial x_1}(a) \cdots \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix}}.$$

Noter que la $j^{\text{ème}}$ colonne de $J_f(a)$ s'obtient en dérivant les composantes de f par rapport à la $j^{\text{ème}}$ variable x_j.

Définition. *La matrice $J_f(a)$ est appelée matrice jacobienne de f en a.*

Exercice. Reprendre l'exercice précédent et retrouver l'expression de $df_{(x,y)}(h, k)$ en utilisant la matrice jacobienne de f en (x, y).

Exercice (coordonnées cylindriques). Calculer la matrice jacobienne de l'application

$$f:\mathbb{R}^3\to\mathbb{R}^3,\ (r,\theta,z)\mapsto(r\cos\theta,r\sin\theta,z)\,.$$

Linéarité - Composition

Proposition (Linéarité). *Si $f:\mathcal{U}\subset\mathbb{R}^n\to\mathbb{R}^p$ et $g:\mathcal{U}\subset\mathbb{R}^n\to\mathbb{R}^p$ sont différentiables en $a\in\mathcal{U}$, alors pour tous scalaires λ et μ, l'application $\lambda f+\mu g$ est différentiable en a et*

$$\boxed{d\left(\lambda f+\mu g\right)_a=\lambda df_a+\mu dg_a}.$$

Proposition (Composition). *Soient $\mathcal{U}$ un ouvert de $E=\mathbb{R}^n$, $\mathcal{V}$ un ouvert de $F=\mathbb{R}^p$ et $f:\mathcal{U}\to\mathcal{V}$ et $g:\mathcal{V}\to G=\mathbb{R}^q$ deux applications. Si f est différentiable en a et si g est différentiable en $f(a)$, alors $(g\circ f)$ est différentiable en a et on a :*

$$\boxed{d\left(g\circ f\right)_a=dg_{f(a)}\circ df_a},$$

soit matriciellement :

$$\boxed{J_{g\circ f}\left(a\right)=J_g\left(f\left(a\right)\right).J_f\left(a\right)}.$$

Remarque. Cette dernière relation généralise la formule donnant la dérivée d'une composée de fonctions dérivables (cas $n=p=q=1$) :

$$\left(g\circ f\right)\left(a\right)=\left(g'\circ f\right)\left(a\right)f'\left(a\right).$$

Exercice. Soient $]a,b[$ un intervalle ouvert de $\mathbb{R}$, $\gamma:\,]a,b[\,\to\mathbb{R}^n$, $t\mapsto\left(x_1\left(t\right),\ldots,x_n\left(t\right)\right)$ une fonction dérivable sur $]a,b[$ et $f:\mathbb{R}^n\to\mathbb{R}$ une fonction différentiable sur $\mathbb{R}^n$. Démontrer que $f\circ\gamma:\,]a,b[\,\to\mathbb{R}$, $t\mapsto f\left(x_1\left(t\right),\ldots,x_n\left(t\right)\right)$ est dérivable sur $]a,b[$ et donner l'expression de sa dérivée.

Exercice. Soit $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^p$ une fonction différentiable sur $\mathcal{U}$ et soient a, b deux points de $\mathcal{U}$ tels que

$$[a, b] =: \{ta + (1-t)\, b \,| t \in \mathbb{R}\} \subset \mathcal{U}.$$

Démontrer que l'application $g : [0,1] \to \mathbb{R}^p$, $t \mapsto f\,(ta + (1-t)\, b)$ est dérivable sur $[0,1]$ et donner l'expression de $g'\,(t)$ pour tout $t \in [0,1]$.

Gradient

On munit $\mathbb{R}^n$ du produit scalaire usuel $\langle ., . \rangle$. Si $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}$ est une fonction différentiable en $a \in \mathcal{U}$, alors il existe un unique vecteur $(\mathrm{grad}\, f)\,(a)$ tel que :

$$\boxed{\forall h \in \mathbb{R}^n, \quad df_a\,(h) = \langle \mathrm{grad}\, f\,(a)\,, h\rangle}.$$

Les coordonnées de $\mathrm{grad}\, f\,(a)$ dans la base canonique de $\mathbb{R}^n$ sont naturellement les dérivées partielles $\frac{\partial f}{\partial x_1}\,(a)\,, \ldots, \frac{\partial f}{\partial x_n}\,(a)$.

