http://wiki.epfl.ch/mathgeo2

GSE Mathématiques II

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'un quantité associée à un tel système

Par exemple la fonction

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

decrit l'altitude d'un corps* soumis a l'attraction terreste ($g = 9{,}81\, m/s^{-2}$) au temps t laché d'une hauteur z_0.

* ponctuel(!)

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

Exemple : si on lance le corps considéré précédement 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 $t \longrightarrow (x(t), y(t), z(t))$ 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 réelles. Cependant un système physique général est souvent beaucoup plus complexe et dependent de nombreux paramètre qui agissent entre eux de manière compliquée.

Ainsi l'accélération gravitationnelle (la constante g) n'est constante qu'en <u>première approximation</u> : sa valeur absolue (le g) voire sa direction dependent du point (x, y, z) ou l'on se trouve

ainsi les équations du mouvement qui decrivent la trajectoire d'un corps lancée avec une vitesse initialle 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 méchanique 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'energie dit que la somme de l'énergie cinetique et de l'énergie potentielle reste constante :

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

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

$$= \frac{1}{2} m(v_{x_0}^2 + v_{y_0}^2 + v_{z_0}^2) + mgz_0$$

en effet $\left(\vec{x}(t), \vec{v}(t)\right) = \left(x_0 + v_{x_0} t, y_0 + v_{y_0} t, z_0 + v_{z_0} t - \frac{1}{2} g t^2, v_{x_0}, v_{y_0}, v_{z_0} - gt\right)$

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~~ dont les coordonnées vérifient

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

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 generalement 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 $(0, 0, -g)$)

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~~ sous-ensemble de $\mathbb{R}^6$ défini par

$$\frac{1}{2} m(\dot{x}^2 + \dot{y}^2 + \dot{z}^2) - m(g_1(x,y,z)x + g_2(x,y,z)y + g_3(x,y,z)z)$$

$$= Cst = \text{valeur en } (x_0, y_0, z_0, v_{x_0}, v_{y_0}, v_{z_0})$$

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

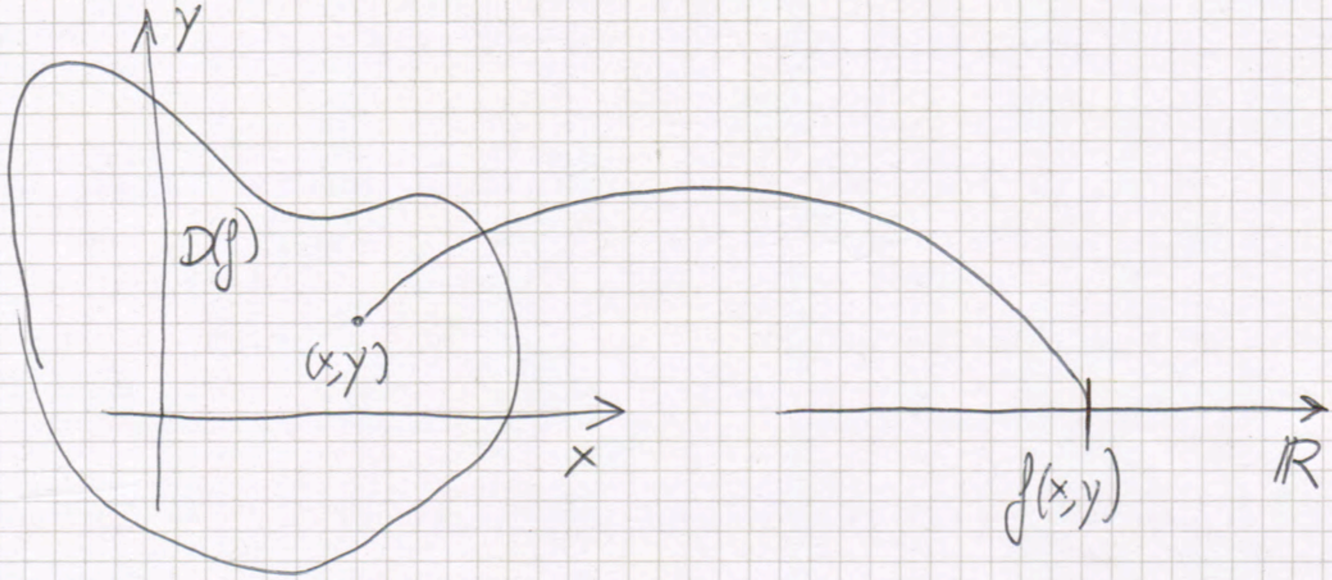
Def: Un espace de paramètres est un produit de copies de $\mathbb{R}$
$\underbrace{\mathbb{R}\times\mathbb{R}\times\ldots\times\mathbb{R}}_{n \text{ fois}} = \mathbb{R}^n$. Un point de cet espace (ou vecteur)
est repéré par ses coordonnées $(x_1, \ldots, x_n)$ x_i dans $\mathbb{R}$
On utilisera souvent une notation vectorielle
$\vec{x} = (x_1, \ldots, x_n)$.
Rappelons qu'on a certaines opérations sur les vecteurs
- somme : $\vec{x}+\vec{x}' = (x_1, x_n)+(x_1', , x_n') = (x_1+x_1', \ldots, x_n+x_n')$
- multiplication par un scalaire (homothétie):
si λ est dans $\mathbb{R}$
$$\lambda.\vec{x} = (\lambda x_1, \ldots, \lambda x_n) \text{ dans } \mathbb{R}^n$$
- produit scalaire
$$\vec{x}.\vec{x}' = \sum_{i=1} x_i x_i' = x_1x_1' + \ldots + x_n x_n' \text{ est un réel}$$
Notons que $\vec{x}.\vec{x} = \sum_{i=1}^{n} x_i^2$ est le carré de la longueur euclidienne du vecteur $\vec{x}$.

