

Partie I: Fonctions de plusieurs variables

L'objectif général du cours de mathématiques au GSE est d'introduire les outils mathématiques qui permettent de décrire de manière précise et efficace le comportement de systèmes physiques

Par exemple la théorie des fonctions d'une variable (Mathématique I) permet de décrire l'évolution en fonction du temps d'une quantité associée à un tel système

Par exemple la fonction

$$t \rightarrow z(t) = z_0 - \frac{1}{2} g t^2$$

décrit l'altitude d'un corps soumis à l'attraction terrestre ($g = 9,81 \text{ m/s}^2$) au temps t lâché d'une hauteur z_0 .

* ponctuel(!)

Bien entendu la détermination de l'état d'un système physique peut nécessiter de connaître plusieurs paramètres en fonction du temps: si on

Exemple: si on lance le corps considéré précédemment dans une certaine direction et à une certaine vitesse, l'objet devra être repéré par ses 3 coordonnées dans l'espace

$$x(t) = x_0 + v_{x_0} t$$

$$y(t) = y_0 + v_{y_0} t$$

$$z(t) = z_0 + v_{z_0} t - \frac{1}{2} g t^2$$

où $\vec{v} = (v_{x_0}, v_{y_0}, v_{z_0})$ décrit le vecteur vitesse initiale

la fonction $t \rightarrow (x(t), y(t), z(t))$ est la trajectoire du corps

ainsi on peut se dire qu'il est possible de décrire l'évolution d'un système physique en "superposant" différentes fonctions d'une variable réelle. Cependant un système physique général est souvent beaucoup plus complexe et dépend de nombreux paramètres qui agissent entre eux de manière compliquée.

Ainsi l'accélération gravitationnelle (la constante g) n'est constante qu'en première approximation : sa valeur absolue (g) voire sa direction dépendent du point (x, y, z) où l'on se trouve

ainsi les équations du mouvement qui décrivent la trajectoire d'un corps lancé avec une vitesse initiale peuvent dépendre d'une fonction de 3 variables

$$(x, y, z) \longrightarrow g(x, y, z)$$

Ainsi il est important de savoir étudier les variations d'une fonction de plusieurs variables sur un espace à plusieurs paramètres

D'autre part, considérons le problème précédent sous l'angle de la mécanique Lagrangienne:

On repère un point par ses coordonnées $\vec{x} = (x, y, z)$ et sa vitesse $\vec{v} = (\dot{x}, \dot{y}, \dot{z})$

Soit un vecteur dans $\mathbb{R}^3 \times \mathbb{R}^3 = \mathbb{R}^6$

le principe de conservation de l'énergie dit que la somme de l'énergie cinétique et de l'énergie potentielle reste constante:

si m est la masse du corps on a pour tout t

$$\frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + mgz = \text{Cst}$$
$$= \frac{1}{2} m (v_{x_0}^2 + v_{y_0}^2 + v_{z_0}^2) + mgz_0$$

en effet $(\vec{x}(t), \vec{v}(t)) = (x_0 + v_{x_0}t, y_0 + v_{y_0}t, z_0 + v_{z_0}t - \frac{1}{2}gt^2, v_{x_0}, v_{y_0}, v_{z_0} - gt)$
et on retrouve bien l'égalité pour tout temps t .

Ainsi le point $P(t) = (x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t))$
 dans l'espace \mathbb{R}^6 reste confiné dans le domaine
 des vecteurs dont les coordonnées vérifient

$$\frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + mgz = \text{Cst}$$

Cst étant la valeur du terme
 de gauche pour $(\dot{x}, \dot{y}, \dot{z}, x, y, z) = (v_{x_0}, v_{y_0}, v_{z_0}, x_0, y_0, z_0)$

Plus généralement si la gravitation en un
 point (x, y, z) est définie par un vecteur

$$\vec{g}(x, y, z) = (g_1(x, y, z), g_2(x, y, z), g_3(x, y, z))$$

(au lieu du vecteur constant $(0, 0, -g)$)

Le tout corps $P(t)$ repéré par ses coordonnées

$$(x(t), y(t), z(t), \dot{x}(t), \dot{y}(t), \dot{z}(t))$$

sera toujours contenu dans le domaine de \mathbb{R}^6
 défini par sous-ensemble

$$\frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + m(g_1(x, y, z)x + g_2(x, y, z)y + g_3(x, y, z)z) = \text{Cst} = \text{valeur en } (x_0, y_0, z_0, v_{x_0}, v_{y_0}, v_{z_0})$$

Dans ce cours, on va donner des éléments permettant d'étudier de tels sous-ensembles des espaces de paramètres \mathbb{R}^n

Def: Un espace de paramètres est un produit de copies de \mathbb{R}
 $\underbrace{\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}}_{n \text{ fois}} = \mathbb{R}^n$. Un point de cet espace (ou vecteur)

est repéré par ses coordonnées (x_1, \dots, x_n) x_i dans \mathbb{R}
On utilisera souvent une notation vectorielle

$$\vec{x} = (x_1, \dots, x_n).$$

Rappelons qu'on a certaines opérations sur les vecteurs

- somme: $\vec{x} + \vec{x}' = (x_1, x_n) + (x'_1, x'_n) = (x_1 + x'_1, \dots, x_n + x'_n)$

- multiplication par un scalaire (homothétie):
si λ est dans \mathbb{R}

$$\lambda \cdot \vec{x} = (\lambda x_1, \dots, \lambda x_n) \text{ dans } \mathbb{R}^n$$

- produit scalaire

$$\vec{x} \cdot \vec{x}' = \sum_{i=1}^n x_i x'_i = x_1 x'_1 + \dots + x_n x'_n \text{ est un réel}$$

Notons que $\vec{x} \cdot \vec{x} = \sum_{i=1}^n x_i^2$ est le carré de la longueur euclidienne du vecteur \vec{x} .

Def: Une fonction réelle (ou bien à valeurs réelles) est une application qui à tout point d'un sous-ensemble $D(f)$ de \mathbb{R}^n associe une valeur réelle $(x_1, \dots, x_n) \rightarrow f(x_1, \dots, x_n)$

$D(f)$ est appelé domaine de définition de f .

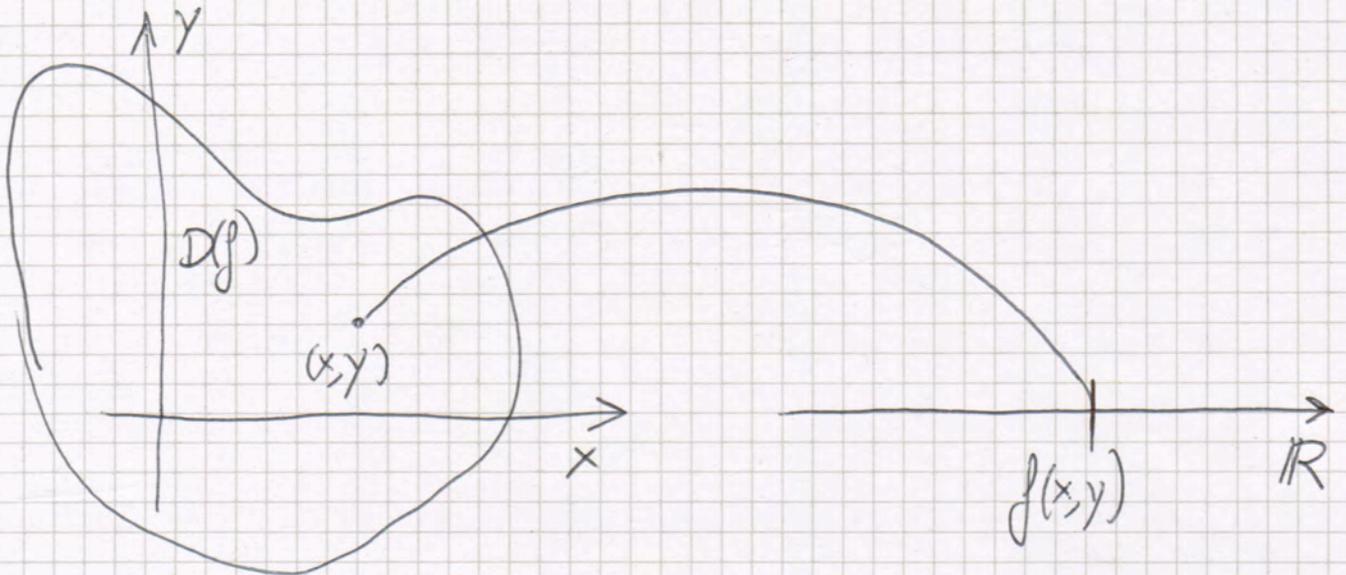
Notations: On dit aussi que f est une fonction de n variables.

$$\begin{aligned} f: D(f) &\longrightarrow \mathbb{R} \\ (x_1, \dots, x_n) &\longrightarrow f(x_1, \dots, x_n) \end{aligned}$$

On va considérer spécialement les cas $n=2$ et $n=3$ et on emploiera souvent les notations suivantes pour les coordonnées

$$(x, y) \rightarrow f(x, y) \quad \text{au lieu de} \quad (x_1, x_2) \rightarrow f(x_1, x_2)$$

$$(x, y, z) \rightarrow f(x, y, z) \quad (x_1, x_2, x_3) \rightarrow f(x_1, x_2, x_3)$$



Exemples. ① $f: \mathbb{R}^2 \rightarrow \mathbb{R}$
 $(x, y) \rightarrow \sqrt{x^2 + y^2}$ (la longueur du vecteur (x, y))

② $f: \mathbb{R}^2 \rightarrow \mathbb{R}$
 $(x, y) \rightarrow 1$ fonction constante égale à 1

③ $f: \mathbb{R}^3 \rightarrow \mathbb{R}$
 $(x, y, z) \rightarrow \frac{1}{x^2 + y^2 + z^2}$ $D(f) = \mathbb{R}^3 - \{(0, 0, 0)\}$

~~④ $f: \mathbb{R}^4 \rightarrow \mathbb{R}$
 $(x_1, x_2, x_3, x_4) \rightarrow \frac{1}{x_1^2 + x_2^2 + x_3^2 + x_4^2}$ $D(f) = \mathbb{R}^4 - \{0\}$~~

On peut "superposer" plusieurs fonctions de n variables

$$f = (f_1, \dots, f_m)$$

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}^m$$
$$(x_1, \dots, x_n) \longrightarrow (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

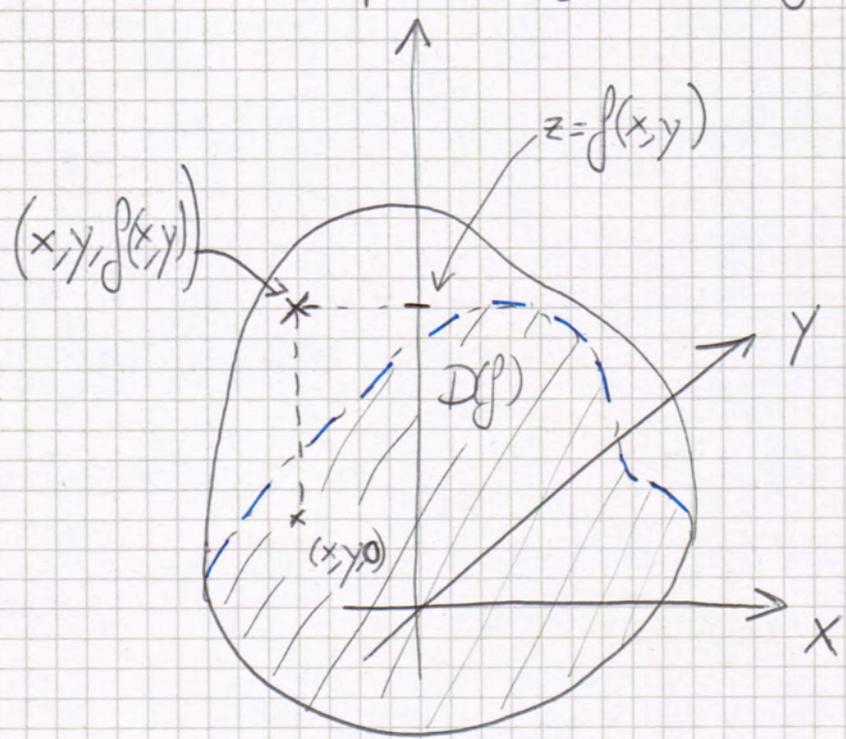
On parle alors de fonction en n variables
à valeurs dans \mathbb{R}^m

