

CALCUL DIFFERENTIEL

PLAN

I : Limites et Continuité

- 1) Fonctions à valeurs réelles
- 2) Limites et continuité

II : Dérivation

- 1) Dérivées partielles
- 2) Gradient
- 3) Dérivées de fonctions composées
- 4) Extremum
- 5) Dérivées successives

III : Courbes et surfaces

- 1) Tangente à une courbe
- 2) Plan tangent à une surface

Annexe : Théorème de Schwarz

I : Limites et continuité

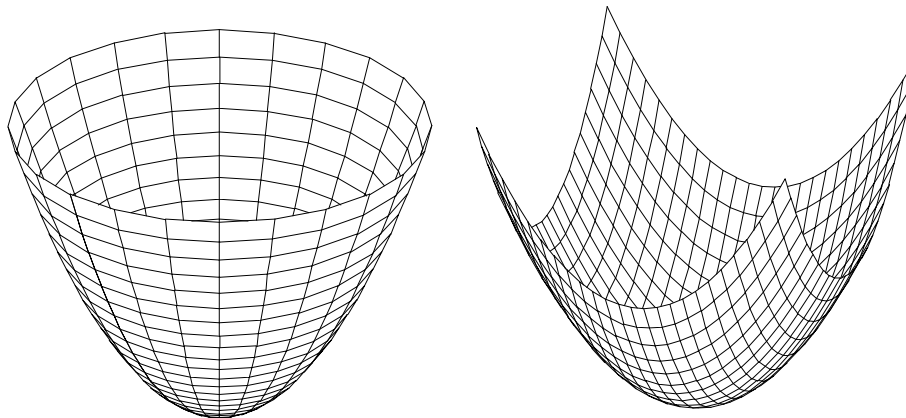
1- Fonctions à valeurs réelles

Dans ce chapitre, on considère des fonctions définies sur un ouvert U de \mathbb{R}^n , à valeurs réelles :

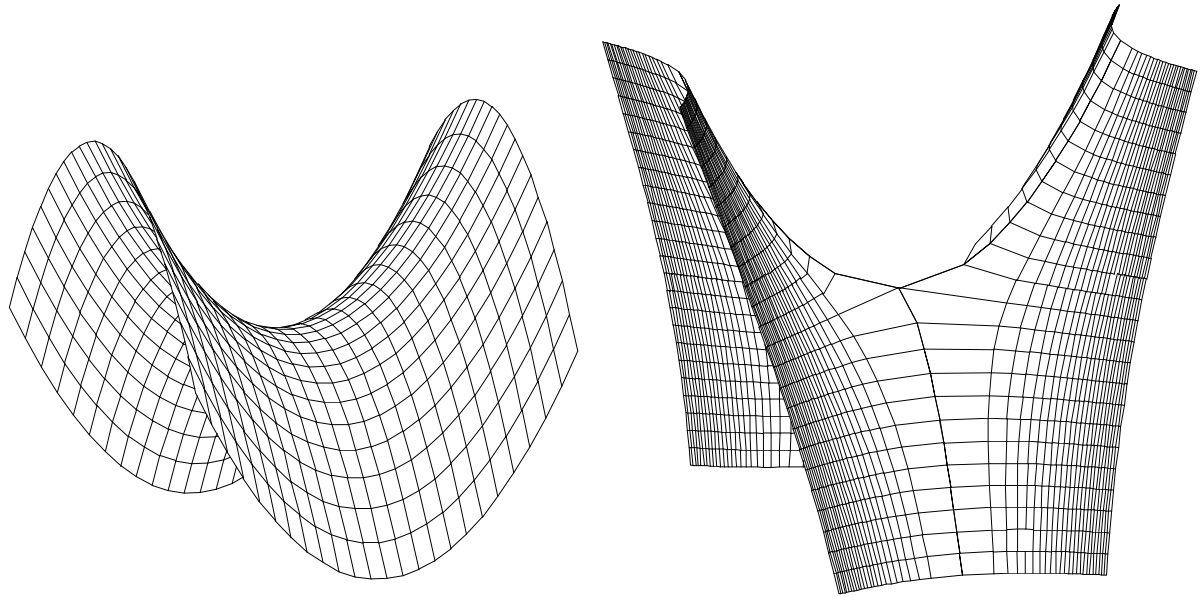
$$f : x = (x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n)$$

Dans le cas où $n = 2$, une telle fonction peut donner lieu à une représentation graphique dans \mathbb{R}^3 par la représentation de la surface $z = f(x, y)$.

EXEMPLE 1 : diverses représentations de la fonction $z = x^2 + y^2$



EXEMPLE 2 : représentations de $z = x^2 - y^2$



Les fonctions f de \mathbb{R}^2 dans \mathbb{R} peuvent également être représentées dans le plan par des lignes de niveaux $f(x, y) = \text{Cte}$. L'utilisation des lignes de niveau dans divers domaines est très fréquente. Citons, entre autres :

- isobares : lignes de même pression
- isobathes : lignes de même profondeur
- isoclines : lignes de même inclinaison magnétique
- isogones : lignes de même déclinaison magnétique
- isohyètes : lignes de même précipitation moyenne
- isohypses : ligne de même altitude
- isothermes : lignes de même température

2- Limites et continuité

Ces notions ont déjà été abordées dans le chapitre *Espaces normés*, mais la continuité a surtout été étudiée dans le cas des applications linéaires. Rappelons que f tend vers la limite l quand (le vecteur) x tend vers (le vecteur) x_0 si :

$$\forall \varepsilon > 0, \exists \alpha > 0, \forall x, \|x - x_0\| < \alpha \Rightarrow |f(x) - l| < \varepsilon$$

f est continue en x_0 si $l = f(x_0)$.

Il convient de noter qu'il ne suffit pas de vérifier la continuité des applications partielles définies ainsi :

$$x_i \rightarrow f(x_1, \dots, x_i, \dots, x_n)$$

pour conclure.

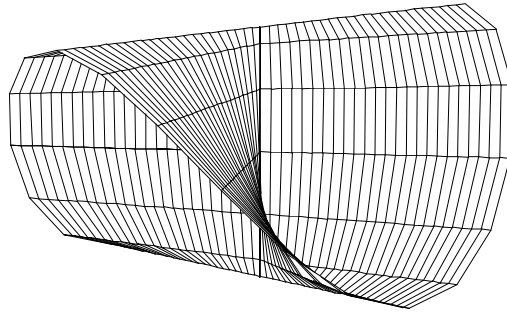
EXEMPLE :

Soit $f(x, y) = \frac{xy}{x^2 + y^2}$ pour $(x, y) \neq (0, 0)$. On note que $f(x, 0) = 0$ et que $f(x, x) = 1$ de sorte que, pour

$\varepsilon = \frac{1}{2}$, il n'existe aucun l permettant de vérifier :

$$\exists \alpha > 0, \forall a, \|a\| < \alpha \Rightarrow |f(a) - l| < \frac{1}{2}$$

puisqu'on doit avoir en même temps $|0 - l| < \frac{1}{2}$ et $|1 - l| < \frac{1}{2}$.



La fonction n'admet pas de limite en $(0, 0)$. Cependant, les deux applications partielles $x \rightarrow f(x, 0)$ et $y = f(0, y)$ admettent une limite en 0 puisqu'elles sont identiquement nulles.

Comment alors, procéder pour déterminer une limite ? Pratiquement, on se ramène au voisinage de 0 en effectuant un changement de variable $x = x_0 + a$. S'il existe une fonction g de \mathbb{R} dans \mathbb{R} telle que

$\lim_{r \rightarrow 0} g(r) = 0$ et $\forall a \in \mathbb{R}^n, |f(a)| \leq g(\|a\|)$, alors on peut conclure que $\lim_{a \rightarrow 0} f(a) = 0$. En effet :

$$\forall \varepsilon > 0, \exists \alpha > 0, \|a\| < \alpha \Rightarrow g(\|a\|) < \varepsilon \Rightarrow |f(a)| < \varepsilon$$

On peut utiliser une norme quelconque, en raison de l'équivalence des normes en dimension finie.