Def: Une fonction réelle (ou bien à valeurs réelles) f est une application qui à tout point d'un sous-ensemble $D(f)$ de $\mathbb{R}^n$ associe une valeur réelle $(x_1, \ldots, x_n) \longrightarrow f(x_1, \ldots, x_n)$
$D(f)$ est appelé domaine de définition de f.
Notations: On dit aussi que f est une fonction de n variables.
$$f : D(f) \longrightarrow \mathbb{R}$$
$$(x_1 \ldots, x_n) \longrightarrow f(x_1, \ldots, x_n)$$

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

$(x,y) \longrightarrow f(x,y)$ au lieu de $(x_1,x_2) \longrightarrow f(x_1,x_2)$

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



<u>Exemples</u> ① $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$(x,y) \longrightarrow \sqrt{x^2+y^2}$ (la longueur du vecteur (x,y))

② $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$(x,y) \longrightarrow 1$ $\quad$ fonction constante égale à 1

③ $f : \mathbb{R}^3-\{0\} \longrightarrow \mathbb{R}$

$(x,y,z) \longrightarrow \dfrac{1}{x^2+y^2+z^2}$ $\quad$ $D(f) = \mathbb{R}^3 - \{(0,0,0)\}$

~~④ $f(x_1,x_2,x_3,x_4) = \dfrac{1}{x_1x_2x_3x_4}$ $D(f) = \mathbb{R}^4 - \{0$~~

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 $\mathbb{R}^m$

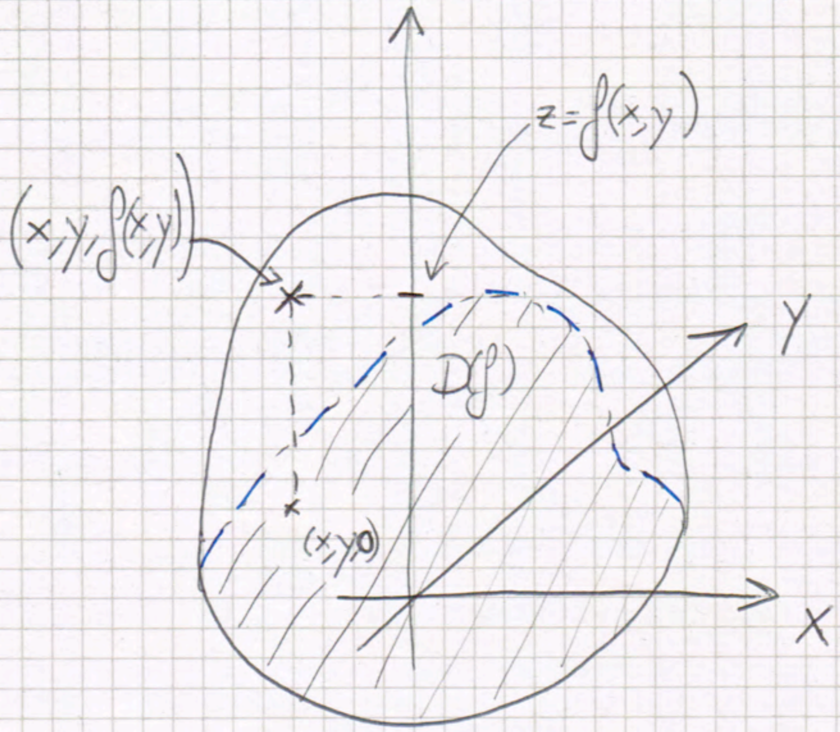
Le <u>graphe</u> d'une fonction de plusieurs variables