Le graphe d'une fonction de plusieurs variables

Def: Soit $f(x_1, \dots, x_n)$ le ~~graphe~~ d'une fonction
~~de n var~~ de n variables. Le graphe de f
est l'ensemble des points de \mathbb{R}^{n+1} de la forme

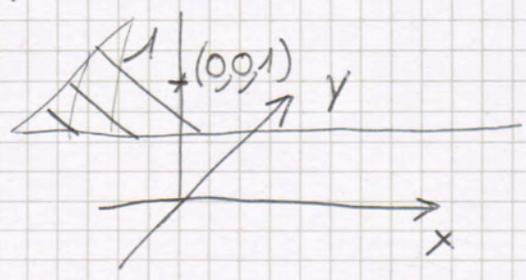
$$(x_1, \dots, x_n, f(x_1, \dots, x_n) = x_{n+1}) \text{ avec}$$
$$(x_1, \dots, x_n) \text{ dans } D(f)$$

Exemple: Si f est une fonction de 2 variables



Exemples: $f(x, y) = 1$

le graphe est un plan

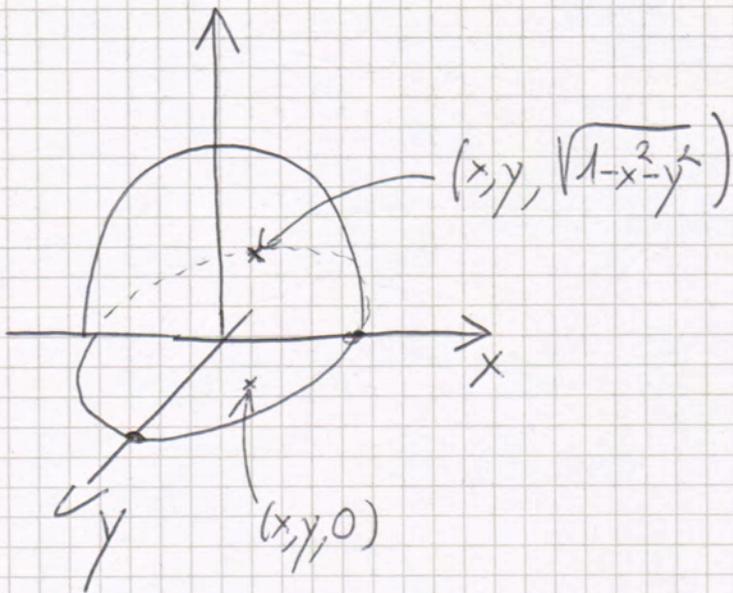


$$f(x, y) = \sqrt{1 - x^2 - y^2} \quad D(f) = \{(x, y) \mid x^2 + y^2 \leq 1\}$$

$D(f)$ est le disque centré en $(0, 0)$ et de rayon 1

$$z = \sqrt{1 - x^2 - y^2} \iff x^2 + y^2 + z^2 = 1 \text{ et } z \geq 0$$

Le graphe de f est donc la demi-sphère
("l'hémisphère nord")



Courbes de niveau Variétés

Soit f une fonction de n variables définie sur un domaine $D(f)$. Les ~~courbes~~ ^{variétés} de niveau de f sont les sous-ensembles de $D(f)$ le long desquelles ~~f~~ reste constante.

Une variété de niveau

Soit C une constante, la variété de niveau C est l'ensemble des points (x_1, \dots, x_n) de $D(f)$ qui sont liés par la relation $f(x_1, \dots, x_n) = C$

Notons qu'une variété de niveau peut être vide:
il suffit que C ne soit pas une valeur prise
par f

Exemple: $f(x, y, z) = x^2 + y^2 + z^2$

la variété de niveau C est vide si $C < 0$
est le point $(0, 0, 0)$ si $C = 0$
" la sphère de rayon \sqrt{C} si
 $C > 0$

Remarque: très souvent, et au moins si l'on se
restreint à des sous-ensembles de $D(f)$ il est
possible en utilisant la relation

$$f(x_1, \dots, x_n) = C$$

d'exprimer un des paramètres x_1, \dots, x_n en fonction
des autres. Ainsi une variété de niveau peut être
vue (au moins sur un sous-ensemble de $D(f)$)
comme le graphe d'une fonction en $n-1$ variables
Elle est donc plus facile à représenter

Les variétés de niveau permettent de se représenter une surface

Exemple: $f(x, y, z) = x^2 + y^2 + z^2$

$C > 0$ si on se restreint à $z > 0$ on a
 $z = \sqrt{C - (x^2 + y^2)}$ qui est définie pour $x^2 + y^2 \leq C$

Vocabulaire: Si f est une fonction en 2 variables on parle de courbe de niveau car la variété de niveau est souvent une courbe

Si f est une fonction en 3 variables on parle de surface de niveau

Exemple: Considérons l'espace à 6 paramètres \mathbb{R}^6
 $(x, y, z, \dot{x}, \dot{y}, \dot{z}) \quad m > 0$

$$E(x, y, z, \dot{x}, \dot{y}, \dot{z}) = \frac{1}{2} m (\dot{x}^2 + \dot{y}^2 + \dot{z}^2) + m g z$$

alors un corps ponctuel de masse m soumis à l'attraction terrestre repéré par sa position

$(x(t), y(t), z(t))$ et sa vitesse $(\dot{x}(t), \dot{y}(t), \dot{z}(t))$ évolue toujours le long d'une variété de niveau

$$E(x, y, z, \dot{x}, \dot{y}, \dot{z}) = E(x(0), y(0), z(0), \dot{x}(0), \dot{y}(0), \dot{z}(0))$$

CHAPITRE 1

Fonctions de plusieurs variables : continuité

1. Introduction, motivation

Voir les notes manuscrites. Elles seront bientôt incluses dans ce texte.

2. Graphes et variétés de niveau

Soit $D \subset \mathbf{R}^n$ un sous ensemble de \mathbf{R}^n et

$$f : \begin{array}{ccc} D & \rightarrow & \mathbf{R} \\ \vec{x} = (x_1, \dots, x_n) & \rightarrow & f(x_1, \dots, x_n) \end{array}$$

une fonction sur D à valeurs réelles. L'ensemble D (noté aussi D_f) est le *domaine de définition* de f .

Pour essayer de représenter f on associe à des sous-ensembles de \mathbf{R}^{n+1} et \mathbf{R}^n appelés le graphe et les variétés de niveau de f

2.1. Graphe d'une fonction.

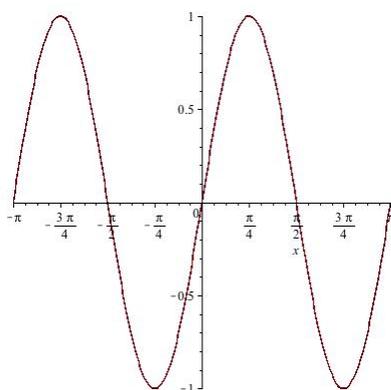
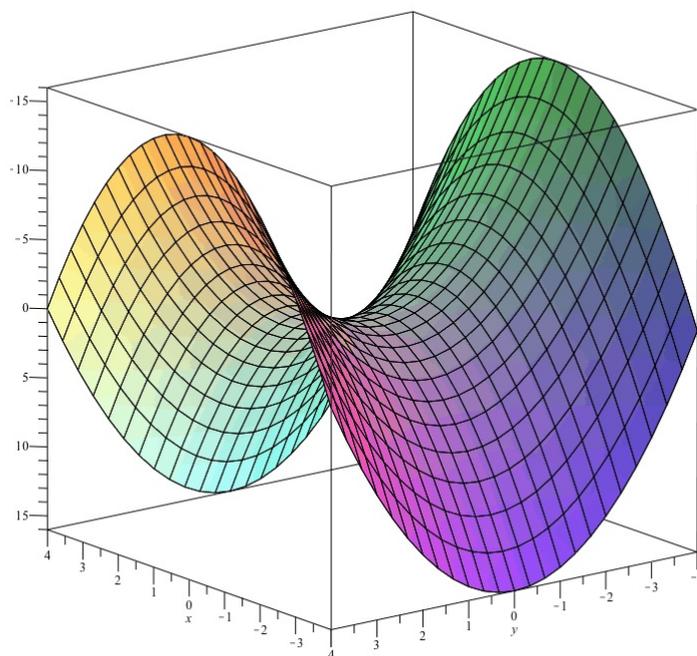


FIGURE 1. Graphe de $f(x) = \sin(2x)$

FIGURE 2. Graphe de $f(x, y) = x^2 - y^2$

DÉFINITION 1.1. Le graphe de f noté \mathcal{G}_f est le sous-ensemble de \mathbf{R}^{n+1}

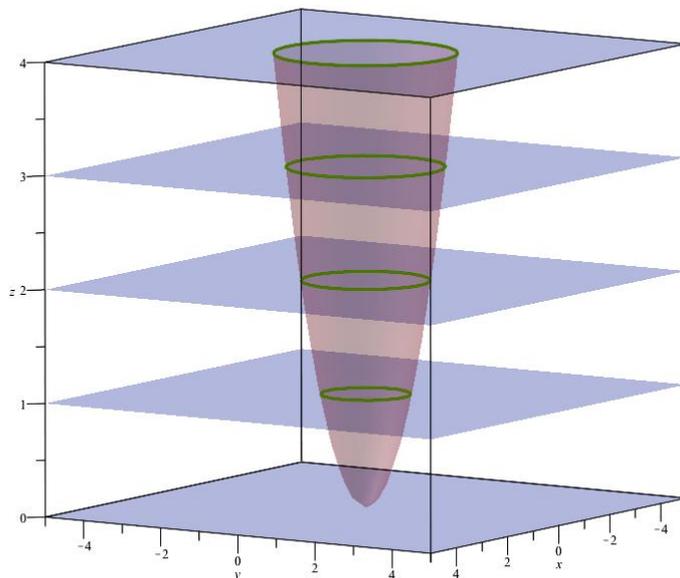
$$\begin{aligned} \mathcal{G}_f &= \{(x_1, \dots, x_n, x_{n+1}), (x_1, \dots, x_n) \in D_f, x_{n+1} = f(x_1, \dots, x_n)\} \\ &= \{(x_1, \dots, x_n, f(x_1, \dots, x_n)), (x_1, \dots, x_n) \in D_f\} \subset \mathbf{R}^{n+1} \end{aligned}$$

- Si $n = 1$, $\mathcal{G}_f \subset \mathbf{R}^2$ n'est autre que le graphe habituel d'une fonction d'une variable.
- Si $n = 2$ et qu'on représente les coordonnées de \mathbf{R}^3 par (x, y, z) , le graphe de $f : (x, y) \rightarrow f(x, y)$ est l'ensemble des points de la forme

$$\mathcal{G}_f = \{(x, y, f(x, y)), (x, y) \in D_f\} \subset \mathbf{R}^3.$$

Graphiquement il s'agit d'une surface qui quand on la projette sur le plat horizontal (des (x, y)) redonne le domaine de définition D_f .

- Pour $n \geq 3$ le graphe de f est un objet de \mathbf{R}^{n+1} qui est difficile de représenter sur un tableau ou un écran.



The intersection of the surface $f(x, y) = x^2 + y^2$ and one or more planes of the form $z = \text{constant}$.

FIGURE 3. Courbes de niveau de $f(x, y) = x^2 + y^2$

2.2. Variétés de niveau. Une autre manière d'appréhender une fonction est de considérer ses variétés de niveau qui sont cette fois des sous-ensemble de \mathbf{R}^n (donc plus faciles à représenter).

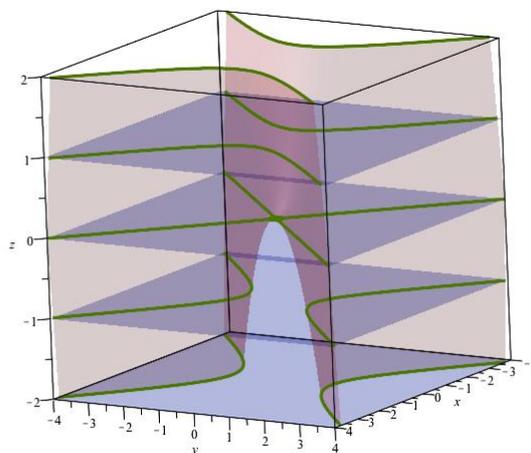
DÉFINITION 1.2. Soit $f : D_f \subset \mathbf{R}^n \rightarrow \mathbf{R}$ une fonction et C une constante, la variété de niveau C est le sous-ensemble de \mathbf{R}^n

$$V_f(C) = \{\vec{x} = (x_1, \dots, x_n), \text{ t.q. } f(x_1, \dots, x_n) = C\}.$$

Notons que variété de niveau peut être vide : si C n'appartient pas à l'image de f , $f(D_f)$.