Définition. *Ce vecteur est appelé gradient de f en a et il est noté $\nabla f\,(a)$ (ce qui se lit « nabla de f en a »).*

Exercice. Reprenons les données de l'exercice précédent dans le cas $n = 2$ et supposons que :

(i) $\forall t \in \,]a, b[,\ \gamma'\,(t) \neq 0_{\mathbb{R}^2}$;
(ii) $\forall t \in \,]a, b[,\ \nabla f\,(\gamma\,(t)) \neq 0_{\mathbb{R}^2}$;
(iii) f est constante sur $\mathcal{C} = \gamma\,(]a, b[)$.

Nous admettrons provisoirement que, sous ces conditions, $\mathcal{C}$ est une courbe (dite *« de niveau »*) admettant en tout point $\gamma\,(t)$ une tangente dirigée par $\gamma'\,(t)$. Démontrer que le vecteur $\nabla f\,(\gamma\,(t))$ est normal à $\mathcal{C}$ en $\gamma\,(t)$, c'est-à-dire perpendiculaire à la tangente à $\mathcal{C}$ en $\gamma\,(t)$. Ainsi, le gradient de f est en tout point $\gamma\,(t)$, un vecteur normal à la ligne de niveau $\mathcal{C}$ en ce point.

Exercice. Soit $f : \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable sur $\mathbb{R}^2$. Démontrer que le gradient de f indique la direction de plus grande pente positive sur Γ_f à partir d'un point.

Différentielle seconde

Définition. *On dit que $f : \mathcal{U} \subset E \to F$ est deux fois différentiable en a si f est différentiable sur voisinage $\mathcal{V}$ de a et si sa différentielle $df : \mathcal{V} \subset E \to \mathcal{L}(E, F)$, $x \mapsto df_x$ est différentiable en a.*

Si c'est le cas, on note $d^2 f_a$ ou $d^2 f(a)$ la différentielle de df en a et on l'appelle la différentielle seconde (ou d'ordre 2) de f en a :

$$d^2 f_a := d(df)_a \,.$$

On dit que f est de classe C^2 sur $\mathcal{U}$ si f est deux fois différentiable sur $\mathcal{U}$ et si $x \mapsto d^2 f_x$ est continue.

Remarque et notation. L'application $d^2 f_a \in \mathcal{L}(E, \mathcal{L}(E, F))$ est linéaire et, pour tout $h \in E$, $d^2 f_a(h)$ est elle-même une application linéaire, par conséquent l'application

$$E \times E \to F, \ (h, k) \mapsto d^2 f_a(h) .k,$$

est bilinéaire. En fait, $\mathcal{L}(E, \mathcal{L}(E, F))$ s'identifie à l'espace $\mathcal{L}^2(E, F)$ des applications bilinéaires (nécessairement continues) de $E = \mathbb{R}^n$ dans $F = \mathbb{R}^p$ (muni de sa norme naturelle) via l'application

$$\begin{array}{c} \mathcal{L}^2(E, F) \to \mathcal{L}(E, \mathcal{L}(E, F)) \\ b \mapsto \begin{cases} E \to \mathcal{L}(E, F) \\ \ x \mapsto b(x, .) \end{cases} \end{array},$$

si bien que cette application bilinéaire sera encore notée $d^2 f_a$ et appelée différentielle seconde de f en a :

$$d^2 f_a(h, k) = d^2 f_a(h) .k.$$

Noter que les positions de h et k ont a priori de l'importance. Mais nous avons le résultat suivant.