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

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

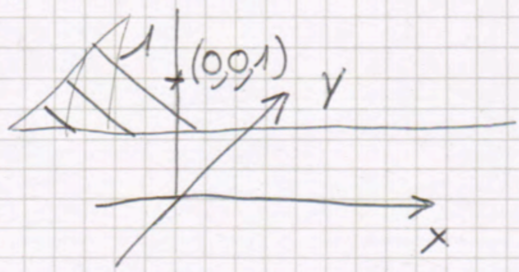
$(x_1, \dots, x_n)$ dans $D(f)$

Exemple : Si f est une fonction de 2 variables



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

le graphe est un plan

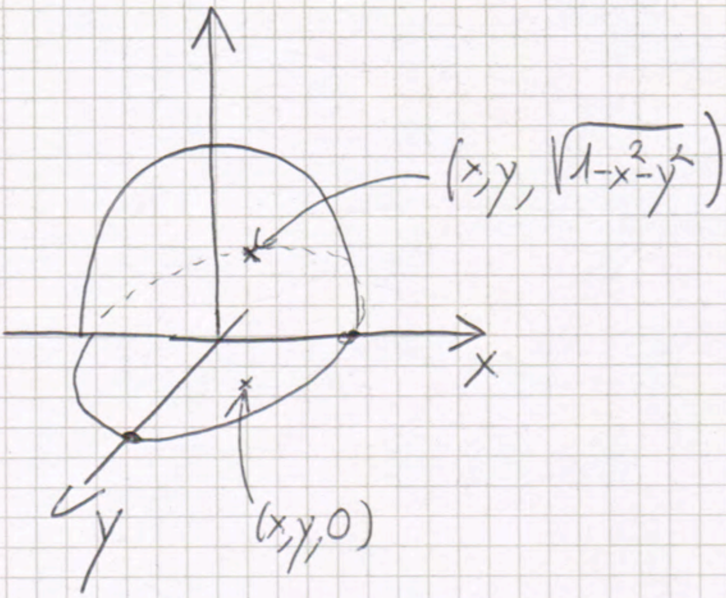


$f(x,y)=\sqrt{1-x^2-y^2}$ $\quad D(f)=\{(x,y)\ \ x^2+y^2\leq 1\}$

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

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

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



~~Courbes~~ Variétés de niveau

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

<u>Exemple</u> : $f(x,y,z) = x^2 + y^2 + z^2$

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 <u>graphe</u> 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 $\mathbb{R}^6$
$(x,y,z,\dot{x},\dot{y},\dot{z})$ $m>0$

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

alors un corps ponctuel de masse m soumis à l'attraction terreste 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

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

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 $\mathbf{R}^n$ appellés le graphe et les variétés de niveau de f

2.1. Graphe d'une fonction.

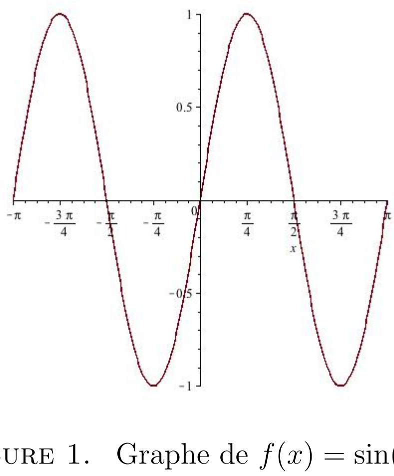


FIGURE 1. Graphe de $f(x) = \sin(2x)$

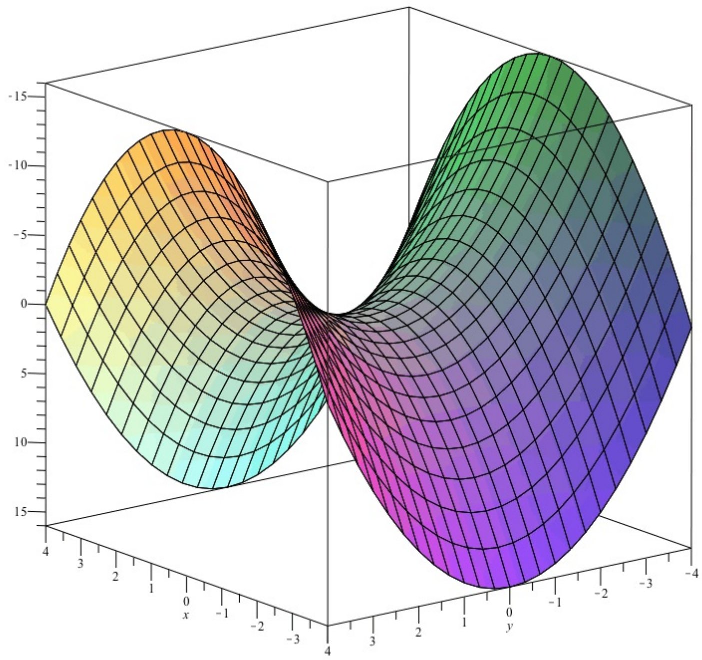


FIGURE 2. Graphe de $f(x,y) = x^2 - y^2$

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

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

– 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 represente les coordonnés de $\mathbf{R}^3$ par (x,y,z), le graphe de $f : (x,y) \to f(x,y)$ 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'un surface qui quand on la projette sur le plat horizontal (des (x,y)) redonne le domaine de définition D_f.
– Pour $n \geqslant 3$ le graphe de f est un objet de $\mathbf{R}^{n+1}$ qui est difficile de representer sur un tableau ou un ecran.