EXEMPLE 1 : $f(x, y) = \frac{xy^2}{x^2 + y^2}$. Si on note $r = \sqrt{x^2 + y^2}$, on a $|f(x, y)| \leq r$. La fonction admet donc une limite nulle en $(0,0)$. (On a ici $g(r) = r$)

En ce qui concerne la continuité, on utilise les trois résultats suivants :

- Les fonctions de projection $(x_1, \dots, x_i, \dots, x_n) \rightarrow x_i$ sont continues (utiliser directement la définition ou bien utiliser le fait qu'elles sont linéaires en dimension finie, et toute application linéaire en dimension finie est continue)
- La somme, produit, quotient, composée de fonctions continues est continue.
- Si la fonction est à valeurs dans \mathbb{R}^p , raisonner composante par composante

EXEMPLE :

La fonction $(x, y, z) \rightarrow \frac{x^2 - 3y}{e^z + 3xyz}$ est continue sur son ensemble de définition.

II : Dérivation

1- Dérivées partielles

Les fonctions qui suivent sont supposées définies sur un ouvert U de \mathbb{R}^n . Soit f une fonction définie de U dans \mathbb{R} , et soit $a = (a_1, \dots, a_n)$ un point de cet ouvert. On appelle dérivées partielles de f en a les dérivées des applications partielles suivantes :

$$x_1 \rightarrow f(x_1, a_2, \dots, a_n) \text{ dont la dérivée en } a_1 \text{ est notée } \partial_1 f(a) \text{ ou } \frac{\partial f}{\partial x_1}(a)$$

...

$x_i \rightarrow f(a_1, \dots, x_i, \dots, a_n)$ dont la dérivée en a_i est notée $\partial_i f(a)$ ou $\frac{\partial f}{\partial x_i}(a)$

...

$x_n \rightarrow f(a_1, \dots, a_{n-1}, x_n)$ dont la dérivée en a_{n-1} est notée $\partial_n f(a)$ ou $\frac{\partial f}{\partial x_n}(a)$

On dérive donc f par rapport à la $i^{\text{ème}}$ composante en considérant les autres composantes comme constantes. Si chaque dérivée partielle est continue, on dit que f est de classe C^1 .

EXEMPLE :

$$f(x, y) = 2x^3y^2 \Rightarrow \frac{\partial f}{\partial x}(x, y) = 6x^2y^2 \text{ et } \frac{\partial f}{\partial y}(x, y) = 4x^3y$$

L'intérêt des dérivées partielles est qu'elles permettent un développement limité de f au voisinage de chaque point. Ainsi, avec l'exemple ci-dessus :

$$\begin{aligned} f(x+h, y+k) &= 2(x+h)^3(y+k)^2 = 2(x^3 + 3x^2h + 3xh^2 + h^3)(y^2 + 2yk + k^2) \\ &= 2x^3y^2 + 6x^2y^2h + 4x^3yk + o(\|(h, k)\|) \\ &= f(x, y) + h \underbrace{\frac{\partial f}{\partial x}(x, y)} + k \underbrace{\frac{\partial f}{\partial y}(x, y)} + o(\|(h, k)\|) \end{aligned}$$

partie linéaire en (h, k) de la variation de f

PROPOSITION :

Soit f de classe C^1 sur un ouvert U de \mathbb{R}^n . Alors, pour tout élément a de U , et tout vecteur h tel que $a+h$ appartienne à U , on a :

$$f(a+h) = f(a) + \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(a) + o(\|h\|)$$

Cette expression s'appelle développement limité de f à l'ordre 1.

Démonstration hors programme, mais est-elle si difficile ?

On fera la démonstration dans le cas $n=2$ pour simplifier. On applique deux fois le théorème des accroissements finis pour les fonctions respectives $x \rightarrow f(x, a_2+h_2)$, et $y \rightarrow f(a_1, y)$:

$$\begin{aligned} f(a_1+h_1, a_2+h_2) - f(a_1, a_2) &= f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2) + f(a_1, a_2+h_2) - f(a_1, a_2) \\ &= h_1 \frac{\partial f}{\partial x_1}(a_1 + \theta_1 h_1, a_2+h_2) + h_2 \frac{\partial f}{\partial x_2}(a_1, a_2 + \theta_2 h_2) \text{ avec } 0 \leq \theta_1 \leq 1 \text{ et } 0 \leq \theta_2 \leq 1 \\ &= h_1 \left(\frac{\partial f}{\partial x_1}(a_1, a_2) + o(1) \right) + h_2 \left(\frac{\partial f}{\partial x_2}(a_1, a_2) + o(1) \right) \end{aligned}$$

où $o(1)$ désigne des fonctions de (h_1, h_2) qui tendent vers 0 lorsque (h_1, h_2) tend vers $(0, 0)$. On trouve bien l'expression annoncée.

L'application linéaire $h = (h_1, \dots, h_n) \rightarrow \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i$ s'appelle différentielle de f en a , notée $df(a)$. Par

ailleurs, les applications $h \rightarrow h_i$ sont notées dx_i , de sorte que :

$$df(a)(h) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i$$

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

$$df = \sum_{i=1}^n \frac{\partial f}{\partial x_i} dx_i$$

En physique, on note souvent de la même façon les fonctions et les valeurs qu'elles prennent ($E(x, y, z)$ est moins la fonction qui à (x, y, z) associe une énergie $E(x, y, z)$ que cette énergie elle-même). Alors que le mathématicien considère dx comme la fonction qui à (h_1, h_2, h_3) associe la variation h_1 de x , dx est considéré par le physicien comme la variation de x elle-même. De même, le mathématicien considère df comme l'application qui, à une variation de position h , associe la partie linéaire de la variation de f , alors que le physicien considère df comme cette variation elle-même, d'autant plus que h peuvent être choisis suffisamment petits pour rendre l'erreur $o(\|h\|)$ indécélable par les instruments de mesure.

EXEMPLE :

$$f(x, y) = x^y \text{ en } (1,2) \Rightarrow \frac{\partial f}{\partial x} = yx^{y-1} = 2 \text{ et } \frac{\partial f}{\partial y} = \ln(x) \times x^y = 0$$

$$df(1,2) = 2dx$$

ainsi, $1,02^{1,99} = 1,04019\dots$ alors que le calcul du développement limité au premier ordre donne :

$$1 + 2 \times 0,02 = 1,04.$$

En $(2,2)$, on a $df(2,2) = 4 dx + 4\ln(2) dy$

ainsi, $1,98^{2,01} = 3,94727\dots$ alors que le calcul du développement limité au premier ordre donne :

$$4 - 4 \times 0,02 + 4\ln(2) \times 0,01 = 3,9477\dots$$

2- Gradient

Dans \mathbb{R}^n muni de sa structure euclidienne, on peut voir la quantité $\sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i$ comme le produit

scalaire du vecteur h par le vecteur de composantes $\frac{\partial f}{\partial x_i}(a)$. Ce dernier vecteur s'appelle le gradient de

f en a , noté $\mathbf{grad}f(a)$ ou $\nabla f(a)$ (en physique). On a donc $df(a)(h) = \langle h, \mathbf{grad}f(a) \rangle$, ou encore :

$$f(a+h) = f(a) + \langle h, \mathbf{grad}f(a) \rangle + o(\|h\|)$$

On remarquera que, pour un déplacement h de longueur donnée, la variation (au premier ordre de f) $df(a)(h)$ est maximale lorsque h est colinéaire à $\mathbf{grad}(f)$, nulle si elle est orthogonale à $\mathbf{grad}(f)$. Cela s'interprète géométriquement par le fait que les surfaces de niveaux $f(x_1, \dots, x_n) = Cte$ sont orthogonales au gradient. La direction du gradient indique la direction suivant laquelle f varie le plus vite, la norme du gradient mesurant l'intensité de cette variation. Par exemple, si f est l'altitude en un point de latitude et longitude (x, y) , $\mathbf{grad}(f)$ est le vecteur orienté dans la direction de la ligne de plus grande pente, de norme égale à la pente locale. Au contraire, si on se déplace orthogonalement au gradient, f ne varie pas. On suit une ligne de niveau.

Cette interprétation est utilisée :

□ en mécanique : $f = -\mathbf{grad}(E)$ où E est l'énergie potentielle. f indique dans quel sens l'énergie potentielle décroît le plus vite. f est la force dérivant de l'énergie potentielle E .