Théorème. *Si $f : \mathcal{U} \subset E \to F$ est deux fois différentiable en $a \in \mathcal{U}$, alors on a :*

$$\forall (h, k) \in E^2, \quad d^2 f_a(h, k) = d^2 f_a(k, h) \,.$$

En d'autres termes, la différentielle seconde de f en a est bilinéaire symétrique.

Définition. *Si la dérivée partielle* $\frac{\partial f}{\partial x_j}$ *est définie au voisinage de* a *et admet une dérivée par rapport à sa* $i^{\text{ème}}$ *variable* x_i *en* a*, cette dernière est notée* $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ *:*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) := \frac{\partial}{\partial x_i}\left[\frac{\partial f}{\partial x_j}\right](a).$$

Les $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$, $(1 \leq i, j \leq n)$, *sont appelées dérivées partielles secondes (ou d'ordre* 2*) de* f *en* a*.*

Proposition. *Si* $f : \mathcal{U} \subset E \to F$ *est deux fois différentiable en* $a \in \mathcal{U}$*, alors les* n^2 *dérivées partielles secondes de* f *en* a *existent et on a :* $\forall h = (h_1, \ldots, h_n) \in E$, $\forall k = (k_1, \ldots, k_n) \in E$,

$$d^2 f_a(h,k) = \sum_{1 \leq i,j \leq n} h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(a).$$

Les deux résultats précédents prouvent que :

Corollaire (Théorème de Schwarz). *Si* $f : \mathcal{U} \subset E \to F$ *est deux fois différentiable en* $a \in \mathcal{U}$*, alors pour tout* $(i,j) \in \{1, \ldots, n\}^2$ *les dérivées partielles croisées* $\frac{\partial^2 f}{\partial x_i \partial x_j}(a)$ et $\frac{\partial^2 f}{\partial x_j \partial xi}(a)$ *existent et sont égales :*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial xi}(a).$$

Exercice. Soit f la fonction réelle définie sur $\mathbb{R}^2$ par :

$$f(0,0) = 0 \quad \text{et} \quad f(x,y) = \frac{x\,y^3}{x^2 + y^2} \quad \text{pour tout } (x,y) \neq 0.$$

a. Démontrer que f est de classe C^1 sur $\mathbb{R}^2$.

b. La fonction f est-elle de classe C^2 sur $\mathbb{R}^2 - \{(0,0)\}$?

c. Montrer que les dérivées partielles $\frac{\partial^2 f}{\partial x \partial y}(0,0)$ et $\frac{\partial^2 f}{\partial y \partial x}(0,0)$ existent et les comparer.

d. Que peut-on en déduire ?

Définition. *On dit que* $f : \mathcal{U} \subset E \to F$ *est de classe* C^2 *sur* $\mathcal{U}$ *si* f *est deux fois différentiable sur* $\mathcal{U}$ *et si* $d^2 f : \mathcal{U} \subset E \to \mathcal{L}^2(E,F)$, $x \mapsto d^2 f_x$ *est continue sur* $\mathcal{U}$*.*

Proposition. *L'application $f : \mathcal{U} \subset E \to F$ est de classe C^2 sur $\mathcal{U}$ si, et seulement si, ses n^2 dérivées partielles d'ordre 2 existent sont continues sur $\mathcal{U}$.*

Théorème (Développement de Taylor-Young de f à l'ordre 2).

Si $f : \mathcal{U} \subset E \to F$ est deux fois différentiable en $a \in \mathcal{U}$, alors :

$$\boxed{f(a+h) = f(a) + df_a(h) + \frac{1}{2} d^2 f_a(h,h) + o\left(\|h\|_E^2\right)}.$$

Remarque. Ce développement est également appelé *développement limité d'ordre 2 de f en a.*

Exercice. Que peut-on dire de la différentielle seconde d'une application linéaire $L : E \to F$ (resp. d'une application $f : E \to F$ de la forme $f(x) = b(x,x)$, où b est une application bilinéaire symétrique) ?

Parenthèse : application à l'optimisation.