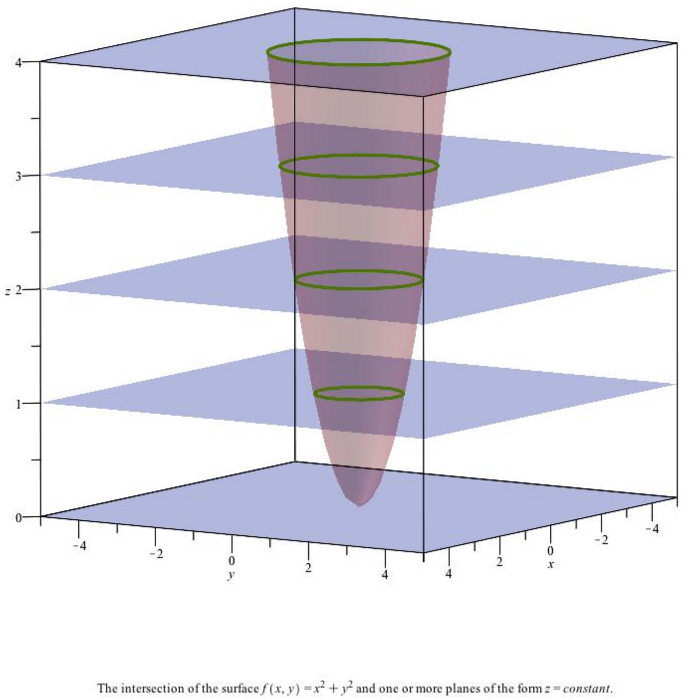


FIGURE 3. Courbes de niveau de $f(x,y) = x^2 + y^2$

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

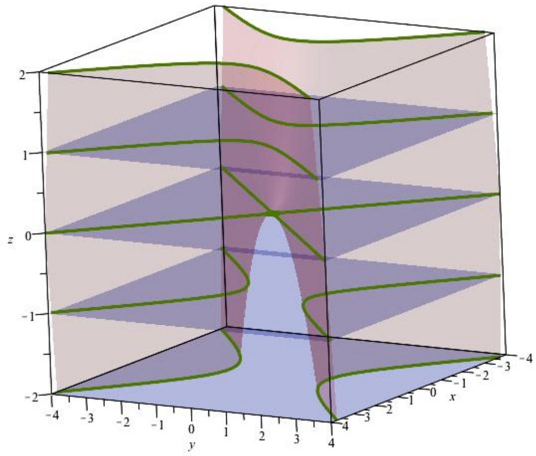
DÉFINITION 1.2. *Soit $f : D_f \subset \mathbf{R}^n \to \mathbf{R}$ une fonction et C une constante, la variété de niveau C est le sous-ensemble de $\mathbf{R}^n$*

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

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

3. Approximation et continuité

Dans ce cours on va chercher à décrire une fonction $\vec{x} \to f(\vec{x})$ quand la variable $\vec{x}$ est "proche" d'un point $\vec{x}_0 = (x_{0,1}, \cdots, x_{0,n})$. 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 $z = constant$.

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, $f : \vec{x} \to Cst$ 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

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

avec $(a_1, \cdots, a_n) \in \mathbf{R}^n$ un vecteur fixe. Les fonctions linéaires satisfont
- $L(\vec{0}) = 0$,
- $L(\vec{x} + \vec{x}') = L(\vec{x}) + L(\vec{x}')$,
- $L(\lambda.\vec{x}) = \lambda L(\vec{x})$.

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

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

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

La somme d'un 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

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

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

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

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 : $\|\vec{h}\|_2 = (h_1^2 + \cdots + h_n^2)^{1/2}$,

(2) La norme L_1 :$\|\vec{h}\|_1 = |h_1| + \cdots + |h_n|$,

(3) La norme sup : $\|\vec{h}\|_\infty = \max_{i=1,\cdots,n} |h_i|$.