Par exemple : la théorie Newtonienne de la gravitation considère que la Terre est soumise à une force centrale dirigée vers le Soleil de la forme $f = \frac{C}{r^2} \mathbf{u}$ où \mathbf{u} est un vecteur normé dirigé du Soleil vers la Terre, ($C < 0$) et $r = \sqrt{x^2 + y^2 + z^2}$. Soit $E = \frac{C}{r}$. On a :

$$\frac{\partial E}{\partial x} = -\frac{C}{r^2} \times \frac{\partial r}{\partial x} = -\frac{Cx}{r^3}$$

de même pour les autres dérivées. D'où $f = -\mathbf{grad}(E)$.

Autre exemple, au voisinage d'un point à la surface de la Terre, si $E = mgz$, alors $f = -mg\mathbf{k}$. E est l'énergie potentielle de pesanteur.

Dans le cas de la Terre dans son ensemble, il est défini en chaque point un champ \mathbf{g} de pesanteur dérivant d'un potentiel. La surface de niveau, perpendiculaire à tout point à ce champ \mathbf{g} de pesanteur, et correspondant au niveau moyen 0 de la mer, est appelé *géoïde*. On peut l'approximer par un ellipsoïde, dont le géoïde peut cependant différer en certains lieux par plusieurs centaines de mètres d'altitude. L'*International Association of Geodesy* a défini le *Geodetic Reference System* IAG GRS 1980 en convenant des longueurs des demi-axes de l'ellipsoïde, à savoir :

$a = 6\,378\,137\text{ m}$ à l'équateur

$b = 6\,356\,752,3141\text{ m}$ aux pôles

Ces longueurs diffèrent de modèles d'ellipsoïdes précédemment définis, par exemple celui de Hayford en 1909 pour lequel :

$a = 6\,378\,388\text{ m}$

$b = 6\,356\,911,9461\text{ m}$

et celui de Clarke en 1880, pour lequel :

$a = 6\,378\,249,2\text{ m}$

$b = 6\,356\,515,0\text{ m}$

Le système géodésique français NTF (*Nouvelle Triangulation Française*) utilisait encore récemment l'ellipsoïde de Clarke, mais le décret n°2000-1276 du 26 décembre 2000 définit un nouveau réseau, le RGF93 (*Réseau Géographique Français* de 1993), basé sur l'ellipsoïde IAG GRS 1980. Cet ellipsoïde est en effet utilisé au niveau mondial par le WGS84 (*World Geodesic System* de 1984) sur lequel est basé le positionnement GPS. Quant à l'Europe, elle utilisait encore l'ellipsoïde de Hayford pour sa base de données ED50 (*Europe Datum* de 1950) avant de s'aligner elle aussi sur le nouvel ellipsoïde dans le cadre de l'ETRS89 (*Europe Terrestrial Reference System* de 1989). L'utilisation d'un même ellipsoïde au niveau français, européen ou mondial est une nécessité afin d'éviter les différences de positionnement selon les systèmes utilisés. L'IGN (<http://www.ign.fr>) donne l'exemple suivant du positionnement d'un même point :

NTF	Ellipsoïde de Clarke	7°44'14.0"	48°36'00.0"
ED50	Ellipsoïde de Hayford	7°44'16.4"	48°36'03.0"
WGS84	Ellipsoïde IAG GRS 1980	7°44'12.2"	48°35'59.9"

Le décalage pourrait atteindre plusieurs centaines de mètres si on prenait les mêmes coordonnées dans les trois systèmes.

□ en électricité : $\mathbf{E} = -\mathbf{grad}(V)$ où V est le potentiel électrique. \mathbf{E} indique dans quel direction le potentiel décroît le plus vite. \mathbf{E} est le champ électrique. Cet exemple est très ressemblant au précédent, car une particule de charge q placée dans un champ électrique \mathbf{E} est soumise à une force $q\mathbf{E}$, qui dérive donc de l'énergie potentielle qV .

Dans un conducteur en équilibre, celui-ci se trouve à un potentiel constant. Le champ est nul à l'intérieur du conducteur, et le champ extérieur est orthogonal à la surface.

L'intérêt du potentiel est que sa connaissance suffit pour connaître le champ de vecteurs, et que les calculs éventuels sur des quantités scalaires sont plus faciles que les calculs sur des quantités vectorielles.

Donnons un dernier exemple : on considère la Terre comme un fluide en équilibre hydrostatique de masse volumique constante (hypothèses très réductrices !!). Dans ce cas :

$$\mathbf{grad} P = \mu \mathbf{g}$$

où P est la pression, μ la masse volumique et \mathbf{g} l'accélération de la pesanteur au point considéré. L'accélération de la pesanteur indique dans quel sens augmente la pression. Elle est orthogonale aux lignes isobares et la variation de pression est d'autant plus importante que μ est grand. Ainsi, au voisinage de la surface terrestre, on a, en fonction de la profondeur z :

$$\frac{dP}{dz} = \mu g \Rightarrow P = P_0 + \mu g z \text{ (avec } z \text{ orienté vers le bas)}$$

Plus généralement, si R est le rayon de la Terre (considérée comme sphérique ici), x la distance du point considéré au centre et g_0 l'accélération à la surface de la Terre, on a $g = \frac{g_0 x}{R}$. En effet,

l'accélération gravitationnelle vaut à la surface $g_0 = \frac{GM}{R^2}$ (avec G constante universelle de gravitation, M masse de la Terre), soit $\frac{4}{3}\pi R^3 \mu \times \frac{G}{R^2}$ ou encore $\frac{4\pi\mu GR}{3}$ alors qu'à la distance x , on a $g = \frac{4\pi\mu Gx}{3}$.

d'où :

$$\frac{dP}{dx} = -\frac{\mu g_0 x}{R} \Rightarrow P = P_0 - \mu g_0 \frac{x^2}{2R} + \mu g_0 \frac{R}{2}$$

Au centre de la Terre, $x = 0$ et $P = P_0 + \mu g_0 \frac{R}{2}$.

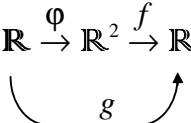
Application numérique : $R = 6370$ km, $M = 6 \cdot 10^{24}$ kg, $G = 6,67 \cdot 10^{-11}$ N.m².kg⁻²

P vaut environ $174 \cdot 10^9$ Pa. (La valeur trouvée dans ce modèle extrêmement simplifié est en fait deux fois plus petite que la valeur actuellement estimée dans des modèles plus élaborés, mais l'ordre de grandeur est bon).

Si on pose $x = R - z$, on obtient : $P_0 + \mu g_0(z - \frac{z^2}{2R})$ soit une erreur par rapport à $P_0 + \mu g z$ égale à $-\mu g_0 \frac{z^2}{2R}$.

3- Dérivées de fonctions composées

a) Considérons d'abord le cas suivant :

$$\mathbb{R} \xrightarrow{\varphi} \mathbb{R}^2 \xrightarrow{f} \mathbb{R}$$


$$t \rightarrow \varphi(t) \rightarrow f(\varphi(t)) = g(t)$$

On suppose que f est C^1 , de même que φ . On notera φ_1 et φ_2 les deux composantes de φ . Alors g est C^1 et :

$$g'(t) = \frac{\partial f}{\partial x_1}(\varphi(t)) \varphi_1'(t) + \frac{\partial f}{\partial x_2}(\varphi(t)) \varphi_2'(t)$$

En effet :

$$\begin{aligned}
 g(t+h) &= f(\varphi_1(t+h), \varphi_2(t+h)) \\
 &= f(\underbrace{\varphi_1(t) + h\varphi_1'(t) + o(h)}_{H_1}, \underbrace{\varphi_2(t) + h\varphi_2'(t) + o(h)}_{H_2}) \\
 &= f(x_1 + H_1, x_2 + H_2) \text{ en posant } x_1 = \varphi_1(t) \text{ et } x_2 = \varphi_2(t) \\
 &= f(x_1, x_2) + H_1 \frac{\partial f}{\partial x_1}(x_1, x_2) + H_2 \frac{\partial f}{\partial x_2}(x_1, x_2) + o(\|(H_1, H_2)\|) \\
 &= g(t) + h(\varphi_1'(t) \frac{\partial f}{\partial x_1}(x_1, x_2) + \varphi_2'(t) \frac{\partial f}{\partial x_2}(x_1, x_2)) + o(h)
 \end{aligned}$$

car $\|(H_1, H_2)\| = |h|O(1)$ où $O(1)$ est borné. Donc $o(\|(H_1, H_2)\|) = o(h)$

Donc g est dérivable en t de dérivée l'expression annoncée. On notera que, si la courbe $t \rightarrow \varphi(t)$ est incluse dans une courbe de niveau de f , alors $g = f \circ \varphi$ est constante, de sorte que la quantité

$$\frac{\partial f}{\partial x_1}(\varphi(t)) \varphi_1'(t) + \frac{\partial f}{\partial x_2}(\varphi(t)) \varphi_2'(t) \text{ est nulle. Cela signifie que le gradient de } f \text{ de composantes } \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \end{pmatrix}$$

est orthogonale à la tangente à la ligne de niveau, dirigée par le vecteur $\begin{pmatrix} \varphi_1' \\ \varphi_2' \end{pmatrix}$. Ce fait a déjà été signalé.