Le calcul différentiel a d'importantes applications à l'optimisation (c'est-à-dire à l'étude des problèmes d'extrema) qui feront l'objet du chapitre 4. Dans cette parenthèse, nous allons déjà en présenter quelques éléments clefs. On se propose d'étudier les extrema d'une fonction $f : \mathcal{U} \subset E \to \mathbb{R}$ différentiable sur $\mathcal{U}$. Le résultat fondamental est le suivant :

Théorème. *Pour que f admette un extremum local en $a \in \mathcal{U}$, il est nécessaire (mais pas suffisant) que $df_a = 0$.*

Démonstration. Pour tout $x \in E$, il résulte des hypothèses que la fonction réelle $\varphi_x(t) := f(a+tx)$ est définie sur un voisinage de 0 dans $\mathbb{R}$ et qu'elle admet un extremum local en ce point. Or, cette fonction φ_x est dérivable en 0 et vérifie

$$\varphi_x'(0) := \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{\varphi_x(t) - \varphi_x(0)}{t} = \lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(a+tx) - f(a)}{t} = (\partial_x f)(a) = df_a(x).$$

Donc, on en déduit que $df_a(x) = 0$. □

Définition. *Un point $a \in \mathcal{U}$ tel que $df_a = 0$ est appelé un point critique de f.*

Remarques. 1. Un point critique de f ne correspond pas nécessairement à un extrema de la fonction f même dans le cas le plus simple des fonctions de $\mathbb{R}$ dans $\mathbb{R}$. Penser à $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$ qui vérifie $f'(0) = 0$, et donc $df_0 = 0$, mais ne présente pas d'extremum local en 0.

2. Le résultat serait erroné si nous ne supposions pas que $\mathcal{U}$ est un ouvert. Penser à $f : [0,1] \to \mathbb{R}$, $x \mapsto x$ qui admet un minimum en 0 et un maximum en 1 sans présenter le moindre point critique.

Interprétation géométrique

Comme $f : \mathcal{U} \subset E \to \mathbb{R}$ est différentiable sur $\mathcal{U}$, son graphe

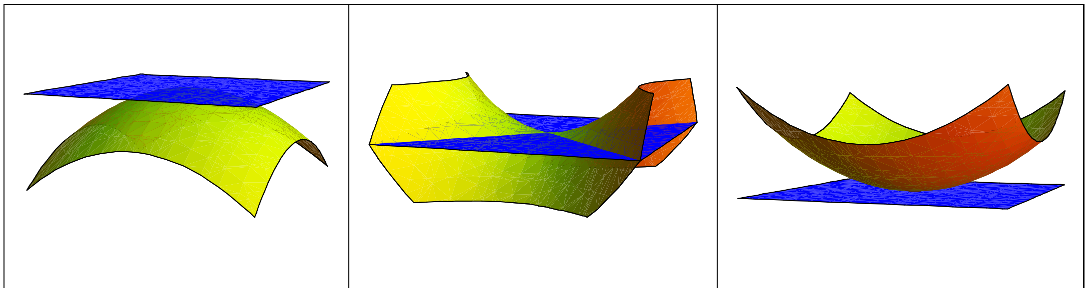
$$\Gamma_f := \{(x_1, \ldots, x_{n+1}) \in \mathcal{U} \times \mathbb{R} \,|\, x_{n+1} = f(x_1, \ldots, x_n)\}$$

est une « hypersurface différentiable » de $\mathbb{R}^{n+1}$ (une « courbe différentiable » si $n = 1$, une « surface différentiable » si $n = 2$, une « sous-variété différentiable de dimension p de $\mathbb{R}^{p+1}$» si $n = p$).

Dire que a est un point critique de la fonction f signifie qu'au point $(a, f(a)) \in \mathcal{U} \times \mathbb{R}$, l'hyperplan tangent au graphe de f est parallèle à l'hyperplan d'équation $x_{n+1} = 0$. D'un point de vue géométrique, on comprend bien qu'il est nécessaire que a soit un point critique de f pour que f admette un extremum local en a. Mais on voit également que cela n'est pas suffisant puisque le graphe de f peut très bien traverser son hyperplan tangent au point $(a, f(a))$.