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

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

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 $\vec{x}$, $\|\vec{x}\| \geqslant 0$*

(2) *(détection du vecteur nul) $\|\vec{x}\| = 0$ si et seulement si $\vec{x} = \vec{0}$,*

(3) *(homogénéité) $\|\lambda.\vec{x}\| = |\lambda|\|\vec{x}\|$.*

(4) *(inégalité triangulaire) $\|\vec{x} + \vec{x}'\| \leqslant \|\vec{x}\| + \|\vec{x}'\|$.*

3.3. Continuité. On rappelle que

DÉFINITION 1.4. *Une fonction $f : D_f \subset \mathbf{R} \to \mathbf{R}$ 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 á 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

$$\lim_{x \to x_0} f(x) = f(x_0),\ \textit{ou}\ f(x) \to f(x_0)\ \textit{quand}\ x \to x_0.$$

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*

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

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

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

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*

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

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

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

DÉFINITION 1.6. *Soit* $f : D_f \in \mathbf{R}^n \mapsto \mathbf{R}$ *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*

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

si pour tout 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* $f : D_f \subset \mathbf{R}^n \to \mathbf{R}$ *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 á dire, si pour tout 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

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

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 eq́uivalente 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 dfinie pour tout* $\vec{h}$ *suffisament proche du vecteur nul* $\vec{0}$ *qui tend vers* 0 *quand* $\vec{h}$ *tend vers* $\vec{0}$ *:*

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

On a alors la définition équivalente suivante de la continuité :

Définition 1.8. *Une fonction* $f : D_f \subset \mathbf{R}^n \to \mathbf{R}$ *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'ecrire sous la forme*

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

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

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

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* $i = 1, ..., n$ *les fonctions coordonnées*

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

  sont continues en tout point de $\mathbf{R}^n$*.*
- *Si* f *et* g *sont continues en* $\vec{x}_0$ *alors* $f + g : \vec{x} \to f(\vec{x}) + g(\vec{x})$ *est continue en* $\vec{x}_0$*.*
- *Si* f *et* g *sont continues en* $\vec{x}_0$ *alors* $f.g : \vec{x} \to f(\vec{x})g(\vec{x})$ *est continue en* $\vec{x}_0$*.*
- *Si* f *et* g *sont continues en* $\vec{x}_0$ *alors* $f/g : \vec{x} \to f(\vec{x})/g(\vec{x})$ *est continue en* $\vec{x}_0$ *si* $g(\vec{x}_0) \neq 0$*.*
- *Si* f *est continue en* $\vec{x}_0$ *et* $g : \mathbf{R} \to \mathbf{R}$ *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

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

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 ($(x, y) \to xy$) sur $\mathbf{R}^2$et d'une fonction d'une variable également continue sur $\mathbf{R}$ (sin) ; sont dénominateur est continu et ne s'annule jamias sur $\mathbf{R}^2$ (car $x^2 + y^2 + 1 \geqslant 1$).

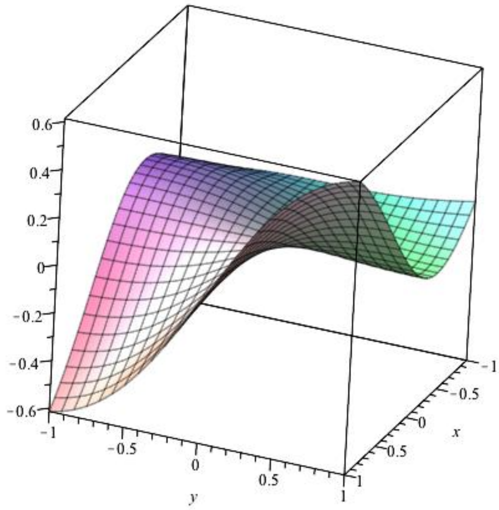


FIGURE 5. graphe de $f(x,y) = \frac{\sin(xy)+x^2y}{x^2+y^2+1}$

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

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

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 $n \to +\infty$ alors que le premier a pour limite

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

Et donc

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

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

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

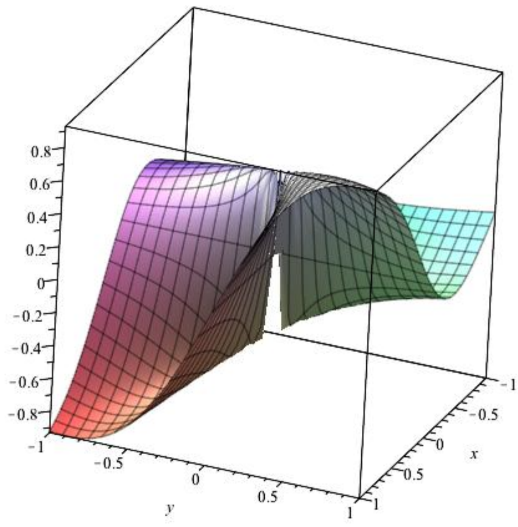


FIGURE 6. graphe de $f(x,y) = \frac{\sin(xy)+x^2y}{x^2+y^2}$

car alors

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

CHAPITRE 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 phénomémes trés différents. Dasn 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 dérivabilité pour les fonctions d'une variable. Cependant pour énoncer cette généralisation il est utile de reformuler la notion de dérivabilité de maniére adéquate.