3. Approximation et continuité

Dans ce cours on va chercher à décrire une fonction $\vec{x} \rightarrow f(\vec{x})$ quand la variable \vec{x} est "proche" d'un point $\vec{x}_0 = (x_{0,1}, \dots, x_{0,n})$. On va montrer que souvent dans ce cas f est "proche" d'une fonction g qui est une somme de fonctions "simples" à étudier.



The intersection of the surface $f(x, y) = x^2 - y^2$ and one or more planes of the form $z = \text{constant}$.

FIGURE 4. Courbes de niveau de $f(x, y) = x^2 - y^2$

3.1. Fonctions de base. Les fonctions simples qu'on va rencontrer sont essentiellement de trois types

- (1) Les fonctions constantes, $f : \vec{x} \rightarrow Cst$ où Cst est une constante (ie. qui ne dépend pas de \vec{x}).
- (2) Les fonctions linéaires; une fonction f est linéaire si elle est de la forme

$$L : (x_1, \dots, x_n) \mapsto a_1x_1 + \dots + a_nx_n$$

avec $(a_1, \dots, a_n) \in \mathbf{R}^n$ un vecteur fixe. Les fonctions linéaires satisfont

- $L(\vec{0}) = 0$,
- $L(\vec{x} + \vec{x}') = L(\vec{x}) + L(\vec{x}')$,
- $L(\lambda \cdot \vec{x}) = \lambda L(\vec{x})$.

- (3) Les fonctions homogènes quadratiques : les fonctions de la forme

$$f(x_1, \dots, x_n) = a_{11}x_1^2 + a_{12}x_1x_2 + \dots + a_{nn}x_n^2$$

avec a_{ij} , $i = 1..n$, $j = 1..n$ des nombres réels fixes.

La somme d'une fonction constante et d'une fonction linéaire sera appelée une fonction *affine* et la somme d'une fonction affine et d'une fonction homogène quadratique sera appelée fonction quadratique ou fonction polynomiale de degré 2.

On doit aussi quantifier la notion d'être "proche". Pour cela on introduit la notion de

3.2. Normes sur \mathbf{R}^n . Soit x et x_0 deux nombres réels; ils sont "proches" si leur distance

$$|x - x_0|$$

est petite.

Soient maintenant deux vecteurs

$$\vec{x} = (x_1, \dots, x_n), \vec{x}_0 = (x_{0,1}, \dots, x_{0,n}) \in \mathbf{R}^n.$$

On dira que deux vecteurs \vec{x} et \vec{x}_0 sont proches si chacune des coordonnées de l'un est proche de la coordonnée correspondante : x_1 est proche de $x_{0,1}$, ..., x_n est proche de $x_{0,n}$, ou encore posant

$$\vec{x} - \vec{x}_0 = (x_1 - x_{0,1}, \dots, x_n - x_{0,n}) = \vec{h} = (h_1, \dots, h_n),$$

si chacune des coordonnées h_i est petite. On mesure la taille d'un vecteur par une *norme* : des exemples de normes sont

- (1) La norme euclidienne : $\|\vec{h}\|_2 = (h_1^2 + \dots + h_n^2)^{1/2}$,
- (2) La norme L_1 : $\|\vec{h}\|_1 = |h_1| + \dots + |h_n|$,
- (3) La norme sup : $\|\vec{h}\|_\infty = \max_{i=1, \dots, n} |h_i|$.

Il est clair que si $\|\vec{h}\|_1$ ou $\|\vec{h}\|_2$ ou $\|\vec{h}\|_\infty$ est petit alors chacune des coordonnées de \vec{h} est petite. On a en effet

$$\|\vec{h}\|_\infty \leq \|\vec{h}\|_2 \leq \|\vec{h}\|_1 \leq n \|\vec{h}\|_\infty.$$

Il y a beaucoup d'autres normes possibles, mais toutes ont des propriétés communes :

DÉFINITION 1.3. Une norme sur \mathbf{R}^n est une fonction

$$\|\cdot\| : \vec{x} \in \mathbf{R}^n \mapsto \|\vec{x}\| \in \mathbf{R}$$

vérifiant les propriétés suivantes

- (1) (*positivité*) Pour tout \vec{x} , $\|\vec{x}\| \geq 0$
- (2) (*détection du vecteur nul*) $\|\vec{x}\| = 0$ si et seulement si $\vec{x} = \vec{0}$,
- (3) (*homogénéité*) $\|\lambda \cdot \vec{x}\| = |\lambda| \|\vec{x}\|$.
- (4) (*inégalité triangulaire*) $\|\vec{x} + \vec{x}'\| \leq \|\vec{x}\| + \|\vec{x}'\|$.

3.3. Continuité. On rappelle que

DÉFINITION 1.4. Une fonction $f : D_f \subset \mathbf{R} \rightarrow \mathbf{R}$ d'une variable définie sur un domaine $D_f \subset \mathbf{R}$ est continue en un point $x_0 \in D_f$ si, quand x tend vers x_0 , $f(x)$ tend vers $f(x_0)$; c'est à dire, si étant donné une suite numérique $(x_k)_k$ qui tend vers x_0 , la suite numérique $(f(x_k))_k$ tend vers $f(x_0)$.

On note cela

$$\lim_{x \rightarrow x_0} f(x) = f(x_0), \text{ ou } f(x) \rightarrow f(x_0) \text{ quand } x \rightarrow x_0.$$

On va étendre cette définition aux fonctions de plusieurs variables. Pour cela nous auront besoin de la notion de convergence d'une suite de vecteurs.

DÉFINITION 1.5. *Une suite de vecteurs*

$$(\vec{x}_k)_k = (x_{k,1}, \dots, x_{k,n})_k$$

tend vers un vecteur $\vec{x}_0 = (x_{0,1}, \dots, x_{0,n})$ et on le note

$$\lim_{k \rightarrow \infty} \vec{x}_k = \vec{x}_0, \text{ ou } \vec{x}_k \rightarrow \vec{x}_0, k \rightarrow +\infty,$$

si pour une (ou de manière équivalente pour toute) norme, la suite numérique

$$(\|\vec{x}_k - \vec{x}_0\|)_k$$

tend vers 0. En considérant la norme

$$\|\vec{x}\|_\infty = \max_{i=1 \dots n} |x_i|$$

on voit que la suite $(\vec{x}_k)_k$ tend vers \vec{x}_0 si et seulement si

$$\text{pour tout } i = 1 \dots, n, \text{ on a } x_{k,i} \rightarrow x_{0,i}, k \rightarrow +\infty.$$

DÉFINITION 1.6. *Soit $f : D_f \subset \mathbf{R}^n \mapsto \mathbf{R}$ une fonction, $\vec{x}_0 \in D_f$ et $l \in \mathbf{R} \cup \{-\infty, +\infty\}$. On dit que f tend vers l quand \vec{x} tend vers \vec{x}_0 , que l'on note*

$$\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = l \text{ ou bien encore } f(\vec{x}) \rightarrow l, \vec{x} \rightarrow \vec{x}_0,$$

si pour toute suite $(\vec{x}_k)_k$ de vecteurs de D_f qui tend vers \vec{x}_0 , la suite numérique $(f(\vec{x}_k))_k$ tend vers l .

On répète alors la définition de la continuité :

DÉFINITION 1.7. *Une fonction $f : D_f \subset \mathbf{R}^n \rightarrow \mathbf{R}$ de n variables définie sur un domaine $D_f \subset \mathbf{R}^n$ est continue en un point $\vec{x}_0 \in D_f$ si, quand \vec{x} tend vers \vec{x}_0 , $f(\vec{x})$ tend vers $f(\vec{x}_0)$; c'est à dire, si pour toute suite de vecteurs $(\vec{x}_k)_k$ tendant vers \vec{x}_0 , la suite numérique $(f(\vec{x}_k))_k$ tend vers $f(\vec{x}_0)$.*

On note cela

$$\lim_{\vec{x} \rightarrow \vec{x}_0} f(\vec{x}) = f(\vec{x}_0), \text{ ou } f(\vec{x}) \rightarrow f(\vec{x}_0) \text{ quand } \vec{x} \rightarrow \vec{x}_0.$$

Si une fonction f est continue en tout point \vec{x}_0 du domaine D_f on dira que f est continue sur D_f .

3.4. Approximation de fonctions continues par des fonctions constantes.

On va donner une définition équivalente de la continuité en terme d'approximation par des fonctions constantes. Pour cela on introduit la notation suivante :

NOTATION. *Dans tout ce cours, on notera $\varepsilon(\vec{h})$ une fonction sur \mathbf{R}^n bien définie pour tout \vec{h} suffisamment proche du vecteur nul $\vec{0}$ qui tend vers 0 quand \vec{h} tend vers $\vec{0}$:*

$$\lim_{\vec{h} \rightarrow \vec{0}} \varepsilon(\vec{h}) = 0 \text{ ou bien } \varepsilon(\vec{h}) \rightarrow 0, \vec{h} \rightarrow \vec{0}.$$

On a alors la définition équivalente suivante de la continuité :

DÉFINITION 1.8. Une fonction $f : D_f \subset \mathbf{R}^n \rightarrow \mathbf{R}$ de n variables définie sur un domaine $D_f \subset \mathbf{R}^n$ est continue en un point $\vec{x}_0 \in D_f$ si elle peut s'écrire sous la forme

$$(3.1) \quad f(\vec{x}) = f(\vec{x}_0) + \varepsilon(\vec{x} - \vec{x}_0) \text{ avec } \lim_{\vec{h} \rightarrow \vec{0}} \varepsilon(\vec{h}) = 0;$$

ou encore (en posant $\vec{h} = \vec{x} - \vec{x}_0$) si on a

$$(3.2) \quad f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \varepsilon(\vec{h}) \text{ avec } \lim_{\vec{h} \rightarrow \vec{0}} \varepsilon(\vec{h}) = 0.$$

Si une fonction f est continue en tout point \vec{x}_0 du domaine D_f on dira que f est continue sur D_f (la fonction $\varepsilon(\vec{h})$ de l'écriture précédente) dépend bien sûr du point \vec{x}_0).

Les expressions (3.1) et (3.2) expriment que quand \vec{x} est proche du point de référence \vec{x}_0 la fonction $f(\vec{x})$ est approximable par la fonction constante $f(\vec{x}_0)$.

3.5. Le mécano de la continuité. On va voir qu'il est très facile de fabriquer des fonctions continues :

THÉORÈME 1.1. On a les critères de continuité suivants :

- Les fonctions constantes sont continues
- Pour $i = 1, \dots, n$ les fonctions coordonnées

$$x_i : \vec{x} = (x_1, \dots, x_n) \in \mathbf{R}^n \mapsto x_i$$

sont continues en tout point de \mathbf{R}^n .

- Si f et g sont continues en \vec{x}_0 alors $f + g : \vec{x} \mapsto f(\vec{x}) + g(\vec{x})$ est continue en \vec{x}_0 .
- Si f et g sont continues en \vec{x}_0 alors $f \cdot g : \vec{x} \mapsto f(\vec{x})g(\vec{x})$ est continue en \vec{x}_0 .
- Si f et g sont continues en \vec{x}_0 alors $f/g : \vec{x} \mapsto f(\vec{x})/g(\vec{x})$ est continue en \vec{x}_0 si $g(\vec{x}_0) \neq 0$.
- Si f est continue en \vec{x}_0 et $g : \mathbf{R} \rightarrow \mathbf{R}$ est une fonction d'une variable qui est définie et continue au point $f(\vec{x}_0)$ alors la fonction composée

$$g \circ f : \vec{x} \mapsto g(f(\vec{x}))$$

est continue en \vec{x}_0 .

EXEMPLE 3.5.1. La fonction

$$f : (x, y) \in \mathbf{R}^2 \mapsto \frac{\sin(xy) + x^2y}{1 + x^2 + y^2}$$

est continue sur \mathbf{R}^2 . En effet le numérateur de f est somme d'un produit de fonctions continues et de la composée d'une fonction continue $((x, y) \rightarrow xy)$ sur \mathbf{R}^2 et d'une fonction d'une variable également continue sur \mathbf{R} (\sin); son dénominateur est continu et ne s'annule jamais sur \mathbf{R}^2 (car $x^2 + y^2 + 1 \geq 1$).