Le résultat précédent, ainsi que sa démonstration, se généralise (avec n indices au lieu de 2) sous la forme suivante :

$$\begin{array}{ccc}
 \mathbb{R} & \xrightarrow{\varphi} & \mathbb{R}^n \xrightarrow{f} \mathbb{R} \\
 & \searrow & \nearrow \\
 & & g
 \end{array}$$

$$t \rightarrow \varphi(t) \rightarrow f(\varphi(t)) = g(t)$$

On a alors :

$$g'(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\varphi(t)) \varphi_i'(t)$$

On peut aussi écrire cette dérivée sous la forme $g'(t) = \langle \mathbf{grad} f(\varphi(t)), \varphi'(t) \rangle$

b) Considérons ensuite le cas :

$$\begin{array}{ccc}
 \mathbb{R}^p & \xrightarrow{\varphi} & \mathbb{R}^n \xrightarrow{f} \mathbb{R} \\
 & \searrow & \nearrow \\
 & & g
 \end{array}$$

$$x \rightarrow y = \varphi(x) \rightarrow f(y) = f \circ \varphi(x) = g(x)$$

$$g(x) = g(x_1, \dots, x_p) = f(\varphi_1(x), \dots, \varphi_n(x))$$

PROPOSITION

Si f et φ sont C^1 , alors il en est de même de g , et, pour tout i variant de 1 à p :

$$\frac{\partial g(x)}{\partial x_i} = \sum_{k=1}^n \frac{\partial f}{\partial y_k}(\varphi(x)) \frac{\partial \varphi_k(x)}{\partial x_i}$$

En effet, $\frac{\partial g}{\partial x_i}$ est la dérivée de l'application partielle en x_i , et l'on applique le résultat du a) sur la fonction partielle $x_1 \rightarrow g(x_1, \dots, x_i, \dots, x_n)$

Les notations utilisées en physique sont peut-être plus faciles pour comprendre et mémoriser les formules. Le physicien en effet, ne donne pas de nom aux fonctions mais traite directement avec le nom des variables.

$$x = \begin{pmatrix} x_1 \\ \dots \\ x_p \end{pmatrix} \rightarrow y = \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} \rightarrow z$$

$$\frac{\partial z}{\partial x_i} = \frac{\partial z}{\partial y_1} \frac{\partial y_1}{\partial x_i} + \frac{\partial z}{\partial y_2} \frac{\partial y_2}{\partial x_i} + \dots + \frac{\partial z}{\partial y_n} \frac{\partial y_n}{\partial x_i}$$

Cette formule est la généralisation de la dérivation des fonctions composées de \mathbb{R} dans \mathbb{R} , que le physicien écrit sous la forme $\frac{dz}{dx} = \frac{dz}{dy} \frac{dy}{dx}$ (*Règle de la chaîne* ou *chain rule* en anglais).

EXEMPLE 1 :

Passage en polaire

$$(r, \theta) \rightarrow (x, y) \rightarrow f(x, y) = g(r, \theta)$$

$$x = r \cos(\theta)$$

$$y = r \sin(\theta)$$

$$\frac{\partial g}{\partial r} = \frac{\partial f}{\partial x} \cos(\theta) + \frac{\partial f}{\partial y} \sin(\theta)$$

$$\frac{1}{r} \frac{\partial g}{\partial \theta} = -\frac{\partial f}{\partial x} \sin(\theta) + \frac{\partial f}{\partial y} \cos(\theta)$$

Ce système permet d'en déduire inversement que :

$$\frac{\partial f}{\partial x} = \frac{\partial g}{\partial r} \cos(\theta) - \frac{1}{r} \frac{\partial g}{\partial \theta} \sin(\theta)$$

$$\frac{\partial f}{\partial y} = \frac{\partial g}{\partial r} \sin(\theta) + \frac{1}{r} \frac{\partial g}{\partial \theta} \cos(\theta)$$

On obtient ainsi l'expression du gradient en polaire :

$$\mathbf{grad} f = \frac{\partial g}{\partial r} \mathbf{e}_r + \frac{1}{r} \frac{\partial g}{\partial \theta} \mathbf{e}_\theta$$

On remarquera que, en physique, les fonctions f et g sont notées de la même façon, la différence des variables étant précisées par l'utilisation d'unités différentes (ex, mètres et mètres pour (x, y) ; mètres et radians pour (r, θ)). Ceci apparaît également dans la situation suivante : une quantité – par exemple l'énergie interne d'un gaz – peut s'exprimer comme fonction de diverses variables, (T, P) ou (T, V) . Dans le premier cas, les physiciens sont *obligés* de noter $(\frac{\partial U}{\partial T})_P$ et $(\frac{\partial U}{\partial T})_V$ les dérivées partielles de U par rapport à T respectivement lorsque les variables sont (T, P) et (T, V) . Le problème ne se

pose pas en Mathématique, car le mathématicien aurait noté de manière différente les fonctionnelles $U = f(T,P)$ et $U = g(T,V)$. Il écrirait alors simplement $\frac{\partial f}{\partial T}$ ou $\frac{\partial g}{\partial T}$ sans ambiguïté.

EXEMPLE 2 :

$$\mathbf{R} \xrightarrow{f} \mathbf{R}^4 \xrightarrow{g} \mathbf{G}$$

$$t \rightarrow \begin{pmatrix} x(t) \\ y(t) \\ z(t) \\ t \end{pmatrix} \rightarrow g(x,y,z,t)$$

La fonction composée est la fonction $t \rightarrow g(x(t), y(t), z(t), t)$. Sa dérivée est $\frac{\partial g}{\partial x} x' + \frac{\partial g}{\partial y} y' + \frac{\partial g}{\partial z} z' + \frac{\partial g}{\partial t}$. Ce cas se présente en physique lorsque l'on considère une particule se déplaçant dans un champ scalaire $E(x, y, z, t)$, où t est le temps. Cela signifie qu'en chaque point de l'espace est définie une fonction E dépendant éventuellement du temps. On souhaite connaître les variations de E que subit la particule, non seulement du fait que E dépend de t , mais aussi du fait que la particule se déplace. (On parle alors de dérivée particulaire). Le physicien aura tendance à noter $E(t)$ la fonction de la seule variable t , composée de E et f . On a alors :

$$\frac{dE}{dt} = \frac{\partial E}{\partial x} x' + \frac{\partial E}{\partial y} y' + \frac{\partial E}{\partial z} z' + \frac{\partial E}{\partial t} = \frac{\partial E}{\partial t} + \langle \mathbf{v}, \nabla E \rangle, \text{ où } \mathbf{v} \text{ est la vitesse de la particule, } \nabla E \text{ est}$$

l'opérateur qui à E associe le vecteur $\begin{pmatrix} \frac{\partial E}{\partial x} \\ \frac{\partial E}{\partial y} \\ \frac{\partial E}{\partial z} \end{pmatrix}$ autrement dit le gradient, et où \langle , \rangle est le produit

scalaire. $\frac{dE}{dt}$ est la dérivée de la fonction composée, alors que $\frac{\partial E}{\partial t}$ est la dérivée partielle de la fonction initiale.