1.1. Rappels sur la dérivabilité. Traditionellement, on dit qu'une fonction $f(x)$ d'une variable est dérivale 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

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

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 elee est continue, ainsi la notion de dérivabilité est un notion de régularité plus forte que la continuité. Pour passer aux fonctions de plusieurs variables, il sera commode de formuler la dérivabiiltée sous une forme équivalente : l'expression

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

est équivalente[1] à l'identité

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

avec $\varepsilon(h) \to 0$ quand $h \to 0$. En particulier, cette formulation montre immédiatement que f est continue en x_0 ; elle s'intérpréte 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)

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

s'interpréte 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 $h = 0$, alrs 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 fonction 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 :

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

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

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

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

et telle qu'on ait l'égalité ($\|.\|$ 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 differentiable 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 terms 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 term 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

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

on a

$$f(x,y) = f(x_0,y_0) + \varepsilon_1(x-x_0, y-y_0),\ \varepsilon_1(u,v) = a_1 u + a_2 v + \|(u,v)\|\varepsilon(u,v);$$

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

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

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

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

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

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

$$f(\vec{x}_0 + \vec{h}) = f(\vec{x}_0) + a_1 h_1 + \cdots + a_n h_n + \|\vec{h}\|\varepsilon(\vec{h}).$$

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, considéront la fonction

$$\varepsilon_1 : \vec{h} = (h_1, \ldots, h_n) \mapsto a_1 h_1 + \cdots + a_n h_n + \|\vec{h}\|\varepsilon(\vec{h});$$

alors $\varepsilon_1(\vec{h})$ tend vers 0 quand $\vec{h} \to 0$ 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 maniére plus compacte en terme de produit scalaire : si

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

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

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

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

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

et donc

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

ou encore

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

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 définit 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-iéme 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)

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

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 replaçant la i-iéme coordonnée par x.

DÉFINITION 2.3. *Le coefficient a_i (ie la dérivéee 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

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

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 considére 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. *Régles de différentiation.* On dispose comme pour les fonctions d'une variable des critéres 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 $\dfrac{f}{g}$ 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)

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

(4)

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

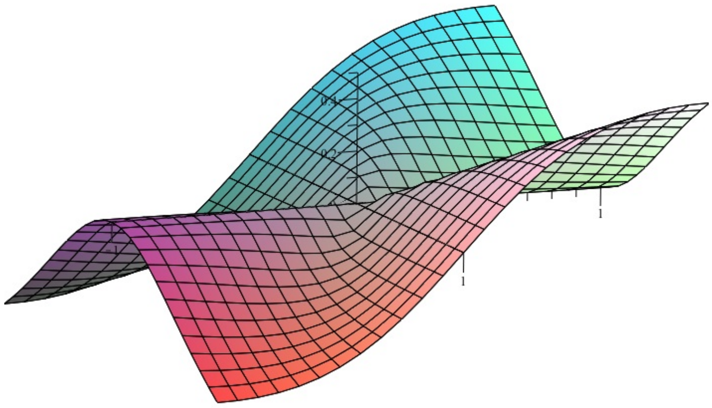


FIGURE 1. $f(x,y) = x^2y/(x^2+y^2)$

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

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

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

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

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

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

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

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

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

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

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 un hypothése de régularité des dérivées partielles :

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

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

existent et que les fonctions

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

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

1.5. Différentielle ; Gradient. Le vecteur des dérivées partielles

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

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

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

On trouve aussi comme notation du gradient

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

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

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

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

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

On écrira de maniéres équivalentes

$$f(\vec{x}) = f(\vec{x}_0) + df(\vec{x}_0)(\vec{x} - \vec{x}_0) + \|\vec{x} - \vec{x}_0\|\varepsilon(\vec{x} - \vec{x}_0),$$

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

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

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

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'interpréte 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 $\vec{x} = (x_1, \cdots, x_n)$ et $\vec{y} = (y_1, \cdots, y_n)$ deux vecteurs de $\mathbf{R}^n$. Les nombres réels $x_i, x_i', \ i = 1, \cdots, n$ sont les *coordonnées* de ces vecteurs ; on rappelle les opérations et notations suivantes

- **Le vecteur nul** : $\vec{0} = (0, \cdots, 0) \in \mathbf{R}^n$
- **Somme** :
$$\vec{x} + \vec{y} = (x_1 + y_1, \cdots, x_n + y_n) \in \mathbf{R}^n,$$
- **Multiplication par un scalaire** : pour $\lambda \in \mathbf{R}$,
$$\lambda.\vec{x} = (\lambda x_1, \cdots, \lambda x_n).$$
- **Produite 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
$$|\vec{x}.\vec{y}| \leqslant \|\vec{x}\|_2 \|\vec{y}\|_2$$
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 défini leur l'angle $\theta \in [0, \pi]$ comme l'unique solution (dans $[0, \pi]$) de léquation
$$\vec{x}.\vec{y} = \cos(\theta)\|\vec{x}\|_2\|\vec{y}\|_2$$
(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* $\vec{a} = (a_1, \cdots, a_n)$, $\vec{x}_0 = (x_{0,1}, \cdots, x_{0,n}) \in \mathbf{R}^n$ *deux vecteurs ; on suppose que* $\vec{a} \neq 0$*, 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*