Illustration dans le cas $n = 2$

La figure suivante représente l'allure du graphe de $f : \mathbb{R}^2 \to \mathbb{R}$ (i.e., de la surface d'équation $z = f(x, y)$) et sa position par rapport au plan d'équation $z = 0$ dans chacun des trois cas suivants : $f(x, y) = -(x^2 + y^2)$, $f(x, y) = x^2 - y^2$ et $f(x, y) = (x^2 + y^2)$. Sur cette figure, l'axe (Oz) qui n'a pas été représenté pointe vers le haut.



"Maximum local"	"Point col"	"Minimum local"

Pour une fonction $f : \mathcal{U} \subset \mathbb{R}^2 \to \mathbb{R}$ deux fois différentiable sur $\mathcal{U}$, le développement limité d'ordre 2 de f en un point $(a, b) \in \mathcal{U}$ prend la forme suivante

$$f(a+h, b+k) = f(a, b) + ph + qk + \frac{1}{2}\left(rh^2 + 2shk + tk^2\right) + o\left(h^2 + k^2\right),$$

où les nombres réels p, q (resp. r, s et t) ne sont autres que les dérivées partielles d'ordre 1 (resp. 2) de f en (a, b) :

$$p := \frac{\partial f}{\partial x}(a, b),\ q := \frac{\partial f}{\partial y}(a, b),\ r := \frac{\partial^2 f}{\partial x^2}(a, b),\ s := \frac{\partial^2 f}{\partial x \partial y}(a, b) \text{ et } t := \frac{\partial^2 f}{\partial^2 y^2}(a, b).$$

Si (a, b) est un point critique de f, c'est-à-dire si le plan tangent au graphe de f en $(a, b, f(a, b))$ est parallèle au plan xOy, alors il vient :

$$f(a+h, b+k) - f(a, b) = \frac{1}{2}\left(rh^2 + 2shk + tk^2\right) + o\left(h^2 + k^2\right).$$

Sous une hypothèse de "non-dégénérescence", cette relation nous permettra de ramener l'étude du signe de $f(a+h, b+k) - f(a, b)$ lorsque $\sqrt{h^2 + k^2} \to 0$ (et donc la position locale du graphe de f par rapport à son plan tangent en $(a, b, f(a, b))$) à l'étude de la forme quadratique

$$\left(H_{(a,b)} f\right)(h, k) := rh^2 + 2shk + tk^2 = (h, k)\begin{pmatrix} r & s \\ s & t \end{pmatrix}\begin{pmatrix} h \\ k \end{pmatrix},$$

appelée la *Hessienne de f en* (a, b).

Différentiabilité d'ordres supérieurs à 2

On procède ensuite par récurrence pour définir de manière analogue et pour tout $k \in \mathbb{N}^*$ les notions de :

• Fonction $f : \mathcal{U} \subset E \to F$ k fois différentiable en $a \in \mathcal{U}$ (resp. sur $\mathcal{U}$) ;

• Différentielle d'ordre k de f en a

$$d^k f_a \in \mathcal{L}^k (E, F),$$

où $\mathcal{L}^k (E, F)$ est l'espace des applications k-linéaires (nécessairement continues) de E dans F (muni de sa norme naturelle) ;

• Différentielle d'ordre k de f sur $\mathcal{U}$

$$\begin{aligned} d^k f : \mathcal{U} &\to \mathcal{L}^n (E, F) \\ x &\mapsto d^k f_x \ ; \end{aligned}$$

• Fonction $f : \mathcal{U} \subset E \to F$ de classe C^k sur $\mathcal{U}$;

• Dérivées partielles d'ordre k de f en a :

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} (a) := \frac{\partial}{\partial x_{i_1}} \left(\frac{\partial^{k-1} f}{\partial x_{i_2} \dots \partial x_{i_k}} \right) (a).$$