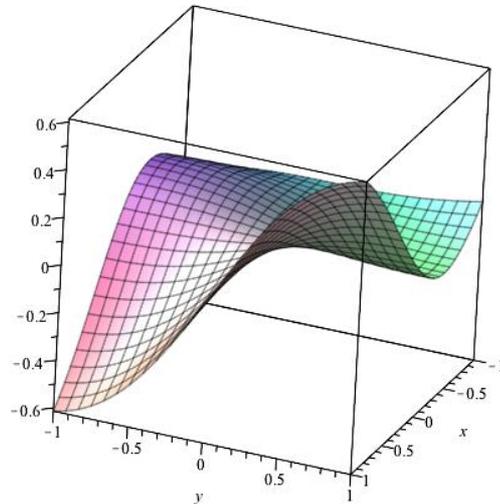


FIGURE 5. graphe de $f(x, y) = \frac{\sin(xy) + x^2y}{x^2 + y^2 + 1}$

EXEMPLE 3.5.2. En revanche la fonction définie par

$$f : (x, y) \in \mathbf{R}^2 \mapsto \begin{cases} \frac{\sin(xy) + x^2y}{x^2 + y^2} & \text{si } (x, y) \neq \vec{0} \\ 0 & \text{si } (x, y) = (0, 0) \end{cases}$$

est continue sur $\mathbf{R}^2 - \{(0, 0)\}$ (par le même raisonnement que dans l'exemple précédent) mais PAS en $(0, 0)$. Il suffit pour cela de montrer que

$$\lim_{(x,y) \rightarrow (0,0)} f(x, y) \neq f(0, 0) = 0.$$

Considérons pour cela la suite

$$\vec{x}_n = (x_n, y_n) = (1/n, 1/n).$$

cette suite tend vers $(0, 0)$ mais

$$f(x_n, x_n) = \frac{\sin(1/n^2) + 1/n^3}{2/n^2} = \frac{\sin(1/n^2)}{2/n^2} + \frac{1}{2n}.$$

Le second terme tend vers 0 quand $n \rightarrow +\infty$ alors que le premier a pour limite

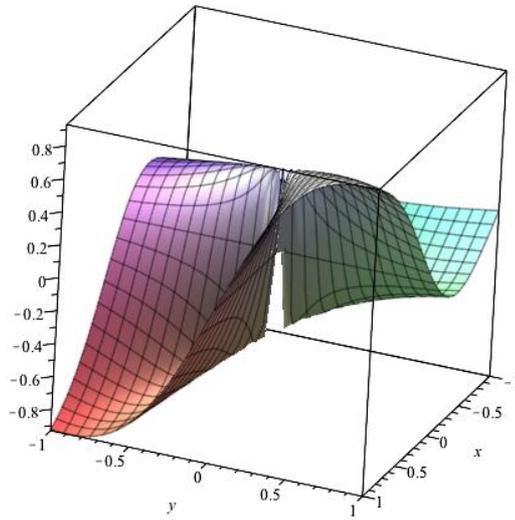
$$\lim_{n \rightarrow \infty} \frac{\sin(1/n^2)}{2/n^2} = \lim_{X \rightarrow 0} \frac{\sin(X)}{2X} = \frac{1}{2}.$$

Et donc

$$\lim_{n \rightarrow \infty} f(1/n, 1/n) = 1/2 \neq 0.$$

Ntons que si $f(0, 0)$ avait été défini comme étant $1/2$ (au lieu de 0) one aurait encore une fonction non-continue. Il suffirait de considérer la suite

$$\vec{x}'_n = (x_n, y_n) = (1/n, 0)$$

FIGURE 6. graphe de $f(x, y) = \frac{\sin(xy) + x^2y}{x^2 + y^2}$

car alors

$$\lim_{n \rightarrow \infty} f(0, 1/n) = \lim_{n \rightarrow \infty} \frac{\sin(0) + 0}{0 + (1/n)^2} = 0 \neq 1/2.$$

Fonctions de plusieurs variables : Calcul différentiel

La condition de continuité pour une fonction de plusieurs variables est une notion de régularité pratique et naturelle mais elle est trop générale et recouvre un ensemble important de phénomènes très différents. Dans ce chapitre on va discuter une condition de régularité plus restreinte : la différentiabilité.

1. Différentiabilité

La différentiabilité généralise aux fonctions de plusieurs variables la notion de dérivabilité pour les fonctions d'une variable. Cependant pour énoncer cette généralisation il est utile de reformuler la notion de dérivabilité de manière adéquate.

1.1. Rappels sur la dérivabilité. Traditionnellement, on dit qu'une fonction $f(x)$ d'une variable est dérivable en $x_0 \in \mathbf{R}$ si la limite du taux d'accroissement en x_0 existe : f est dérivable en x_0 si et seulement si

$$\lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

existe ; on appelle alors cette limite *la dérivée* de f en x_0 et on la note $f'(x_0)$. On montre alors que si f est dérivable en x_0 alors elle est continue, ainsi la notion de dérivabilité est une notion de régularité plus forte que la continuité. Pour passer aux fonctions de plusieurs variables, il sera commode de formuler la dérivabilité sous une forme équivalente : l'expression

$$(1.1) \quad \lim_{\substack{x \rightarrow x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)$$

est équivalente¹ à l'identité

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + (x - x_0)\varepsilon(x - x_0),$$

avec $\varepsilon(h) \rightarrow 0$ quand $h \rightarrow 0$. En particulier, cette formulation montre immédiatement que f est continue en x_0 ; elle s'interprète en disant que *quand x est proche de x_0 la fonction $f(x)$ est bien approximée par la fonction affine*

$$x \mapsto f(x_0) + f'(x_0)(x - x_0).$$

1. le vérifier !

REMARQUE 1.1. On a vu que la continuité (en x_0)

$$f(x) = f(x_0) + \varepsilon_1(x - x_0) \text{ avec } \varepsilon_1(h) \rightarrow 0 \text{ quand } h \rightarrow 0.$$

s'interprète en disant que *quand x est proche de x_0 $f(x)$ est approximée par la fonction constante $x \mapsto f(x_0)$* . La dérivabilité implique la continuité car la fonction

$$h \mapsto f'(x_0)h + h\varepsilon(h)$$

tend bien vers 0 quand h tend vers 0 cependant la dérivabilité dit plus car pour la continuité, l'erreur d'approximation par une fonction constante est simplement une fonction $\varepsilon_1(h)$ qui tend vers 0 en $h = 0$, alors que pour la dérivabilité l'erreur d'approximation par une fonction affine est une fonction de la forme $h\varepsilon(h)$ qui tend vers 0 en $h = 0$ mais *a priori plus vite* puisque produit de deux fonctions qui tendent vers 0 : $\varepsilon(h)$ et la fonction h . On parle pour la continuité d'approximation à l'ordre 0 et pour la dérivabilité d'approximation à l'ordre 1.

1.2. Définition de la différentiabilité.

1.2.1. *fonctions de deux variables.* On commence par le cas de deux variables qui est plus simple du point de vue des notations :

$$f : (x, y) \in D(f) \subset \mathbf{R}^2 \rightarrow \mathbf{R}$$

une fonction de deux variables et $(x_0, y_0) \in D(f)$ un point de référence.

DÉFINITION 2.1. *On dit que f est différentiable au point (x_0, y_0) si il existe deux nombres réels, $a_1, a_2 \in \mathbf{R}$, une fonction $\varepsilon : \mathbf{R}^2 \rightarrow \mathbf{R}$ vérifiant*

$$\varepsilon(u, v) \rightarrow 0 \text{ quand } (u, v) \rightarrow (0, 0)$$

et telle qu'on ait l'égalité ($\|\cdot\|$ désigne une norme quelconque)

$$f(x, y) = f(x_0, y_0) + a_1(x - x_0) + a_2(y - y_0) + \|(x - x_0, y - y_0)\|\varepsilon(x - x_0, y - y_0).$$

Alternativement, posant

$$(u, v) = (x - x_0, y - y_0),$$

f est différentiable au point (x_0, y_0) si on a l'égalité

$$f(x_0 + u, y_0 + v) = f(x_0, y_0) + a_1u + a_2v + \|(u, v)\|\varepsilon(u, v).$$

On dit que la fonction $f(x, y)$ est différentiable sur $D(f) \subset \mathbf{R}^2$ si elle est différentiable en tout point (x_0, y_0) de $D(f)$.

REMARQUE 1.2. En d'autres termes une fonction est différentiable si quand (x, y) est proche de (x_0, y_0) la fonction est bien approximable par la fonction affine

$$(x, y) \mapsto f(x_0, y_0) + a_1(x - x_0) + a_2(y - y_0) = a_1u + a_2v$$

le terme d'erreur de cette approximation étant de la forme $\|(x - x_0, y - y_0)\|\varepsilon(x - x_0, y - y_0) = \|(u, v)\|\varepsilon(u, v)$ qui tend "très vite" vers 0 puisque $\|(x - x_0, y - y_0)\|$ tend vers 0 et $\|\varepsilon(x - x_0, y - y_0)\|$ tend vers 0 : On parle d'approximation à l'ordre 1. Notons également que comme la fonction

$$(u, v) \mapsto a_1u + a_2v \rightarrow 0, \quad (u, v) \rightarrow (0, 0)$$

on a

$$f(x, y) = f(x_0, y_0) + \varepsilon_1(x - x_0, y - y_0), \quad \varepsilon_1(u, v) = a_1u + a_2v + \|(u, v)\|\varepsilon(u, v);$$

la fonction f est donc continue en (x_0, y_0) .

1.2.2. *Définition dans le cas général.* Dans le cas d'une fonction de plusieurs variables générale on répète la définition avec des notations différentes. Soit $n \geq 1$ et une fonction de n variables

$$f : \vec{x} = (x_1, \dots, x_n) \in D(f) \subset \mathbf{R}^n \mapsto f(\vec{x}) \in \mathbf{R}$$

DÉFINITION 2.2. On dit que f est différentiable en $\vec{x}_0 = (x_{0,1}, \dots, x_{0,n}) \in D(f)$, si il existe $\vec{a} = (a_1, \dots, a_n) \in \mathbf{R}^n$ et une fonction $\varepsilon(\vec{h})$ qui tend vers 0 quand $\vec{h} \rightarrow \vec{0}$ tel que on ait

$$(1.2) \quad f(\vec{x}) = f(\vec{x}_0) + a_1(x_1 - x_{0,1}) + \dots + a_n(x_n - x_{0,n}) + \|\vec{x} - \vec{x}_0\|\varepsilon(\vec{x} - \vec{x}_0).$$

De manière équivalente en écrivant $\vec{x} = \vec{x}_0 + \vec{h}$

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + a_1h_1 + \dots + a_nh_n + \|\vec{h}\|\varepsilon(\vec{h}).$$

On voit que pour $n = 1$ cette définition n'est autre que celle de la dérivabilité avec $a_1 = f'(x_0)$.

De plus si f est différentiable en \vec{x}_0 alors elle est continue au même point : est effet, considérons la fonction

$$\varepsilon_1 : \vec{h} = (h_1, \dots, h_n) \mapsto a_1h_1 + \dots + a_nh_n + \|\vec{h}\|\varepsilon(\vec{h});$$

alors $\varepsilon_1(\vec{h})$ tend vers 0 quand $\vec{h} \rightarrow 0$ et on a $f(\vec{x}) = f(\vec{x}_0) + \varepsilon_1(\vec{x} - \vec{x}_0)$.