Le cas d'un champ vectoriel \mathbf{E} est identique. Il suffit de raisonner composante par composante. On notera ici :

$$\frac{d\mathbf{E}}{dt} = \frac{\partial \mathbf{E}}{\partial t} + \langle \mathbf{v}, \nabla \rangle \mathbf{E}$$

où $\langle \mathbf{v}, \nabla \rangle$ est l'opérateur qui, au champ de composantes $\begin{pmatrix} E_x \\ E_y \\ E_z \end{pmatrix}$ associe le vecteur de composantes

$\begin{pmatrix} \langle \mathbf{v}, \nabla E_x \rangle \\ \langle \mathbf{v}, \nabla E_y \rangle \\ \langle \mathbf{v}, \nabla E_z \rangle \end{pmatrix}$. Un cas important se présente en cinématique des fluides : en tout point du fluide règne

un champ des vitesses \mathbf{V} de ce fluide, dépendant éventuellement du temps. A l'instant t , une particule se trouvant en (x, y, z) possède la vitesse $\mathbf{v} = \mathbf{V}(x, y, z, t)$. Son accélération est, d'après la formule précédente :

$$\mathbf{a} = \frac{\partial \mathbf{V}}{\partial t} + \langle \mathbf{V}, \nabla \rangle \mathbf{V}$$

qui a pour composantes $\left(\frac{\partial}{\partial t} + V_x \frac{\partial}{\partial x} + V_y \frac{\partial}{\partial y} + V_z \frac{\partial}{\partial z}\right) \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \begin{pmatrix} \frac{\partial V_x}{\partial t} + V_x \frac{\partial V_x}{\partial x} + V_y \frac{\partial V_x}{\partial y} + V_z \frac{\partial V_x}{\partial z} \\ \frac{\partial V_y}{\partial t} + V_x \frac{\partial V_y}{\partial x} + V_y \frac{\partial V_y}{\partial y} + V_z \frac{\partial V_y}{\partial z} \\ \frac{\partial V_z}{\partial t} + V_x \frac{\partial V_z}{\partial x} + V_y \frac{\partial V_z}{\partial y} + V_z \frac{\partial V_z}{\partial z} \end{pmatrix}$ On

pourra vérifier que cette expression est égale à :

$$\mathbf{a} = \frac{\partial \mathbf{V}}{\partial t} + \frac{1}{2} \mathbf{grad}(V^2) + \mathbf{Rot}(\mathbf{V}) \wedge \mathbf{V}$$

qui est l'autre expression possible de l'accélération de la particule, avec :

$$\frac{1}{2} \mathbf{grad}(V^2) = \begin{pmatrix} V_x \frac{\partial V_x}{\partial x} + V_y \frac{\partial V_y}{\partial x} + V_z \frac{\partial V_z}{\partial x} \\ V_x \frac{\partial V_x}{\partial y} + V_y \frac{\partial V_y}{\partial y} + V_z \frac{\partial V_z}{\partial y} \\ V_x \frac{\partial V_x}{\partial z} + V_y \frac{\partial V_y}{\partial z} + V_z \frac{\partial V_z}{\partial z} \end{pmatrix}$$

$$\mathbf{Rot}(\mathbf{V}) \wedge \mathbf{V} = \begin{pmatrix} \frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z} \\ \frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x} \\ \frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y} \end{pmatrix} \wedge \begin{pmatrix} V_x \\ V_y \\ V_z \end{pmatrix} = \begin{pmatrix} \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right)V_z - \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)V_y \\ \left(\frac{\partial V_y}{\partial x} - \frac{\partial V_x}{\partial y}\right)V_x - \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right)V_z \\ \left(\frac{\partial V_z}{\partial y} - \frac{\partial V_y}{\partial z}\right)V_y - \left(\frac{\partial V_x}{\partial z} - \frac{\partial V_z}{\partial x}\right)V_x \end{pmatrix}$$

Terminons par une propriété :

PROPOSITION

Soit f une fonction C^1 définie sur un ouvert convexe U de \mathbb{R}^n , à valeurs réelles. Alors f est constante si et seulement si son gradient est identiquement nul.

Démonstration :

U est dit convexe si : $\forall a \in U, \forall b \in U, \forall t \in [0, 1], ta + (1 - t)b \in U$. L'ensemble des $ta + (1 - t)b$ lorsque t décrit $[0, 1]$ est noté $[a, b]$. Des exemples de convexes sont donnés par un demi-plan, un disque, etc...

Si f est constante, toutes ses dérivées partielles sont nulles.

Réciproquement, si toutes les dérivées partielles sont nulles et si a et b appartiennent à U , alors la fonction $g = t \rightarrow f(ta + (1 - t)b)$ a pour dérivée $\langle a - b, \mathbf{grad} f(ta + (1 - t)b) \rangle$ donc est nul puisque le gradient est identiquement nul. g est donc constante, donc $f(a) = g(1) = g(0) = f(b)$, ce qui prouve que la valeur de f ne dépend pas du point choisi.

4- Extremum

Si $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ admet un extremum en $a = (a_1, \dots, a_n)$, alors pour tout i , la fonction partielle $x_i \rightarrow f(a_1, \dots, x_i, \dots, a_n)$ admet un extremum en a_i . Par conséquent, pour tout i , $\frac{\partial f}{\partial x_i}(a) = 0$, ce qu'on peut résumer en énonçant que $\mathbf{grad} f(a) = 0$. Un tel point est dit point critique.

Ce n'est qu'une condition nécessaire. Elle n'est pas suffisante (elle est fautive déjà pour les fonctions de \mathbb{R} dans \mathbb{R} . Exemple : x^3 en 0). Etudions de plus près ce qui se passe en dimension 2. Pour simplifier les notations, on suppose que $a = 0$ et que f admet un développement limité à l'ordre 2 au voisinage de $(0,0)$ sous la forme :

$$f(x, y) = f(0, 0) + \alpha x^2 + \beta xy + \gamma y^2 + o(r^2)$$

où $r^2 = x^2 + y^2$. Il n'y a pas de termes linéaires car $\frac{\partial f}{\partial x}(0,0)$ et $\frac{\partial f}{\partial y}(0,0)$ sont supposés être nuls.

- **Cas 1** : la quantité $\alpha x^2 + \beta xy + \gamma y^2$ strictement positive pour tout (x, y) non nul. Cette condition se produit si et seulement si le discriminant $\beta^2 - 4\alpha\gamma < 0$ et $\alpha > 0$, si et seulement si elle peut s'écrire comme somme de deux carrés. En effet, dans ce cas, le trinôme considéré comme fonction de x n'admet pas de racine sauf si $x = y = 0$. Si on pose $x = r\cos(\theta)$ et $y = r\sin(\theta)$, cette expression s'écrit $r^2(\alpha\cos^2(\theta) + \beta\cos(\theta)\sin(\theta) + \gamma\sin^2(\theta))$. Lorsque θ décrit $[0, 2\pi]$, la fonction $\theta \rightarrow \alpha\cos^2(\theta) + \beta\cos(\theta)\sin(\theta) + \gamma\sin^2(\theta)$ est continue, strictement positive, donc admet un minimum strictement positif m . On a donc :

$$f(x, y) \geq f(0,0) + r^2 m + o(r^2) = f(0,0) + r^2(m + o(1)) \geq f(0,0) \text{ pour } r \text{ assez petit}$$

Donc f admet un minimum local en $(0,0)$. De même si $\beta^2 - 4\alpha\gamma < 0$ et $\alpha < 0$, si et seulement si $\alpha x^2 + \beta xy + \gamma y^2$ s'écrit comme l'opposé d'une somme de deux carrés, f admet un maximum local en $(0,0)$

- **Cas 2** : l'expression $(x, y) \rightarrow \alpha x^2 + \beta xy + \gamma y^2$ change de signe, ce qui se produit si et seulement si son discriminant est strictement positif, si et seulement si elle peut s'écrire comme différence de deux carrés. Dans ce cas il n'y a pas d'extremum car si (x_1, y_1) est tel que $\alpha x_1^2 + \beta x_1 y_1 + \gamma y_1^2 > 0$, alors pour t assez petit, $f(tx_1, ty_1) = f(0,0) + t^2(\alpha x_1^2 + \beta x_1 y_1 + \gamma y_1^2) + o(t^2)$ est strictement supérieur à $f(0,0)$, et si (x_2, y_2) est tel que $\alpha x_2^2 + \beta x_2 y_2 + \gamma y_2^2 < 0$, alors pour t assez petit, $f(tx_2, ty_2) = f(0,0) + t^2(\alpha x_2^2 + \beta x_2 y_2 + \gamma y_2^2) + o(t^2)$ est strictement inférieur à $f(0,0)$
- **Cas 3** : l'expression est de signe constant mais sans être dans le cas 1. Cette condition se produit si et seulement si le discriminant $\beta^2 - 4\alpha\gamma = 0$, si et seulement si l'expression est un carré parfait ou l'opposé d'un carré parfait. Dans ce cas, on ne peut rien conclure.