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

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 $\vec{h} = (h_1, \cdots, h_n)$ tels que $\vec{a}.\vec{h} = a_1 h_1 + \cdots + a_n.h_n = 0$),

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

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

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

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

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

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

On utilisera donc également la notation

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

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

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

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 equation

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

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

$$\vec{x} \mapsto f(\vec{x}_0) + \vec{a}.(\vec{x} - \vec{x}_0) \tag{2.1}$$

avec

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

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 $(x_1, \ldots, x_n, x_{n+1})$

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

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

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

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 á 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 réecrit

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

C'est donc l'hyperplan associé au vecteur normal

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

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 $(\vec{x}_0, f(\vec{x}_0)) \in \mathbf{R}^{n+1}$.*

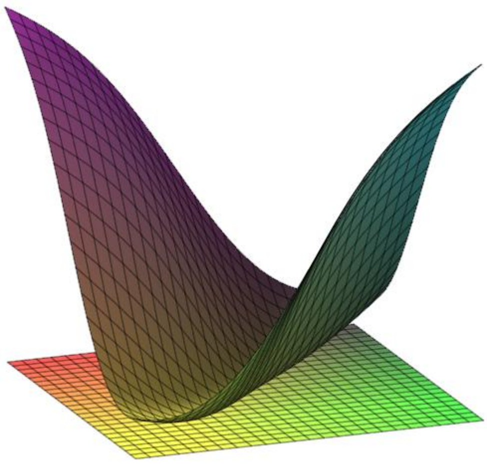


FIGURE 2. Graphe de $f(x,y) = ((x-y)^2 + x^4 + y^4)/(1 + 10(x^2 + y^2))$ et son plan tangent en $(0,0,0)$

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

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

soit en simplifiant

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

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

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

quand $\vec{x} \to \vec{x}_0$: 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 preseque 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 $f : \mathbf{R}^n \to \mathbf{R}$ différentiable en $\vec{x}_0$ et soit $C = f(\vec{x}_0)$. Si $\nabla(f)(\vec{x}_0) \neq \vec{0}$ alors l'hyperlan de $\mathbf{R}^n$ donné par l'équation*

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

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

$$x^2 + y^2 + z^2 = R^2, \ R = 6378Km$$

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 $\mathbf{R}^{n+1}$,

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

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

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

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

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

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.

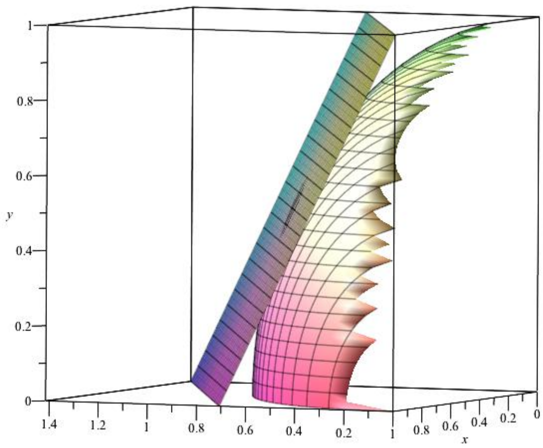


FIGURE 3. La sphere et un plan tangent.

3. Dérivée directionnelle d'une fonction

La i-éme 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 deplace à partir de $\vec{x}_0$ parallélelement 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 $\vec{e} = (e_1, \ldots, e_n)$ 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*

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

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

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

et donc

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

Ainsi

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

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

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

Applquant cela à la décomposition

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

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

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

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 $D_{\vec{e}}f(\vec{x}_0) = \nabla f(\vec{x}_0).\vec{e}$ mesure donc *l'amplitude* de la variation de la fonction f quand la variable $\vec{x}$ se deplace depuis $\vec{x}_0$ dans la direction $\vec{e}$. Par exemple cette amplitude est *nulle* (f varie trés peu) si

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

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

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

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

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

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

$$|\nabla f(\vec{x}_0).\vec{e}| \leqslant \|\nabla f(\vec{x}_0)\| \|\vec{e}\| = \|\nabla f(\vec{x}_0)\|$$

avec égalité si et seulement si $\vec{e}$ et $\nabla f(\vec{x}_0)$ son 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 parallélement 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, troisiéme,...) 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 une domaine $D(f)$; on dispose alors en chaque point (x, y) de ce domaine des deux dérivées partielles

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

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 démander si ces fonctions sont continues différentiables sur $D(f)$.