1.2.3. *Notation sous forme de produit scalaire.* On peut écrire la différentiabilité de manière plus compacte en terme de produit scalaire : si

$$\vec{x} = (x_1, \dots, x_n) \text{ et } \vec{y} = (y_1, \dots, y_n)$$

sont des vecteurs de \mathbf{R}^n , leur produit scalaire est donné par et noté

$$\vec{x} \cdot \vec{y} = x_1y_1 + \dots + x_ny_n.$$

Posant alors $\vec{a} = (a_1, \dots, a_n)$ on a

$$a_1(x_1 - x_{0,1}) + \dots + a_n(x_n - x_{0,n}) = \vec{a} \cdot (\vec{x} - \vec{x}_0)$$

et donc

$$f(\vec{x}) = f(\vec{x}_0) + \vec{a} \cdot (\vec{x} - \vec{x}_0) + \|\vec{x} - \vec{x}_0\|\varepsilon(\vec{x} - \vec{x}_0)$$

ou encore

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + \vec{a} \cdot \vec{h} + \|\vec{h}\|\varepsilon(\vec{h}).$$

1.3. Différentielle ; dérivées partielles. Soit f une fonction différentiable en \vec{x}_0 , alors les coefficients a_1, \dots, a_n (ou ce qui revient au même le vecteur \vec{a}) est défini de manière unique : soit $i = 1 \dots n$, définissons la fonction d'une variable

$$f_i : x \mapsto f(x_{0,1}, \dots, x, \dots, x_{0,n})$$

ou la variable x est placée à la i -ième coordonnée : on a en particulier

$$f_i(x_{0,i}) = f(x_{0,1}, \dots, x_{0,i}, \dots, x_{0,n}) = f(\vec{x}_0)$$

et on voit par (1.2)

$$\begin{aligned} f_i(x) &= f_i(x_{0,i}) + a_1 \cdot 0 + \dots + a_i \cdot (x - x_{0,i}) + \dots + a_n \cdot 0 \\ &\quad + \|(0, \dots, x - x_{0,i}, \dots, 0)\| \varepsilon((0, \dots, x - x_{0,i}, \dots, 0)) \\ &= f_i(x_{0,i}) + a_i \cdot (x - x_{0,i}) + |x - x_{0,i}| \varepsilon_i(x - x_{0,i}) \end{aligned}$$

en posant

$$\varepsilon_i(h) = \varepsilon((0, \dots, h, \dots, 0)).$$

Cette fonction tend vers 0 quand h tend vers 0 et donc la fonction $x \mapsto f_i(x)$ est une fonction dérivable en $x_{0,i}$ de dérivée

$$f'_i(x_{0,i}) = a_i.$$

En d'autres termes le coefficient a_i est la dérivée au point $x_{0,i}$ de la fonction d'une variable obtenue à partir de $f(x_1, \dots, x_i, \dots, x_n)$ en fixant toutes les coordonnées x_j pour $j \neq i$ à la valeur $x_{0,j}$ et en remplaçant la i -ième coordonnée par x .

DÉFINITION 2.3. Le coefficient a_i (ie la dérivée de $f_i(x)$ en $x_{0,i}$) s'appelle la dérivée partielle de f dans la direction x_i au point \vec{x}_0 et on la note

$$a_i = f'_i(x_{0,i}) = \frac{\partial f}{\partial x_i}(\vec{x}_0).$$

EXEMPLE 1.3.1. Soit $f(x, y) = x^2y + xy^3 + 2xy$; on va voir que cette fonction est différentiable en tout point (x_0, y_0) de \mathbf{R}^2 . calculons ses dérivées partielles

$$\frac{\partial f}{\partial x}(x_0, y_0), \quad \frac{\partial f}{\partial y}(x_0, y_0).$$

Pour calculer $\frac{\partial f}{\partial x}(x_0, y_0)$, on dérive suivant x en considérant y_0 comme une constante et on évalue le résultat en x_0 : on regarde la dérivée en x_0 de la fonction

$$x \mapsto f(x, y_0) = x^2y_0 + xy_0^3 + 2xy_0.$$

Cette dérivée vaut

$$(x \mapsto f(x, y_0) = x^2y_0 + xy_0^3 + 2xy_0)' = 2xy_0 + y_0^3 + 2y_0$$

et donc

$$\frac{\partial f}{\partial x}(x_0, y_0) = 2x_0y_0 + y_0^3 + 2y_0.$$

De même pour la dérivée en y on considère la dérivée en y

$$(x \mapsto f(x_0, y) = x_0^2y + x_0y^3 + 2x_0y)' = x_0^2 + 3x_0y^2 + 2x_0$$

et donc

$$\frac{\partial f}{\partial y}(x_0, y_0) = x_0^2 + 3x_0y_0^2 + 2x_0.$$

1.3.2. *Règles de différentiation.* On dispose comme pour les fonctions d'une variable des critères suivants pour déterminer si une fonction est différentiable :

- (1) Les fonctions constantes et les fonctions coordonnées

$$x_i : \vec{x} = (x_1, \dots, x_n) \mapsto x_i$$

sont différentiables sur \mathbf{R}^n ; les dérivées partielles des fonctions constantes sont nulles et

$$\frac{\partial x_i}{\partial x_j}(\vec{x}) = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

- (2) Sommes, produits : Si $f, g : \mathbf{R}^n \mapsto \mathbf{R}$ sont différentiables en \vec{x}_0 alors $f + g$ et $f \times g$ sont différentiables en \vec{x}_0 .

- (3) Quotients : Si de plus $g(\vec{x}_0) \neq 0$ alors $\frac{f}{g}$ est différentiable en \vec{x}_0 .

- (4) Composition : Si $f : \mathbf{R}^n \mapsto \mathbf{R}$ différentiable en \vec{x}_0 et $\varphi : \mathbf{R} \mapsto \mathbf{R}$ est dérivable (il suffit en fait qu'elle soit dérivable au point $f(\vec{x}_0)$) alors

$$\varphi \circ f : \vec{x} \mapsto \varphi(f(\vec{x}))$$

est différentiable en \vec{x}_0 .

On a alors les valeurs suivants pour les dérivées partielles : pour tout $j = 1 \dots n$

- (1)

$$\frac{\partial f + g}{\partial x_j}(\vec{x}_0) = \frac{\partial f}{\partial x_j}(\vec{x}_0) + \frac{\partial g}{\partial x_j}(\vec{x}_0).$$

- (2)

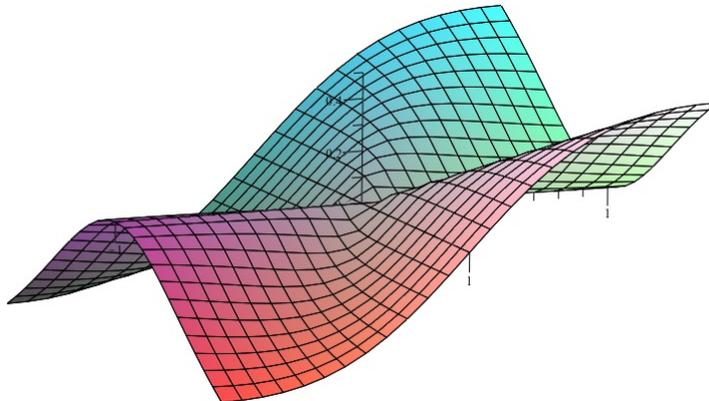
$$\frac{\partial f \times g}{\partial x_j}(\vec{x}_0) = \frac{\partial f}{\partial x_j}(\vec{x}_0)g(\vec{x}_0) + f(\vec{x}_0)\frac{\partial g}{\partial x_j}(\vec{x}_0).$$

- (3)

$$\frac{\partial f/g}{\partial x_j}(\vec{x}_0) = \frac{1}{g(\vec{x}_0)^2} \left(\frac{\partial f}{\partial x_j}(\vec{x}_0)g(\vec{x}_0) - f(\vec{x}_0)\frac{\partial g}{\partial x_j}(\vec{x}_0) \right).$$

- (4)

$$\frac{\partial \varphi \circ f}{\partial x_j}(\vec{x}_0) = \varphi'(f(\vec{x}_0))\frac{\partial f}{\partial x_j}(\vec{x}_0).$$

FIGURE 1. $f(x, y) = x^2y/(x^2 + y^2)$

1.4. Différentiabilité et dérivées partielles : ATTENTION! Comme on vient de la voir la différentiabilité d'une fonction $f : D(f) \subset \mathbf{R}^n \rightarrow \mathbf{R}$ en \vec{x}_0 implique l'existence des dérivées partielles en ce point

$$\frac{\partial f}{\partial x_j}(\vec{x}_0) = \lim_{x_j \rightarrow x_{0,j}} \frac{f_j(x_j) - f(\vec{x}_0)}{x_j - x_{0,j}}, \quad j = 1 \dots n.$$

ATTENTION la réciproque n'est PAS vraie : par exemple la fonction

$$f(x, y) = \frac{x^2y}{x^2 + y^2}, \quad (x, y) \neq (0, 0), \quad f(0, 0) = 0$$

est différentiable sur $\mathbf{R}^2 - \{(0, 0)\}$ mais en $(0, 0)$ elle n'est pas différentiable bien que ses deux dérivées partielles existent en $(0, 0)$

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{x \rightarrow 0} \frac{f(x, 0) - f(0, 0)}{x} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{y \rightarrow 0} \frac{f(0, y) - f(0, 0)}{y} = 0.$$

Supposons cette fonction différentiable en $(0, 0)$, on aura alors pour (x, y) proche de $(0, 0)$

$$f(x, y) = f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y + \|(x, y)\|\varepsilon(x, y) = \|(x, y)\|\varepsilon(x, y).$$

Prenons $(x, y) = (x, x)$ et calculons

$$\lim_{x \rightarrow 0} \frac{f(x, x) - f(0, 0)}{x}.$$

D'après la dernière égalité cela vaut ($\|(x, y)\| = |x|\|(1, 1)\|$)

$$\lim_{x \rightarrow 0} \frac{|x|\varepsilon(x, x)}{x} = 0.$$

D'autre part

$$\frac{f(x, x) - f(0, 0)}{x} = \frac{1}{x} \times \frac{x^2 \times x}{x^2 + x^2} = \frac{1}{2},$$

contradiction. \square

On a cependant le résultat plus positif suivant en faisant une hypothèse de régularité des dérivées partielles :

THÉORÈME 2.1. *Soit $f : D(f) \subset \mathbf{R}^n \rightarrow \mathbf{R}$ une fonction de n variables. On suppose que en tout point \vec{x}_0 de $D(f)$ les dérivées partielles*

$$\frac{\partial f}{\partial x_j}(\vec{x}_0) = \lim_{x_j \rightarrow x_{0,j}} \frac{f_j(x_j) - f(\vec{x}_0)}{x_j - x_{0,j}}, \quad j = 1 \dots n$$

existent et que les fonctions

$$\frac{\partial f}{\partial x_j} : \vec{x}_0 \rightarrow \frac{\partial f}{\partial x_j}(\vec{x}_0)$$

sont continues en \vec{x}_0 , alors f est différentiable sur $D(f)$.

1.5. Différentielle ; Gradient. Le vecteur des dérivées partielles

$$\vec{a} = (a_1, \dots, a_n) = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right)$$

est appelé *le gradient* de f au point \vec{x}_0 et est noté

$$\nabla(f)(\vec{x}_0) = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right).$$

On trouve aussi comme notation du gradient

$$\overrightarrow{\text{Grad}}(f)(\vec{x}_0).$$

La *différentielle* de f en \vec{x}_0 est la fonction sur \mathbf{R}^n

$$df(\vec{x}_0) : \vec{h} = (h_1, \dots, h_n) \mapsto \vec{\nabla}(f)(\vec{x}_0) \cdot \vec{h} = \frac{\partial f}{\partial x_1}(\vec{x}_0)h_1 + \dots + \frac{\partial f}{\partial x_n}(\vec{x}_0)h_n.$$

C'est une fonction linéaire (également appelée *forme linéaire*) :

$$df(\vec{x}_0)(\vec{h} + \vec{h}') = df(\vec{x}_0)(\vec{h}) + df(\vec{x}_0)(\vec{h}'), \quad df(\vec{x}_0)(\lambda \vec{h}) = \lambda df(\vec{x}_0)(\vec{h}).$$

On écrira de manières équivalentes

$$\begin{aligned} f(\vec{x}) &= f(\vec{x}_0) + df(\vec{x}_0)(\vec{x} - \vec{x}_0) + \|\vec{x} - \vec{x}_0\| \varepsilon(\vec{x} - \vec{x}_0), \\ f(\vec{x}) &= f(\vec{x}_0) + \nabla(f)(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \|\vec{x} - \vec{x}_0\| \varepsilon(\vec{x} - \vec{x}_0), \\ f(\vec{x}_0 + \vec{h}) &= f(\vec{x}_0) + df(\vec{x}_0)(\vec{h}) + \|\vec{h}\| \varepsilon(\vec{h}), \\ f(\vec{x}_0 + \vec{h}) &= f(\vec{x}_0) + \nabla(f)(\vec{x}_0) \cdot \vec{h} + \|\vec{h}\| \varepsilon(\vec{h}). \end{aligned}$$

1.6. Approximation linéaire. La fonction

$$\vec{x} \mapsto f(\vec{x}_0) + \frac{\partial f}{\partial x_1}(\vec{x}_0).(x_1 - x_{0,1}) + \cdots + \frac{\partial f}{\partial x_n}(\vec{x}_0).(x_n - x_{0,n})$$

est une fonction affine (somme d'une constante et d'une fonction linéaire). La définition de la différentiabilité

$$f(\vec{x}) = f(\vec{x}_0) + \frac{\partial f}{\partial x_1}(\vec{x}_0).(x_1 - x_{0,1}) + \cdots + \frac{\partial f}{\partial x_n}(\vec{x}_0).(x_n - x_{0,n}) + \|\vec{x} - \vec{x}_0\|\varepsilon(\vec{x} - \vec{x}_0)$$

s'interprète en disant que la fonction $\vec{x} \mapsto f(\vec{x})$ est bien approximée par cette fonction affine quand \vec{x} est proche de \vec{x}_0 . Notons que (si le vecteur $df(\vec{x}_0) = \nabla f(\vec{x}_0)$ est non nul, le terme d'erreur fait dans cette approximation $\|\vec{x} - \vec{x}_0\|\varepsilon(\vec{x} - \vec{x}_0)$ est un produit de deux termes tendant vers 0 et en tous cas, tend vers zéro "plus vite" que le terme $\frac{\partial f}{\partial x_1}(\vec{x}_0)(x_1 - x_{0,1}) + \cdots + \frac{\partial f}{\partial x_n}(\vec{x}_0)(x_n - x_{0,n})$.