On retiendra les résultats suivants :

Théorème. *Si $f : \mathcal{U} \subset E \to F$ est k fois différentiable en a, alors l'application k-linéaire $d^k f_a$ est symétrique.*

Proposition. *Si $f : \mathcal{U} \subset E \to F$ est k fois différentiable en a, alors les n^k dérivées partielles d'ordre k de f en a existent et on a :*

$$d^k f_a (h_1, \dots, h_k) = \sum_{1 \leq i_1, \dots, i_k \leq n} h_{1,i_1} \dots h_{k,i_k} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} (a),$$

où, pour tout $i \in \{1, \dots, n\}$, $\left(h_{i,1}, \dots, h_{i,n}\right)$ sont les coordonnées de h_i dans la base canonique de $\mathbb{R}^n$.

Corollaire (Théorème de Schwarz). *Si $f : \mathcal{U} \subset E \to F$ est k fois différentiable en a, alors pour tout k-uplet $(i_1, \ldots, i_k) \in \{1, \ldots, n\}^k$ et toute permutation σ de $\{1, \ldots, k\}$, on a :*

$$\frac{\partial^k f}{\partial x_{i_{\sigma(1)}} \partial x_{i_{\sigma(2)}} \ldots \partial x_{i_{\sigma(k)}}}(a) = \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \ldots \partial x_{i_k}}(a).$$

En d'autres termes, les valeurs des dérivées partielles d'ordre k de f en a sont indépendantes de l'ordre dans lequel on effectue les dérivations successives.

Théorème (Développement de Taylor-Young de f à l'ordre k).

Si $f : \mathcal{U} \subset E \to F$ est k fois différentiable en $a \in \mathcal{U}$, alors :

$$\boxed{f(a+h) = f(a) + df(a).h + \tfrac{1}{2!} d^2 f(a).h^2 + \ldots + \tfrac{1}{k!} d^k f(a).h^k + o\left(\|h\|_E^k\right)},$$

où $d^i f(a).h^i := d^i f_a(h, \ldots, h)$, $(1 \leq i \leq k)$.

Remarque. 1. Ce développement est également appelé *développement limité d'ordre k de f en a.*

2. Noter qu'il s'agit là d'un **résultat local** : la seule chose que l'on sache sur le reste $o\left(\|h\|_E^k\right)$ est qu'il est négligeable devant $\|h\|_E^k$ lorsque $h \to 0_E$, c'est-à-dire de la forme $\|h\|_E^k \varepsilon(h)$, où $\lim\limits_{\substack{h \to 0_E \\ h \neq 0_E}} \varepsilon(h) = 0_F$.

Définition. *On dit que la fonction $f : \mathcal{U} \subset E \to F$ est de classe C^∞ sur $\mathcal{U}$ si elle est de classe C^n sur $\mathcal{U}$ pour tout $n \in \mathbb{N}$.*

Difféomorphismes

Les « difféomorphismes » sont les applications qui nous permettront d'effectuer des changements de variables (calculs d'intégrales multiples, résolution d'équations aux dérivées partielles (EDP), etc).

Définition. *Soient $\mathcal{U}$ et $\mathcal{V}$ deux ouverts de $\mathbb{R}^n$ et k un entier ≥ 1. Nous appellerons difféomorphisme de classe C^k de $\mathcal{U}$ sur $\mathcal{V}$ (ou C^k-difféomorphisme, ou simplement difféomorphisme si $k = 1$, de $\mathcal{U}$ sur $\mathcal{V}$) toute bijection $f : \mathcal{U} \to \mathcal{V}$ telle que $f : \mathcal{U} \to \mathcal{V} \subset \mathbb{R}^n$ et $f^{-1} : \mathcal{V} \to \mathcal{U} \subset \mathbb{R}^n$ sont de classe C^k.*