EXEMPLE 1 :

Si $f(x, y) = x^2 + 3xy + y^2 + o(r^2) = (x + \frac{3y}{2})^2 - \frac{5y^2}{4} + o(r^2)$, il n'y a pas d'extremum.

Si $f(x, y) = x^2 + 3xy + 3y^2 + o(r^2) = (x + \frac{3y}{2})^2 + \frac{3y^2}{4} + o(r^2)$, il y a un minimum local en $(0,0)$

Si $f(x, y) = -x^2 - 3xy - 3y^2 + o(r^2) = -(x + \frac{3y}{2})^2 - \frac{3y^2}{4} + o(r^2)$, il y a un maximum local en $(0,0)$

Si $f(x, y) = y^2 + o(r^2)$, on ne peut rien dire.

EXEMPLE 2 :

Le dernier cas est illustré par l'exemple suivant :

$$f(x, y) = y^2 - 2yx^2 - y^3 + x^4 = y^2 + o(r^2)$$

Il est faux de dire que, si $y \neq 0$, alors $f(x, y) \sim y^2 \geq 0$, et que pour $y = 0$, $f(x, 0) = x^4 \geq 0$ et d'en conclure que f admet un minimum en $(0,0)$. En effet, la première équivalence est incorrecte : pour $y \neq 0$, $y^2 + o(r^2) = y^2 (1 + o(\frac{r^2}{y^2}))$, mais on ne peut absolument pas conclure que $o(\frac{r^2}{y^2})$ tend vers 0

quand y tend vers 0. D'ailleurs, pour tout $t \neq 0$:

$$f(t, t^2) = t^4 - 2t^4 - t^6 + t^4 = -t^6$$

et f prend des valeurs négatives dans tout voisinage de $(0,0)$. Il n'y a pas d'extremum.

5- Dérivées successives

Les dérivées partielles premières sont de nouvelles fonctions de \mathbb{R}^n dans \mathbb{R} . On peut évidemment itérer le procédé de dérivation, définir des dérivées partielles secondes, et parler de fonctions de classe C^2 , lorsque ces dérivées sont continues, et plus généralement de fonctions de classes C^k pour des fonctions dont les k premières dérivées partielles existent et sont continues.

On note $\frac{\partial^2 f}{\partial x_i^2}$ ou $\partial_i^2 f$ la dérivée seconde obtenue en dérivant deux fois de suite par rapport à x_i , $\frac{\partial^2 f}{\partial x_i \partial x_j}$ ou $\partial_i \partial_j f$ la dérivée seconde obtenue en dérivant d'abord par rapport à x_j puis à x_i , et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ ou $\partial_j \partial_i f$ la dérivée seconde obtenue en dérivant d'abord par rapport à x_i puis à x_j . Ces deux dernières dérivées peuvent être différentes. Cependant, la plupart du temps, elles sont égales.

THEOREME DE SCHWARZ

Soit f de classe C^2 sur un ouvert de \mathbb{R}^n . Alors, pour tout (i, j) $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$

Cette propriété est d'usage courant en physique. Une démonstration est donnée en annexe I. Plus généralement, pour une fonction de classe C^k , l'ordre de dérivation des k premières dérivées est sans importance.

III : Courbes et surfaces

1- Tangente à une courbe

Une courbe de \mathbb{R}^2 est le plus souvent donnée

- par un paramétrage $t \rightarrow \varphi(t)$:

Si $\varphi'(t)$ est non nul (point régulier), il s'agit du vecteur tangent en $\varphi(t)$ à la courbe.

- dans le plan, par une équation $f(x, y) = 0$:

Soit $m = (a, b)$ un point de la courbe. On a vu que le gradient $\begin{pmatrix} \frac{\partial f}{\partial x}(m) \\ \frac{\partial f}{\partial y}(m) \end{pmatrix}$ est orthogonal à la tangente

à la courbe en m . Si ce gradient est non nul, on dit que le point est régulier. L'équation de la tangente est alors :

$$\frac{\partial f}{\partial x}(m) (x - a) + \frac{\partial f}{\partial y}(m) (y - b) = 0$$

EXEMPLE 1 :

$$f(x, y) = x^2 + y^2 - 1 = 0$$

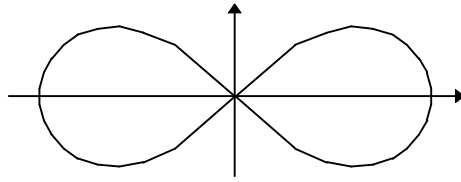
$\text{grad } f = 2 \begin{pmatrix} x \\ y \end{pmatrix}$ et l'on retrouve le résultat bien connu que le rayon d'un cercle est perpendiculaire à la tangente à ce cercle au point de contact.

EXEMPLE 2 :

$$f(x, y) = (x^2 + y^2)^2 - (x^2 - y^2) = 0$$

En polaire, on obtient :

$$r^2 = \cos(2\theta)$$



On constate qu'à l'origine, il y a deux tangentes. Ce point n'est pas régulier. De fait, le gradient s'y annule.

EXEMPLE 3 :

L'ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0$ admet pour tangente en (x_0, y_0) la droite d'équation :

$$(x - x_0) \frac{x_0}{a^2} + (y - y_0) \frac{y_0}{b^2} = 0$$

équation qu'on peut encore écrire $\frac{xx_0}{a^2} + \frac{yy_0}{b^2} - 1 = 0$.

L'équation de la tangente à une courbe de niveau peut se retrouver par le raisonnement suivant. Supposons qu'au voisinage de (x_0, y_0) , la relation $f(x, y) = 0$ permet de définir localement y comme fonction de x , soit $y = \Phi(x)$. On a donc identiquement $f(x, \Phi(x)) = 0$ sur un intervalle centré en x_0 . Cette relation permet de donner la valeur de $\Phi'(x)$ en fonction des dérivées partielles de f . On dérive par rapport à x la fonction composée :

$$x \rightarrow \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ \Phi(x) \end{pmatrix} \rightarrow f(x, y) = f(x, \Phi(x))$$

Cette fonction composée étant identiquement nulle, il en est de même de sa dérivée.

$$0 = \frac{\partial f}{\partial x}(x, y) + \Phi'(x) \frac{\partial f}{\partial y}(x, y)$$

$$\text{Donc : } \Phi'(x) = -\frac{\frac{\partial f}{\partial x}(x, y)}{\frac{\partial f}{\partial y}(x, y)} \text{ et en particulier pour } x = x_0 \text{ (et } y = y_0) : \Phi'(x_0) = -\frac{\frac{\partial f}{\partial x}(x_0, y_0)}{\frac{\partial f}{\partial y}(x_0, y_0)}$$

(On remarquera que Φ' s'annule là où $\frac{\partial f}{\partial x}$ s'annule). Le physicien notera plutôt cette relation sous la forme suivante :

$$f(x, y) = 0 \Rightarrow \frac{dy}{dx} = -\frac{\frac{\partial f}{\partial x}}{\frac{\partial f}{\partial y}}$$

2- Surfaces

Une surface dans \mathbb{R}^3 est le plus souvent donnée

- par un paramétrage :

$$(u, v) \rightarrow \begin{pmatrix} x(u, v) \\ y(u, v) \\ z(u, v) \end{pmatrix} = f(u, v)$$

où f est de classe C^1 sur un ouvert U de \mathbb{R}^2 .