2. Plan tangent

2.1. Notions concernant les vecteurs et l'algèbre linéaire. Etant donné $\vec{x} = (x_1, \dots, x_n)$ et $\vec{y} = (y_1, \dots, y_n)$ deux vecteurs de \mathbf{R}^n . Les nombres réels $x_i, x'_i, i = 1, \dots, n$ sont les *coordonnées* de ces vecteurs ; on rappelle les opérations et notations suivantes

– **Le vecteur nul** : $\vec{0} = (0, \dots, 0) \in \mathbf{R}^n$

– **Somme** :

$$\vec{x} + \vec{y} = (x_1 + y_1, \dots, x_n + y_n) \in \mathbf{R}^n,$$

– **Multiplication par un scalaire** : pour $\lambda \in \mathbf{R}$,

$$\lambda.\vec{x} = (\lambda x_1, \dots, \lambda x_n).$$

– **Produit scalaire** :

$$\vec{x}.\vec{y} = x_1 y_1 + \cdots + x_n y_n \in \mathbf{R}.$$

Ce dernier vérifie

(1) **Linéarité** : $(\vec{x} + \vec{y}).\vec{z} = \vec{x}.\vec{z} + \vec{y}.\vec{z}$, $(\lambda\vec{x}).\vec{y} = \lambda(\vec{x}.\vec{y})$.

(2) **Symétrie** : $\vec{x}.\vec{y} = \vec{y}.\vec{x}$.

(3) **Longueur euclidienne** :

$$(\vec{x}.\vec{x})^{1/2} = (x_1^2 + \cdots + x_n^2)^{1/2} = \|\vec{x}\|_2$$

(4) **Inégalité de Cauchy-Schwarz** : pour $\vec{x}, \vec{y} \in \mathbf{R}^n$ on a

$$|\vec{x}.\vec{y}| \leq \|\vec{x}\|_2 \|\vec{y}\|_2$$

et l'égalité à lieu si et seulement si \vec{x} et \vec{y} sont colinéaires (si $\vec{x} \neq \vec{0}$, il existe $\lambda \in \mathbf{R}$ tel que $\vec{y} = \lambda\vec{x}$).

(5) **Angle de deux vecteurs.** si \vec{x} et \vec{y} sont non-nuls, on définit leur l'angle $\theta \in [0, \pi]$ comme l'unique solution (dans $[0, \pi]$) de l'équation

$$\vec{x}.\vec{y} = \cos(\theta) \|\vec{x}\|_2 \|\vec{y}\|_2$$

(par l'inégalité de Cauchy-Schwarz on sait que $\vec{x}.\vec{y}/\|\vec{x}\|_2 \|\vec{y}\|_2 \in [-1, 1]$).

2.2. Hyperplans de \mathbf{R}^n .

DÉFINITION 2.4. Soit $\vec{a} = (a_1, \dots, a_n)$, $\vec{x}_0 = (x_{0,1}, \dots, x_{0,n}) \in \mathbf{R}^n$ deux vecteurs ; on suppose que $\vec{a} \neq \vec{0}$, l'hyperplan $HP(\vec{a}, \vec{x}_0)$ de \mathbf{R}^n perpendiculaire à \vec{a} et passant par \vec{x}_0 est l'ensemble des vecteurs \vec{x} tels que $\vec{x} - \vec{x}_0$ est perpendiculaire à \vec{a} , c'est à dire qui vérifient

$$0 = \vec{a} \cdot (\vec{x} - \vec{x}_0) = a_1(x_1 - x_{0,1}) + \dots + a_n(x_n - x_{0,n}).$$

REMARQUE 2.1. Si $\vec{a} = \vec{0}$, l'équation devient

$$0 = 0$$

et on obtient \mathbf{R}^n en entier, c'est pourquoi on exclut ce cas dans la suite.

Alternativement, on voit en posant $\vec{x} = \vec{x}_0 + \vec{h}$ ($x_i = x_{0,i} + h_i$) que $HP(\vec{a}, \vec{x}_0)$ est obtenu à partir de l'hyperplan $HP(\vec{a}, \vec{0})$ formé des vecteurs perpendiculaires à \vec{a} (les vecteurs $\vec{h} = (h_1, \dots, h_n)$ tels que $\vec{a} \cdot \vec{h} = a_1 h_1 + \dots + a_n h_n = 0$),

$$HP(\vec{a}, \vec{x}_0) = \vec{x}_0 + HP(\vec{a}, \vec{0}).$$

Le vecteur \vec{a} est appelé un *vecteur normal* à l'hyperplan, et \vec{x}_0 est un *vecteur de translation* associé à cet hyperplan. On dit encore que l'hyperplan "passe" par \vec{x}_0 puisque \vec{x}_0 est contenu dedans.

Soit

$$b = -a_1 x_{0,1} + \dots - a_n x_{0,n}$$

on voit alors qu'une équation de cet hyperplan est donnée par

$$0 = a_1 x_1 + \dots + a_n x_n + b.$$

$$Eq(\vec{a}, b) : \vec{a} \cdot \vec{x} = a_1 x_1 + \dots + a_n x_n = b$$

On utilisera donc également la notation

$$Hyp(\vec{a}, b) = \{ \vec{x} \in \mathbf{R}^n, \vec{a} \cdot \vec{x} = a_1 x_1 + \dots + a_n x_n = b \} = HP(\vec{a}, \vec{x}_0).$$

Réciproquement toute relation linéaire de la forme précédente (avec $(a_1, \dots, a_n) = \vec{a} \neq \vec{0}$) définit un hyperplan de vecteur normal \vec{a} et dont un vecteur de translation est

$$\vec{x}_0 = (0, \dots, -b/a_i, \dots, 0)$$

pour i tel que $a_i \neq 0$.

2.2.1. *Le cas $n = 1$.* Dans le cas $n = 1$, les vecteurs $\vec{a}, \vec{x}_0, \vec{x}$ sont des nombres $a, x_0, x \in \mathbf{R}$ et l'hyperplan $HP(a, x_0)$ a pour équation

$$0 = a(x - x_0)$$

c'est à dire est réduit à $\{x_0\}$.

2.2.2. *Le cas $n = 2$.* on a $\vec{a} = (a_1, a_2)$, $\vec{x} = (x, y)$, $\vec{x}_0 = (x_0, y_0)$ et $HP(\vec{a}, \vec{x}_0)$ a pour équation

$$0 = a_1(x - x_0) + a_2(y - y_0)$$

qui est l'équation de la droite passant par (x_0, y_0) et qui est perpendiculaire au vecteur (a_1, a_2) .

2.2.3. *Le cas $n = 3$.* on a $\vec{a} = (a_1, a_2, a_3)$, $\vec{x} = (x, y, z)$, $\vec{x}_0 = (x_0, y_0, z_0)$ et $HP(\vec{a}, \vec{x}_0)$ a pour équation

$$0 = a_1(x - x_0) + a_2(y - y_0) + a_3(z - z_0)$$

qui est l'équation du plan passant par (x_0, y_0, z_0) et qui est perpendiculaire au vecteur (a_1, a_2, a_3) .

2.3. Plan tangent au graphe d'une fonction. Comme nous l'avons expliqué, si une fonction est différentiable en un point \vec{x}_0 alors "au voisinage" de \vec{x}_0 , la fonction $f(\vec{x})$ est bien approximée par la fonction affine

$$(2.1) \quad \vec{x} \mapsto f(\vec{x}_0) + \vec{a} \cdot (\vec{x} - \vec{x}_0)$$

avec

$$\vec{a} = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0) \right) = \nabla f(\vec{x}_0).$$

Géométriquement cela signifie encore que dans \mathbf{R}^{n+1} , "au voisinage" du point $(\vec{x}_0, f(\vec{x}_0))$, le graphe de la fonction (2.1) est "proche" du graphe de f . Ce dernier graphe est un hyperplan dont l'équation est (les coordonnées de \mathbf{R}^{n+1} étant notées $(x_1, \dots, x_n, x_{n+1})$)

$$(2.2) \quad x_{n+1} = z_0 + \vec{a} \cdot (\vec{x} - \vec{x}_0),$$

où on a posé $z_0 = f(\vec{x}_0)$. C'est hyperplan est appelé *hyperplan tangent* au graphe de f au point

$$(\vec{x}_0, f(\vec{x}_0)) = (x_{0,1}, \dots, x_{0,n}, z_0).$$

2.3.1. *Le cas $n = 1$.* Dans ce cas on dispose d'une fonction d'une variable $x \in \mathbf{R} \mapsto f(x)$ dont le graphe est l'ensemble

$$G_f := \{(x, f(x)), x \in D_f\} \subset \mathbf{R}^2.$$

Au point x_0 le gradient de f est simplement la dérivée

$$\nabla f(x_0) = f'(x_0)$$

et l'hyperplan tangent au graphe au point $(x_0, f(x_0))$ a pour équation

$$y = f(x_0) + f'(x_0)(x - x_0)$$

c'est à dire la droite tangente au graphe au point $(x_0, f(x_0))$.

2.4. Vecteur normal au graphe d'une fonction. L'équation du plan tangent se réécrit

$$0 = \vec{a} \cdot (\vec{x} - \vec{x}_0) - (z - z_0) = (a_1, \dots, a_n, -1) \cdot (x_1 - x_{0,1}, \dots, x_n - x_{0,n}, x_{n+1} - z_0).$$

C'est donc l'hyperplan associé au vecteur normal

$$\vec{n}_f(\vec{x}_0) = (\nabla f(\vec{x}_0), -1) = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0), -1 \right)$$

et passant par le point $(\vec{x}_0, f(\vec{x}_0))$.