Exemple. a. Prouver que l'application

$$\begin{aligned} \varphi :]0,+\infty[\times]-\pi,\pi[&\to \mathbb{R}^2 - D \\ (r,\theta) &\longmapsto (r\cos\theta, r\sin\theta), \end{aligned}$$

où D est la demi-droite définie par $x \leq 0$ et $y = 0$, est un difféomorphisme de $]0,+\infty[\times]-\pi,\pi[$ sur $\mathbb{R}^2 - D$.

b. Étant donné f et g des fonctions de classe C^1 telles que $g(r,\theta) = f(r\cos\theta, r\sin\theta)$ pour tout $(r,\theta) \in]0,+\infty[\times]-\pi,\pi[$, donner les formules de passage entre les dérivées partielles de f et celles de g.

Propriété. Si f est un C^1- difféomorphisme de $\mathcal{U}$ sur $\mathcal{V}$ alors df_x est inversible pour tout $x \in \mathcal{U}$.

Démonstration. Par différentiation de la relation $f^{-1} \circ f = Id_{\mathcal{U}}$, il vient : $\forall x \in \mathcal{U}$,

$$d\left(f^{-1} \circ f\right)_x = d\left(Id_{\mathcal{U}}\right)_x$$

soit

$$\left(df^{-1}\right)_{f(x)} \circ df_x = Id_{\mathbb{R}^n},$$

autrement dit df_x est inversible et

$$\left(df_x\right)^{-1} = \left(df^{-1}\right)_{f(x)}.$$

□

Exercice. Exprimer matriciellement cette dernière relation.

Cette propriété admet une réciproque qui assure qu'une fonction $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^n$ de classe C^1 est inversible au voisinage de $a \in \mathcal{U}$ pourvu que df_a soit inversible.

Théorème d'inversion locale

Définition. *On dit que $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^n$ est un difféomorphisme local en $a \in \mathcal{U}$ s'il existe un voisinage ouvert $\mathcal{V}$ de a et un voisinage ouvert $\mathcal{W}$ de $f(a) \in F$ tel que $f : \mathcal{V} \to \mathcal{W}$ soit un difféomorphisme de $\mathcal{V}$ sur $\mathcal{W}$.*

Théorème d'inversion locale. *Soit $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction de classe C^1 sur $\mathcal{U}$ et soit $a \in \mathcal{U}$. Si df_a est inversible, alors f est un difféomorphisme local en a.*

La preuve du théorème d'inversion local repose sur le théorème du point fixe pour les applications contactantes dans un espace de Banach.

Exercice. a. Prouver que $f : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) \mapsto (x + y^4, y + x^3 y)$ est un difféomorphisme local en $(0, 0)$.

b. Que peut-on en déduire au sujet du système d'équations

$$\begin{cases} x + y^6 = a \\ \\ y + x^3 y = b \end{cases}$$

pour tout (a, b) assez proche de $(0, 0)$ dans $\mathbb{R}^2$?

Corollaire. *Soit $f : \mathcal{U} \subset \mathbb{R}^n \to \mathbb{R}^n$ une fonction de classe C^1 sur $\mathcal{U}$. Si f est injective et si df_x est inversible pour tout $x \in \mathcal{U}$, alors f est un difféomorphisme de $\mathcal{U}$ sur $f(\mathcal{U})$.*

Le théorème des fonctions implicites

Soit $f : \mathcal{U} \to \mathbb{R}$ une fonction réelle sur un ouvert $\mathcal{U}$ de $\mathbb{R}^n$. Lorsque l'ensemble

$$\Gamma := \{(x_1, \ldots, x_n) \in \mathcal{U} \,|\, f(x_1, \ldots, x_n) = 0\}$$

est le graphe d'une fonction $\varphi : \mathcal{V} \to \mathbb{R}$, $(x_1, \ldots, x_{n-1}) \mapsto \varphi(x_1, \ldots, x_{n-1})$ on dit que la relation $f(x_1, \ldots, x_n) = 0$ détermine implicitement x_n en fonction de $(x_1, \ldots, x_{n-1})$. On conçoit aisément qu'il est a priori plus commode d'envisager Γ comme le graphe d'une fonction φ (si toutefois cela est possible). Mais malheureusement une telle fonction φ n'existe pas toujours et lorsqu'elle existe, nous ne pouvons généralement pas la déterminer explicitement. Le problème en pratique est donc de décider de l'existence d'une telle fonction et, si elle existe, de pouvoir l'étudier sans la connaître explicitement. Cela sera l'objet du théorème suivant.