Pour définir un arc paramétré au sein de la surface, il suffit de se donner deux fonctions $t \rightarrow u(t)$, $t \rightarrow v(t)$, indiquant comment les deux paramètres repérant le point de la surface varient en fonction du paramètre t .

Le vecteur tangent à cette courbe en un point considéré est donné par la dérivée de $t \rightarrow \begin{pmatrix} x(u(t),v(t)) \\ y(u(t),v(t)) \\ z(u(t),v(t)) \end{pmatrix}$

à savoir $u'(t) \frac{\partial f}{\partial u}(u,v) + v'(t) \frac{\partial f}{\partial v}(u,v)$ où $\frac{\partial f}{\partial u}(u,v)$ (respectivement $\frac{\partial f}{\partial v}(u,v)$) est le vecteur dont les composantes sont les dérivées partielles de x , y et z par rapport à u (respectivement v). En général, ces deux vecteurs sont linéairement indépendants. Si c'est le cas, les vecteurs tangents en un point M donné à tous les arcs paramétrés passant par $M = f(u, v)$ sont donc tous éléments du plan vectoriel engendré par les deux vecteurs $\frac{\partial f}{\partial u}(u,v)$ et $\frac{\partial f}{\partial v}(u,v)$ calculés en M. On définit le plan affine passant par

M et de direction le plan vectoriel précédent $\{f(u, v) + h \frac{\partial f}{\partial u}(u,v) + k \frac{\partial f}{\partial v}(u,v) \mid h \in \mathbb{R}, k \in \mathbb{R}\}$ comme étant le plan tangent en M à la surface. Il s'agit du plan approchant le mieux la surface au voisinage de $f(u,v)$. On reconnaît en effet dans l'expression $f(u, v) + h \frac{\partial f}{\partial u}(u,v) + k \frac{\partial f}{\partial v}(u,v)$ le développement limité de $f(u + h, v + k)$ sans son reste en $o(\|(h, k)\|)$. Le vecteur normal à ce plan est obtenu en faisant le produit vectoriel de $\frac{\partial f}{\partial u}(u,v)$ par $\frac{\partial f}{\partial v}(u,v)$.

- par une équation implicite $F(x, y, z) = 0$:

La surface s'appelle surface de niveau de la fonction F. Une courbe incluse dans cette surface est donnée par un paramétrage $t \rightarrow \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix}$ de façon que, pour tout t , $F(x(t), y(t), z(t)) = 0$. Si on dérive cette relation, on obtient :

$$x'(t) \frac{\partial F}{\partial x}(x(t), y(t), z(t)) + y'(t) \frac{\partial F}{\partial y}(x(t), y(t), z(t)) + z'(t) \frac{\partial F}{\partial z}(x(t), y(t), z(t)) = 0$$

ce qui exprime que $\begin{pmatrix} x'(t) \\ y'(t) \\ z'(t) \end{pmatrix}$, vecteur directeur tangente à la courbe si le point est régulier), est

orthogonal à $\begin{pmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial z} \end{pmatrix}$ au point considéré, c'est-à-dire au gradient. Comme pour les courbes, le

gradient est orthogonal aux surfaces de niveau.

Etablissons un lien entre les deux situations. Si on suppose que l'une des variables, z par exemple, est, au moins localement, fonction des autres, alors on a, dans un voisinage du point considéré :

$$F(x, y, z) = 0$$

$$\Leftrightarrow z = f(x,y) \text{ pour une certaine fonction } f$$

$$\Rightarrow \forall x,y, F(x,y,f(x,y)) = 0$$

$$\Rightarrow \frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \times \frac{\partial f}{\partial x} = 0 \text{ et } \frac{\partial F}{\partial y} + \frac{\partial F}{\partial z} \times \frac{\partial f}{\partial y} = 0$$

$$\Rightarrow \frac{\frac{\partial f}{\partial x}}{\frac{\partial F}{\partial z}} = -\frac{\frac{\partial x}{\partial F}}{\frac{\partial z}{\partial F}} \text{ et } \frac{\frac{\partial f}{\partial y}}{\frac{\partial F}{\partial z}} = -\frac{\frac{\partial y}{\partial F}}{\frac{\partial z}{\partial F}}$$

Il nous faut donc supposer que $\frac{\partial F}{\partial z}$ est non nulle au point considéré, et nous admettrons que, pour f de classe C^1 , cette condition est également suffisante pour l'existence locale de f .

Remarquons au passage que les physiciens noteront les relations précédentes sous la forme suivante, en utilisant x , y et z :

$$F(x, y, z) = 0 \Rightarrow \left(\frac{\partial z}{\partial x}\right)_y = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} \text{ et } \left(\frac{\partial z}{\partial y}\right)_x = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial z}}$$

Si la relation permet de définir $z = \Phi_z(x, y)$, $y = \Phi_y(x, z)$ et $x = \Phi_x(y, z)$, on aura :

$$\frac{\partial \Phi_x}{\partial y} \frac{\partial \Phi_y}{\partial z} \frac{\partial \Phi_z}{\partial x} = -\frac{\frac{\partial F}{\partial y}}{\frac{\partial F}{\partial x}} \times -\frac{\frac{\partial F}{\partial z}}{\frac{\partial F}{\partial y}} \times -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial z}} = -1$$

relation que les physiciens noteront plutôt sous la forme :

$$\left(\frac{\partial x}{\partial y}\right)_z \left(\frac{\partial y}{\partial z}\right)_x \left(\frac{\partial z}{\partial x}\right)_y = -1$$

On définit par exemple les coefficients calorimétriques suivants, en thermodynamique où la pression P , la température T et le volume V d'un gaz sont reliés par une relation $F(T, P, V) = 0$:

$$l = T \left(\frac{\partial P}{\partial T}\right)_V \text{ coefficient calorimétrique de dilatation}$$

$$h = -T \left(\frac{\partial V}{\partial T}\right)_P \text{ coefficient calorimétrique de compression}$$

$$\text{On a donc } l \left(\frac{\partial V}{\partial P}\right)_T = T \left(\frac{\partial P}{\partial T}\right)_V \left(\frac{\partial V}{\partial P}\right)_T = -T \left(\frac{\partial V}{\partial T}\right)_P = h$$

Revenons à notre surface, dont le paramétrage de la surface est $\begin{pmatrix} x \\ y \\ f(x,y) \end{pmatrix}$. Les vecteurs du plan

tangent au point (x,y,z) considéré sont $\begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x} \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y} \end{pmatrix}$. Un vecteur normal au plan tangent est donc

$\begin{pmatrix} -\frac{\partial f}{\partial x} \\ -\frac{\partial f}{\partial y} \\ 1 \end{pmatrix}$ ou encore par $\begin{pmatrix} \frac{\partial F}{\partial x} \\ \frac{\partial F}{\partial y} \\ \frac{\partial F}{\partial z} \end{pmatrix}$ On retrouve donc le fait que, pour une surface donnée sous la forme

implicite $F(x, y, z) = 0$, un vecteur normal à la surface est donné par le gradient de la fonction F .

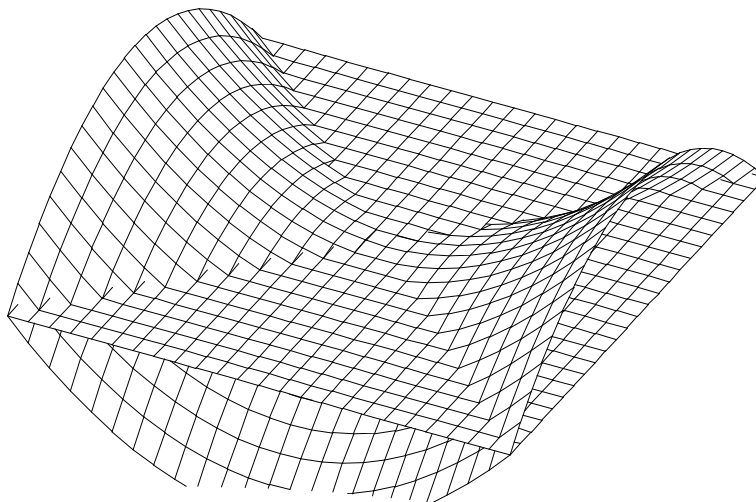
EXEMPLE : Soit la surface $z = x^2 - y^2$. Donner l'équation du plan tangent au point $(1, 2, -3)$.