DÉFINITION 2.5. *Le vecteur $\vec{n}_f(\vec{x}_0)$ est le vecteur normal au graphe de f au point $(\vec{x}_0, f(\vec{x}_0)) \in \mathbf{R}^{n+1}$.*

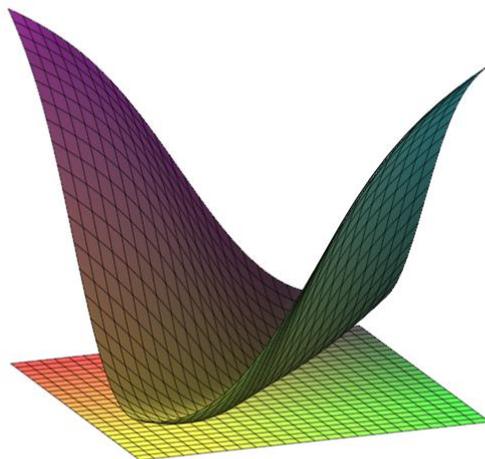


FIGURE 2. Graphe de $f(x, y) = ((x - y)^2 + x^4 + y^4)/(1 + 10(x^2 + y^2))$ et son plan tangent en $(0, 0, 0)$

2.5. Plan tangent en une variété de niveau. On a déjà vu que le gradient permet de former le *vecteur normal* au graphe de f en $(\vec{x}_0, f(\vec{x}_0))$:

$$\vec{n}_f(\vec{x}_0) = (\nabla(f)(\vec{x}_0), -1).$$

Soit $C = f(\vec{x}_0)$ et considérons maintenant la variété de niveau

$$V_f(C) = \{\vec{x} \in \mathbf{R}^n, f(\vec{x}) = C\}$$

Cette variété est contenue dans \mathbf{R}^n et contient évidemment le point \vec{x}_0 . On voudrait décrire la structure de la variété de niveau $V_f(C)$ au voisinage de \vec{x}_0 : soit \vec{x} sur $V_f(C)$ et proche de \vec{x}_0 ; \vec{x} vérifie

$$f(\vec{x}) = f(\vec{x}_0)$$

(ie. est dans la variété de niveau) d'autre part, part la formule d'approximation linéaire, on a

$$f(\vec{x}) = f(\vec{x}_0) = f(\vec{x}_0) + \nabla(f)(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) + \|\vec{x} - \vec{x}_0\| \varepsilon(\vec{x} - \vec{x}_0)$$

soit en simplifiant

$$\nabla(f)(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = -\|\vec{x} - \vec{x}_0\| \varepsilon(\vec{x} - \vec{x}_0)$$

ou encore si $\vec{x} \neq \vec{x}_0$

$$\nabla(f)(\vec{x}_0) \cdot \frac{\vec{x} - \vec{x}_0}{\|\vec{x} - \vec{x}_0\|} = \varepsilon(\vec{x} - \vec{x}_0) \rightarrow 0$$

quand $\vec{x} \rightarrow \vec{x}_0$: en d'autres termes le vecteur de longueur 1 $\frac{\vec{x}-\vec{x}_0}{\|\vec{x}-\vec{x}_0\|}$ est "presque" orthogonal au vecteur $\nabla(f)(\vec{x}_0)$; si $\nabla(f)(\vec{x}_0) \neq 0$ cela signifie aussi que \vec{x} est presque contenu dans l'hyperplan $HP(\nabla f(\vec{x}_0), \vec{x}_0)$ passant par \vec{x}_0 et perpendiculaire à $\nabla f(\vec{x}_0)$.

DÉFINITION 2.6. Soit $f : \mathbf{R}^n \rightarrow \mathbf{R}$ différentiable en \vec{x}_0 et soit $C = f(\vec{x}_0)$. Si $\nabla(f)(\vec{x}_0) \neq \vec{0}$ alors l'hyperplan de \mathbf{R}^n donné par l'équation

$$0 = \nabla(f)(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0) = \frac{\partial f}{\partial x_1}(\vec{x}_0)(x_1 - x_{0,1}) + \cdots + \frac{\partial f}{\partial x_n}(\vec{x}_0)(x_n - x_{0,n})$$

est appelé plan tangent en \vec{x}_0 à la variété de niveau $V_f(C)$. Le gradient $\nabla f(\vec{x}_0)$ est le vecteur normal en \vec{x}_0 à la variété de niveau $V_f(C)$.

On interprète la discussion précédente en disant que quand on est "près" de \vec{x}_0 , la variété de niveau $V_f(C)$ "ressemble" à un hyperplan autrement dit :

2.5.1. *La terre est plate.* En effet la surface de la terre est donné par l'équation

$$x^2 + y^2 + z^2 = R^2, \quad R = 6378 \text{ Km}$$

et si un humain (0.0018 Km) se tient au point (x_0, y_0, z_0) de la surface terrestre,, il aura l'impression que la surface de la terre ressemble au plan tangent d'équation

$$0 = x_0(x - x_0) + y_0(y - y_0) + z_0(z - z_0).$$

2.6. Plan tangent au graphe/plan tangent à une variété de niveau. On notera que cette définition du vecteur normal à une variété de niveau en terme du gradient est "compatible" avec la notion de vecteur normal associé au graphe d'une fonction. En effet le graphe d'une fonction f de \mathbf{R}^n , n'est autre que la variété de niveau 0, $V_F(0)$ associée à la fonction sur \mathbf{R}^{n+1} ,

$$F(x_1, \dots, x_n, z) := f(x_1, \dots, x_n) - z;$$

En effet le graphe de f est défini par l'équation

$$F(x_1, \dots, x_n, z) = f(x_1, \dots, x_n) - z = 0.$$

On a, en posant $z_0 = f(\vec{x}_0)$,

$$\nabla F(\vec{x}_0, z_0) = \left(\frac{\partial f}{\partial x_1}(\vec{x}_0), \dots, \frac{\partial f}{\partial x_n}(\vec{x}_0), \frac{\partial F}{\partial z}(z_0) \right) = (\nabla(f)(\vec{x}_0), -1) = n_f(\vec{x}_0).$$

2.7. Point critique. Si $\nabla f(\vec{x}_0) = 0$, le plan tangent à $V_f(f(\vec{x}_0))$ n'est pas défini : on obtient l'équation $0 = 0$ satisfaite par tous les points de \mathbf{R}^n : on dit alors que \vec{x}_0 est un *point critique*. Voir le chapitre suivant pour une discussion des points critique dans le cas des fonctions de deux variables.

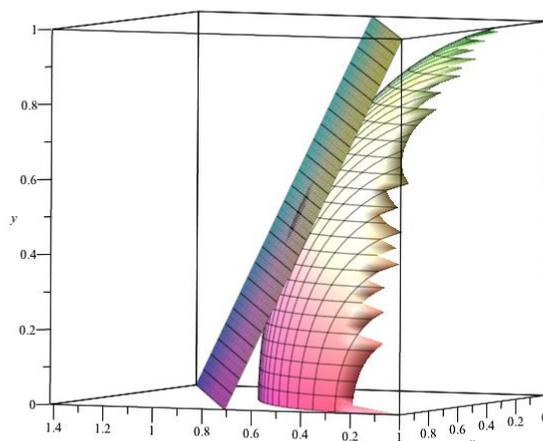


FIGURE 3. La sphere et un plan tangent.

3. Dérivée directionnelle d'une fonction

La i -ième dérivée partielle $\frac{\partial f}{\partial x_i}(\vec{x}_0)$ d'une fonction différentiable décrit la variation de f au voisinage de \vec{x}_0 quand on se déplace à partir de \vec{x}_0 parallèlement à l'axe de la coordonnée x_i . La notion de dérivée directionnelle est une extension de cette notion quand on se déplace le long d'un axe général :

DÉFINITION 2.7. Soit f une fonction sur \mathbf{R}^n différentiable en \vec{x}_0 et soit $\vec{e} = (e_1, \dots, e_n)$ un vecteur de \mathbf{R}^n ; la dérivée de f en \vec{x}_0 dans la direction \vec{e} , $D_{\vec{e}}f(\vec{x}_0)$ est donnée par la limite

$$D_{\vec{e}}f(\vec{x}_0) = \lim_{\substack{h \rightarrow 0 \\ h \neq 0}} \frac{f(\vec{x}_0 + h\vec{e}) - f(\vec{x}_0)}{h}.$$

Notons que cette limite existe : en effet par la formule d'approximation linéaire, on a

$$f(\vec{x}_0 + h\vec{e}) - f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot h\vec{e} + |h| \|\vec{e}\| \varepsilon(h\vec{e})$$

et donc

$$\frac{f(\vec{x}_0 + h\vec{e}) - f(\vec{x}_0)}{h} = \nabla f(\vec{x}_0) \cdot \vec{e} + \|\vec{e}\| \varepsilon(h\vec{e}) \rightarrow \nabla f(\vec{x}_0) \cdot \vec{e} \quad h \rightarrow 0.$$

Ainsi

$$D_{\vec{e}}f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{e} = \frac{\partial f}{\partial x_1}(\vec{x}_0)e_1 + \dots + \frac{\partial f}{\partial x_n}(\vec{x}_0)e_n.$$

REMARQUE 3.1. Si $\vec{e} = \vec{e}_i = (0, \dots, 1, 0, \dots, 0)$ (le 1 en i -ème position)

$$D_{\vec{e}_i} f = \frac{\partial f}{\partial x_i}.$$

Ainsi la dérivée directionnelle généralise la notion de dérivées partielles.

3.1. Propriétés de la dérivée directionnelle. Soient f et g des fonction différentiables en \vec{x}_0 et λ un réel, on vérifie facilement que la dérivée directionnelle est *linéaire en \vec{e}* :

$$D_{\vec{e} + \vec{e}'}(f)(\vec{x}_0) = D_{\vec{e}}(f)(\vec{x}_0) + D_{\vec{e}'}(f)(\vec{x}_0)$$

et

$$D_{\lambda \vec{e}}(f)(\vec{x}_0) = \lambda D_{\vec{e}}(f)(\vec{x}_0).$$

Appliquant cela à la décomposition

$$\vec{e} = (e_1, \dots, e_n) = e_1 \vec{e}_1 + \dots + e_n \vec{e}_n$$

on en déduit immédiatement des identités analogues à celles des dérivées partielles $\frac{\partial f}{\partial x_i}$ les identités

(1)

$$D_{\vec{e}}(f + g)(\vec{x}_0) = D_{\vec{e}}(f)(\vec{x}_0) + D_{\vec{e}}(g)(\vec{x}_0).$$

(2)

$$D_{\vec{e}}(f \times g)(\vec{x}_0) = D_{\vec{e}}(f)(\vec{x}_0)g(\vec{x}_0) + f(\vec{x}_0)D_{\vec{e}}(g)(\vec{x}_0).$$

(3)

$$D_{\vec{e}}(f/g)(\vec{x}_0) = \frac{1}{g(\vec{x}_0)^2}(D_{\vec{e}}(f)(\vec{x}_0)g(\vec{x}_0) - f(\vec{x}_0)D_{\vec{e}}(g)(\vec{x}_0).$$

(4)

$$D_{\vec{e}}(\varphi \circ f)(\vec{x}_0) = \varphi'(f(\vec{x}_0))D_{\vec{e}}(f)(\vec{x}_0).$$

3.2. Variation d'une fonction dans une direction donnée. Supposons que \vec{e} est un vecteur *unitaire* : ie un vecteur de longueur euclidienne 1

$$\|\vec{e}\| = \sqrt{e_1^2 + \dots + e_n^2} = 1.$$

La formule d'approximation linéaire s'écrit

$$f(\vec{x}_0 + h\vec{e}) = f(\vec{x}_0) + hD_{\vec{e}}f(\vec{x}_0) + h\varepsilon(h\vec{e}).$$

Comme $h\varepsilon(h\vec{e})$ est une erreur, la dérivée directionnelle $D_{\vec{e}}f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{e}$ mesure donc *l'amplitude* de la variation de la fonction f quand la variable \vec{x} se déplace depuis \vec{x}_0 dans la direction \vec{e} . Par exemple cette amplitude est *nulle* (f varie très peu) si

$$D_{\vec{e}}f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{e} = 0$$

c'est à dire si la direction \vec{e} du déplacement est *perpendiculaire* au gradient, ou ce qui revient au même, si le déplacement à lieu "le long" du plan tangent à la variété de niveau

$$f(\vec{x}) = f(\vec{x}_0).$$

Inversement, si \vec{e} est *colinéaire* à $\nabla f(\vec{x}_0)$ alors

$$D_{\vec{e}}f(\vec{x}_0) = \nabla f(\vec{x}_0) \cdot \vec{e} = \pm \|\nabla f(\vec{x}_0)\|$$

et le coefficient de variation est le plus grand possible : en effet par l'inégalité de Cauchy-Schwarz on a toujours

$$|\nabla f(\vec{x}_0) \cdot \vec{e}| \leq \|\nabla f(\vec{x}_0)\| \|\vec{e}\| = \|\nabla f(\vec{x}_0)\|$$

avec égalité si et seulement si \vec{e} et $\nabla f(\vec{x}_0)$ sont colinéaires. Ainsi un déplacement de \vec{x} dans la direction du gradient produit une variation *maximale* de la valeur de la fonction f .

REMARQUE 3.2. Cela est naturellement intuitif : si on regarde une carte topographique, la variation d'altitude est la plus faible si on se déplace parallèlement aux courbes de niveaux et la variation d'altitude est la plus forte si on se déplace perpendiculairement aux courbes de niveaux.

4. Dérivées partielles d'ordre supérieur

La notion de dérivées partielles d'ordre supérieur généralise celle de dérivée d'ordre supérieur (dérivée seconde, troisième,...) pour les fonction d'une variable.