Théorème des fonctions implicites. *Soit $\mathcal{U}$ un ouvert de $\mathbb{R}^n$, soit $f : \mathcal{U} \to \mathbb{R}$ une fonction réelle de classe C^k sur $\mathcal{U}$, $(k \in \mathbb{N}^*)$, et soit $a = (a_1, \ldots, a_n)$ un point de $\mathcal{U}$ pour lequel*

$$f(a_1,\ldots,a_n)=0 \quad \text{et} \quad \frac{\partial f}{\partial x_n}(a_1,\ldots,a_n)\neq 0.$$

Il existe un voisinage ouvert $\mathcal{V}$ de $(a_1,\ldots,a_{n-1})$ dans $\mathbb{R}^{n-1}$ et un voisinage ouvert $\mathcal{W}$ de a_n dans $\mathbb{R}^n$ tels que l'ensemble

$$\{(x_1,\ldots,x_n)\in\mathcal{V}\times\mathcal{W}\,|\,f(x_1,\ldots,x_n)=0\}$$

soit le graphe d'une fonction réelle $\varphi:\mathcal{V}\to\mathcal{W}$ de classe C^k sur $\mathcal{U}$:

$$\forall(x_1,\ldots,x_n)\in\mathcal{V}\times\mathcal{W},\quad f(x_1,\ldots,x_n)=0\Leftrightarrow x_n=\varphi(x_1,\ldots,x_{n-1}).$$

Les dérivées partielles de φ vérifient :

$$\frac{\partial\varphi}{\partial x_i}(x_1,\ldots,x_{n-1})=-\frac{\frac{\partial f}{\partial x_i}(x_1,\ldots,x_{n-1},\varphi(x_1,\ldots,x_{n-1}))}{\frac{\partial f}{\partial x_n}(x_1,\ldots,x_{n-1},\varphi(x_1,\ldots,x_{n-1}))}$$

pour tout $(x_1,\ldots,x_{n-1})\in\mathcal{V}$.

Exercice. Démontrer cette dernière assertion.

Exercice. a. Montrer que la relation

$$\sin y+y+e^x=1$$

définit implicitement y en fonction de x au voisinage de $(0,0)$.

b. Prouver que la fonction ainsi définie est de classe C^∞ et préciser son développement limité d'ordre 3 en 0.

Exercice. *Soit $f:\mathcal{U}\to\mathbb{R}$, $(x,y,z)\mapsto f(x,y,z)$ une fonction réelle de classe C^1 sur un ouvert $\mathcal{U}$ de $\mathbb{R}^3$. On suppose que $df_{(x,y,z)}\neq 0$ pour tout $(x,y,z)\in\mathcal{U}$. Nous admettrons que sous ces hypothèses, l'ensemble $\mathcal{S}$ défini par l'équation $f(x,y,z)=0$ est une surface (cela résulte du théorème précédent qui permet d'affirmer qu'au voisinage de chacun de ses points, $\mathcal{S}$ peut se voir comme le graphe d'une fonction différentiable de deux variables réelles). Démontrer que pour tout $(x_0,y_0,z_0)\in\mathcal{U}$, le plan tangent à $\mathcal{S}$ au point (x_0,y_0,z_0) est défini par l'équation*

$$(x-x_0)\frac{\partial f}{\partial x}(x_0,y_0,z_0)+(y-y_0)\frac{\partial f}{\partial y}(x_0,y_0,z_0)+(z-z_0)\frac{\partial f}{\partial z}(x_0,y_0,z_0)=0.$$