- Si on considère le paramétrage $(x, y) \rightarrow \begin{pmatrix} x \\ y \\ x^2 - y^2 \end{pmatrix}$, alors le plan tangent passe par $(1, 2, -3)$ et est dirigé par les deux vecteurs dérivés $\begin{pmatrix} 1 \\ 0 \\ 2x \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ -2y \end{pmatrix}$, calculés en ce point, à savoir $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$ et $\begin{pmatrix} 0 \\ 1 \\ -4 \end{pmatrix}$.

Un vecteur normal est donné par le produit vectoriel de ces deux vecteurs, ce qui donne $\begin{pmatrix} -2 \\ 4 \\ 1 \end{pmatrix}$.

- On peut aussi paramétriser (mais quelle idée !!) la partie de la surface contenue dans le demi-espace $z \geq 0$ par le paramétrage $(r, t) \rightarrow \begin{pmatrix} r \operatorname{ch}(t) \\ r \operatorname{sh}(t) \\ r^2 \end{pmatrix}$. Les vecteurs dérivés sont proportionnels cette fois $\begin{pmatrix} \operatorname{ch}(t) \\ \operatorname{sh}(t) \\ 2r \end{pmatrix}$ et $\begin{pmatrix} \operatorname{sh}(t) \\ \operatorname{ch}(t) \\ 0 \end{pmatrix}$, dont le produit vectoriel est égal à $\begin{pmatrix} -2r \operatorname{ch}(t) \\ 2r \operatorname{sh}(t) \\ 1 \end{pmatrix}$ qui vaut $\begin{pmatrix} -2 \\ 4 \\ 1 \end{pmatrix}$ au point considéré, comme ci-dessus.
- On peut aussi considérer la surface de niveau $z - x^2 + y^2 = 0$. Le gradient, orthogonal à cette surface est égal à $\begin{pmatrix} -2x \\ 2y \\ 1 \end{pmatrix}$ qui vaut $\begin{pmatrix} -2 \\ 4 \\ 1 \end{pmatrix}$ au point considéré, comme ci-dessus.

Une équation du plan tangent est donc $-2x + 4y + z = 3$. Dans le cas présent, le plan tangent coupe la surface.



Annexe : Théorème de Schwarz

Dans cette annexe, on démontre le théorème de Schwarz dans le cas où $n = 2$. Soit f une fonction de classe C^2 sur un ouvert de \mathbb{R}^2 . Nous voulons montrer que $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$. Nous considérons la fonction suivante :

$$\Delta(h_1, h_2) = f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) - f(x_1, x_2 + h_2) + f(x_1, x_2)$$

i) Considérons $\Phi_1(t) = f(t, x_2 + h_2) - f(t, x_2)$. On a, en appliquant le théorème des accroissements finis à Φ_1 :

$$\begin{aligned} \Delta(h_1, h_2) &= \Phi_1(x_1 + h_1) - \Phi_1(x_1) = h_1 \Phi_1'(x_1 + \theta h_1) \text{ avec } 0 < \theta < 1 \\ \Rightarrow \Delta(h_1, h_2) &= h_1 \left(\frac{\partial f}{\partial x_1}(x_1 + \theta h_1, x_2 + h_2) - \frac{\partial f}{\partial x_1}(x_1 + \theta h_1, x_2) \right) \\ &= h_1 h_2 \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1 + \theta h_1, x_2 + \theta' h_2) \text{ avec } 0 < \theta' < 1 \end{aligned}$$

en appliquant le théorème des accroissements finis sur la fonction $x_2 \rightarrow \frac{\partial f}{\partial x_1}(x_1 + \theta h_1, x_2)$. On en

déduit que $\lim_{(h_1, h_2) \rightarrow (0, 0)} \frac{\Delta(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_1, x_2)$, car $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ est continue.

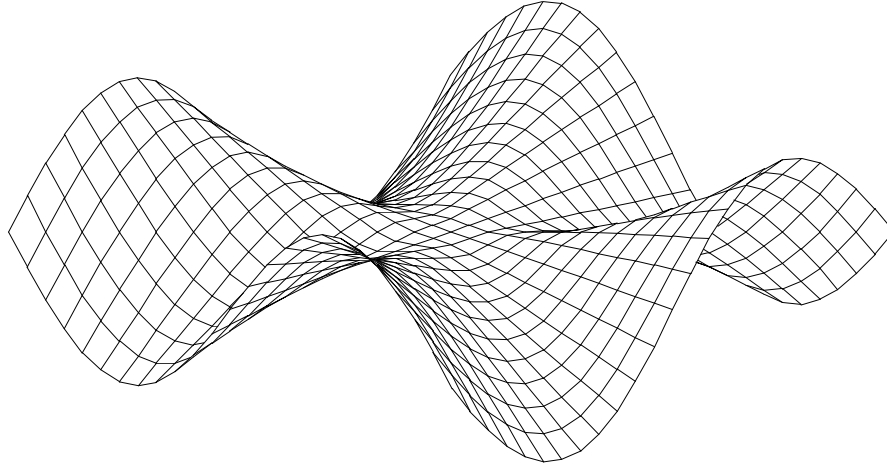
ii) On peut également considérer $\Phi_2(t) = f(x_1 + h_1, t) - f(x_1, t)$. Par un raisonnement analogue à celui qui précède, on a :

$$\begin{aligned} \Delta(h_1, h_2) &= \Phi_2(x_2 + h_2) - \Phi_2(x_2) = h_2 \Phi_2'(x_2 + \tau h_2) \text{ avec } 0 < \tau < 1 \\ &= h_2 \left(\frac{\partial f}{\partial x_2}(x_1 + h_1, x_2 + \tau h_2) - \frac{\partial f}{\partial x_2}(x_1, x_2 + \tau h_2) \right) \\ &= h_2 h_1 \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1 + \tau' h_1, x_2 + \tau h_2) \text{ avec } 0 < \tau' < 1 \end{aligned}$$

Donc $\lim_{(h_1, h_2) \rightarrow (0, 0)} \frac{\Delta(h_1, h_2)}{h_1 h_2} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_1, x_2)$

On en déduit donc l'égalité des deux dérivées partielles secondes

Voici un contre-exemple, dans le cas d'une fonction non C^2 . Soit $f(x_1, x_2) = \frac{x_1 x_2 (x_1^2 - x_2^2)}{x_1^2 + x_2^2}$ en dehors de $(0, 0)$ et $f(0, 0) = 0$



$$\frac{\partial f}{\partial x_1}(x_1, x_2) = \frac{x_2(x_1^2 - x_2^2)}{x_1^2 + x_2^2} + \frac{4x_1^2 x_2^3}{(x_1^2 + x_2^2)^2} \text{ en dehors de } (0,0)$$

$$\frac{\partial f}{\partial x_1}(0,0) = 0, \text{ dérivée en } 0 \text{ de la fonction } x_1 \rightarrow f(x_1, 0) = 0$$

$$\Rightarrow \frac{\partial f}{\partial x_1}(0, x_2) = -x_2 \text{ pour tout } x_2$$

$$\Rightarrow \frac{\partial^2 f}{\partial x_2 \partial x_1}(0, 0) = -1$$

De même :

$$\frac{\partial f}{\partial x_2}(x_1, x_2) = \frac{x_1(x_1^2 - x_2^2)}{x_1^2 + x_2^2} - \frac{4x_1^3 x_2^2}{(x_1^2 + x_2^2)^2}$$

$$\frac{\partial f}{\partial x_2}(0, 0) = 0$$

$$\Rightarrow \frac{\partial f}{\partial x_2}(x_1, 0) = x_1 \text{ pour tout } x_1$$

$$\Rightarrow \frac{\partial^2 f}{\partial x_1 \partial x_2}(0, 0) = 1$$

On remarque donc que $\frac{\partial^2 f}{\partial x_2 \partial x_1}(0, 0) \neq \frac{\partial^2 f}{\partial x_1 \partial x_2}(0, 0)$