Pour plus de clarté, on restreint la discussion qui vient aux fonctions de deux variables

$$(x, y) \mapsto f(x, y).$$

Etant donné une telle fonction ; supposons qu'elle soit différentiable sur un domaine $D(f)$; on dispose alors en chaque point (x, y) de ce domaine des deux dérivées partielles

$$\frac{\partial f}{\partial x}(x, y), \quad \frac{\partial f}{\partial y}(x, y),$$

et on dispose donc de deux fonctions de deux variables, notées $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sur $D(f)$

$$\begin{aligned} \frac{\partial f}{\partial x} : (x, y) &\mapsto \frac{\partial f}{\partial x}(x, y), \\ \frac{\partial f}{\partial y} : (x, y) &\mapsto \frac{\partial f}{\partial y}(x, y) \end{aligned}$$

on peut alors se demander si ces fonctions sont continues différentiables sur $D(f)$.

Supposons qu'elles soient différentiables, on a alors pour chacune d'elle deux dérivées partielles ce qui nous donne quatre dérivées partielles *d'ordre 2*

$$\frac{\partial f}{\partial x} \left(\frac{\partial f}{\partial x} \right), \quad \frac{\partial f}{\partial y} \left(\frac{\partial f}{\partial x} \right), \quad \frac{\partial f}{\partial x} \left(\frac{\partial f}{\partial y} \right), \quad \frac{\partial f}{\partial y} \left(\frac{\partial f}{\partial y} \right),$$

et à partir de ces quatre fonctions on peut éventuellement former (si chacune d'elle est différentiable) 8 dérivées partielles *d'ordre 3*. et ainsi de suite...

DÉFINITION 2.8. Une fonction de deux variables $f : D(f) \subset \mathbf{R}^2 \rightarrow \mathbf{R}$ est différentiable à l'ordre 1 sur un domaine $D(f)$ si elle est différentiable en tout point de ce domaine : elle admet alors deux dérivées partielles $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ qui sont des fonctions sur $D(f)$.

DÉFINITION 2.9. Une telle fonction est différentiable à l'ordre k si elle est différentiable et si ses deux dérivées partielles $\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}$ sont différentiables à l'ordre $k - 1$. elle admet alors 2^k dérivées partielles d'ordre k qui sont obtenues en appliquant à f k fois f une dérivation dans la direction x ou dans la direction y :

$$\left(\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\dots f\right)\right)\right), k \text{ fois.}$$

Une fonction k -fois différentiable pour tout entier $k \geq 1$, sera dite infiniment différentiable.

EXEMPLE 4.0.1. Soit $f(x, y) = x \cos(xy)$; cette fonction est infiniment différentiable sur \mathbf{R}^2 . On a

$$\frac{\partial f}{\partial x}(x, y) = \cos(xy) - xy \sin(xy), \quad \frac{\partial f}{\partial y}(x, y) = -x^2 \sin(xy)$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right)(x, y) = -2y \sin(xy) - xy^2 \cos(xy), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)(x, y) = -2x \sin(xy) - x^2 y \cos(xy)$$

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)(x, y) = -2x \sin(xy) - x^2 y \cos(xy), \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)(x, y) = -x^3 \cos(xy)$$

NOTATION. Si on effectue plusieurs fois (k fois) une opération de dérivation suivant la même variable (disons x), on notera cette opération

$$\frac{\partial^k}{\partial x^k}.$$

Ainsi

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial y^2};$$

de même on écrira

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}$$

ou encore

$$\frac{\partial}{\partial x}\left(\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right)\right) = \frac{\partial^3 f}{\partial^2 x \partial y}$$

4.1. Commutativité de l'ordre des dérivations. Etant donné une fonction f au moins deux fois différentiable : on peut calculer sa dérivée partielle en x puis sa dérivée partielle en y , obtenant la dérivée partielle d'ordre 2

$$\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) = \frac{\partial^2 f}{\partial y \partial x}$$

alternativement on pourrait commencer par dériver en y puis en x obtenant

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) = \frac{\partial^2 f}{\partial x \partial y}.$$

On remarque que dans l'exemple précédent ($f(x, y) = x \cos(xy)$) qu'on obtient la même fonction

$$\frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) = -2x \sin(xy) - x^2 y \cos(xy).$$

C'est en fait un phénomène général :

THÉORÈME 2.2. *Soit $f : D(f) \subset \mathbf{R}^2 \rightarrow \mathbf{R}$ une fonction différentiable à l'ordre 2. Supposons que ses dérivées partielles d'ordre 2,*

$$\frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x}, \quad \frac{\partial^2 f}{\partial y^2}$$

sont toutes continues (par exemple si f est différentiable à l'ordre 3) alors on a l'égalité

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$$

En d'autres termes pour une fonction suffisamment régulière, l'ordre dans lequel on effectue les opérations de dérivation partielle ne compte pas.

Généralisant ce théorème on voit que si f est différentiable à l'ordre au moins $k + 1$ ses dérivées partielles d'ordre k sont toutes de la forme

$$\frac{\partial^k f}{\partial^{k_1} x \partial^{k_2} y}, \quad k_1 + k_2 = k.$$

CHAPITRE 3

Etude des extrema d'une fonction

1. Extrema : Rappels sur les fonctions d'une variable

Dans cette section on veut généraliser à plusieurs variable la discussion suivante concernant les fonctions d'une variable :

Soit f une fonction définie sur un intervalle I de \mathbf{R} ; on désire connaître les points x de I où $f(x)$ prend une valeur maximale ou minimale (on veut déterminer les *extrema* de f). Pour cela

- On commence par calculer les valeurs de f aux extrémités de I , $f(a)$, $f(b)$ au moins quand ces valeurs sont définies.
- on étudie alors f à l'intérieur de I : dans l'intervalle "ouvert" $]a, b[$.

On suppose que f est 2 fois dérivable.

PROPOSITION 3.1. *Si f admet un extremum au point x dans $]a, b[$ alors*

$$f'(x) = 0.$$

Cette proposition nous amène à trouver les solutions de l'équation

$$f'(x) = 0.$$

Les solutions sont appelés point *critiques* ou *stationnaires*. On regarde alors la dérivée seconde en un tel point

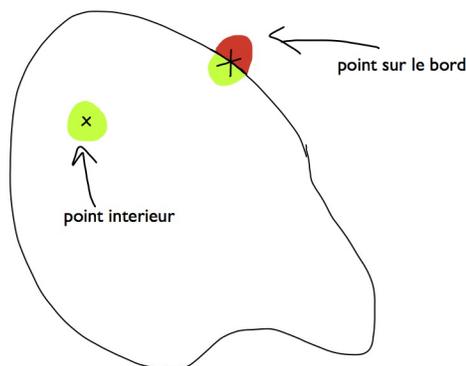
THÉORÈME 3.1. *Soit $x \in]a, b[$ tel que $f'(x) = 0$ alors, si*

- si $f''(x) < 0$, il existe un intervalle ouvert $I_x =]a_x, b_x[$ contenant x tel que f restreinte à I_x prend sa valeur maximale en x .
- si $f''(x) > 0$, il existe un intervalle ouvert $I_x =]a_x, b_x[$ contenant x tel que f restreinte à I_x prend sa valeur minimale en x .
- Si $f''(x) = 0$ plusieurs choses sont possible. En tout cas, on dit que f a un point d'inflexion en x .

Dans les deux premiers cas on dit que f admet un *extremum* (un *minimum* ou un *maximum*) local en x . Evidemment les extrema locaux sont des candidats à être des *extrema globaux* (sur I tout entier). A priori pour les point d'inflexions, tels que $f''(x) = 0$, on ne peut rien dire a priori et une étude plus approfondie est nécessaire :

EXEMPLE 1.0.1. les fonctions f suivantes admettent un point d'inflexion en $x = 0$

- $f(x) = x^4$: minimum en 0.
- $f(x) = -x^4$: maximum en 0
- $f(x) = x^3$ ni minimum ni maximum (pas même local), f est strictement croissante.



2. Cas des fonctions de deux variables

On va généraliser la discussion précédente aux fonction à deux variables. On se donne f définie sur un domaine D de \mathbf{R}^2 et on désire déterminer les $\vec{x} = (x, y)$ où $f(\vec{x})$ prend des valeurs extrêmes.

On va commencer par donner l'analogie des extrémités de I .

DÉFINITION 3.1. On définit l'intérieur de D , D° comme l'ensemble des \vec{x} dans D tels qu'il existe une boule $B(\vec{x}, r)$ de rayon > 0 entièrement contenue dans D . On définit le bord de D , ∂D comme étant le complémentaire de D° (les points de D qui ne sont pas dans l'intérieur).

Pour déterminer les extrema de f sur D , on procède ainsi

- (1) On étudie les extrema de f sur le bord ∂D : le bord est en général la réunion d'une courbe et de point isolés. L'étude de f le long du bord se ramène la plupart du temps à étudier les extrema d'une fonction d'une variable.
- (2) On étudie les extrema dans l'intérieur D° : on déterminera *a priori* des extrema locaux.
- (3) On compare les différent extrema qu'on a trouvé.

Comme on l'a dit la première étape se ramène à l'étude des variations d'une fonction d'une seule variable.

On passe donc à la deuxième étape : étudier les extrema de f contenus dans l'intérieur D° de D . Pour ce faire, on aura besoin de calculer les dérivées partielles de f et on supposera que f est différentiable à l'ordre au moins 3 sur D .

PROPOSITION 3.2. Soit $\vec{x}_0 = (x_0, y_0)$ un point de l'intérieur de D . Si \vec{x}_0 est un extremum de f sur D° alors

$$\nabla f(\vec{x}_0) = \left(\frac{\partial f}{\partial x}(\vec{x}_0), \frac{\partial f}{\partial y}(\vec{x}_0) \right) = (0, 0).$$

DÉFINITION 3.2. *Un point \vec{x}_0 tel que*

$$\nabla f(\vec{x}_0) = \left(\frac{\partial f}{\partial x}(\vec{x}_0), \frac{\partial f}{\partial y}(\vec{x}_0) \right) = (0, 0)$$

est appelé point critique pour f .

Preuve: Soit $\vec{x}_0 = (x_0, y_0)$ un point de l'intérieur et $B(\vec{x}_0, r)$, $r > 0$ une boule centrée en \vec{x}_0 et contenue dans D . Considérons la fonction d'une variable

$$f_1 : x \mapsto f(x, y_0);$$

f_1 est définie sur l'intervalle $]x_0 - r, x_0 + r[$ et par hypothèse admet un extremum en x_0 . On a donc

$$f_1'(x_0) = \frac{\partial f}{\partial x}(x_0, y_0) = 0.$$

De même, la fonction d'une variable

$$f_2 : y \mapsto f(x_0, y);$$

f_2 est définie sur l'intervalle $]y_0 - r, y_0 + r[$ et par hypothèse admet un extremum en y_0 . On a donc

$$f_2'(y_0) = \frac{\partial f}{\partial y}(x_0, y_0) = 0.$$

□

3. Le Hessian et la formule d'approximation quadratique

Soit $\vec{x}_0 \in D^\circ$, le *Hessian* de f en \vec{x}_0 est défini par la formule suivante

$$\text{Hess}(f)(\vec{x}_0) := \left(\frac{\partial^2 f}{\partial x \partial y} \right)^2(\vec{x}_0) - \frac{\partial^2 f}{\partial x^2} \frac{\partial^2 f}{\partial y^2}(\vec{x}_0).$$

On a alors le résultat suivant

THÉORÈME 3.2. *Soit f définie sur $B(\vec{x}_0, r)$, $r > 0$, différentiable à l'ordre 3 au moins. Supposons que \vec{x}_0 soit un point critique pour f alors si*

- Si $\text{Hess}(f)(\vec{x}_0) < 0$, et si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) < 0$ alors il existe $0 < r' < r$ tel que f restreinte à la boule $B(\vec{x}_0, r')$ prend sa valeur maximale en \vec{x}_0 . On dit que \vec{x}_0 est un maximum local pour f .
- Si $\text{Hess}(f)(\vec{x}_0) < 0$, et si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) > 0$, il existe $0 < r' < r$ tel que f restreinte à la boule $B(\vec{x}_0, r')$ prend sa valeur minimale en \vec{x}_0 . On dit que \vec{x}_0 est un minimum local pour f .
- Si $\text{Hess}(f)(\vec{x}_0) > 0$, f n'est ni un maximum, ni un minimum local en \vec{x}_0 . On dit que \vec{x}_0 est un point selle, ou un point col de f .
- Si $\text{Hess}(f)(\vec{x}_0) = 0$ plusieurs configurations sont possibles. En tous cas le point \vec{x}_0 sera dit dégénéré.