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

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

et à patir de ces quaters 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 $f : D(f) \subset \mathbf{R}^2 \to \mathbf{R}$ 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'rodre k qui sont obtenues en appliquant à f k fois f une dérivation dans la direction x ou dans la direction y :*

$$(\frac{\partial}{\partial x}(\frac{\partial}{\partial x}(\frac{\partial}{\partial y}...f)),\ k\ fois.$$

Une function k-fois différentiable pour tout entier $k \geqslant 1$, 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

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

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

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

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

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

de même on écrira

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

ou encore

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

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

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

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

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

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

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

C'est en fait un phénoméme général :

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

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

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'autre termes pour une fonction suffisament réguliére, l'ordre dans lequel on effectue les opérations de dérivation partielle ne compte pas.

Généralisant ce théoréme 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

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

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 définit 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 vent 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 améne 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

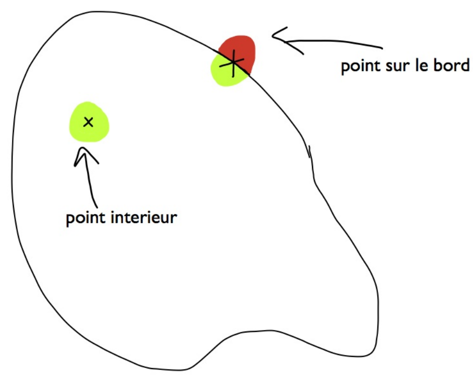
THÉORÈME 3.1. *Soit $x \in]a, b[$ tel que $f'(x) = 0$ alors,si*

- *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. Evidement 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 nécéssaire :

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'analogue 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 entiérement 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 procéde 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*

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

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

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

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 hypothése 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 hypothése 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 Hessien et la formule d'approximation quadratique

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

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

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* $\mathrm{Hess}(f)(\vec{x}_0) < 0$*, 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* $\mathrm{Hess}(f)(\vec{x}_0) < 0$*, 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* $\mathrm{Hess}(f)(\vec{x}_0) > 0$*,* 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* $\mathrm{Hess}(f)(\vec{x}_0) = 0$ *plusieurs configurations sont possibles. En tous cas le point* $\vec{x}_0$ *sera dit dégénéré.*

Remarque 3.1. Dans le cas où $\mathrm{Hess}(f)(\vec{x}_0) < 0$, pourquoi considérer $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$ plutôt que $\frac{\partial^2 f}{\partial y^2}(\vec{x}_0)$? En fait cela ne fait aucune différence : si

$$\mathrm{Hess}(f)(\vec{x}_0) := (\frac{\partial^2 f}{\partial x \partial y})^2(\vec{x}_0) - \frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}(\vec{x}_0) < 0$$

alors nécéssairement

$$\frac{\partial^2 f}{\partial x^2}\frac{\partial^2 f}{\partial y^2}(\vec{x}_0) > 0$$

et les dérivées d'ordre 2 en x et en y sont les même.

3.1. Exemples. Dans cette section on va donner des exemples caractéristiques des situations précédentes.

3.1.1. *Cas d'un minimum local.* Soit

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

On a

$$\nabla f(x,y) = (2x, 2y)$$

et donc le seul point critique est le point $(0,0)$ et $f(0,0) = 0$. On a

$$\frac{\partial^2 f}{\partial^2 x}(x,y) = 2, \frac{\partial^2 f}{\partial x \partial y}(x,y) = 0, \ \frac{\partial^2 f}{\partial^2 y}(x,y) = 2$$

et donc

$$\mathrm{Hess}(f)(0,0) = -4.$$

On a $\frac{\partial^2 f}{\partial^2 x}(0,0) = 2 > 0$ donc f a un minimum local en $(0,0)$: c'est même un minimum absolu et il est facile de voir cela directement : pour tout (x,y)

$$f(x,y) = x^2 + y^2 \geqslant 0 = f(0,0).$$

3.1.2. *Cas d'un maximum local.* Il suffit de considérer l'opposé du cas précédent Soit

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

La discussion précédente s'applique et on voit bien que $(0,0)$ est un maximum absolu :

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

3.1.3. *Cas d'un point col.* Soit

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

On a

$$\nabla f(x,y) = (2x, -2y)$$

et donc le seul point critique est le point $(0,0)$ et $f(0,0) = 0$. On a

$$\frac{\partial^2 f}{\partial^2 x}(x,y) = 2, \frac{\partial^2 f}{\partial x \partial y}(x,y) = 0, \ \frac{\partial^2 f}{\partial^2 y}(x,y) = -2$$

et donc

$$\mathrm{Hess}(f)(0,0) = 4 > 0.$$

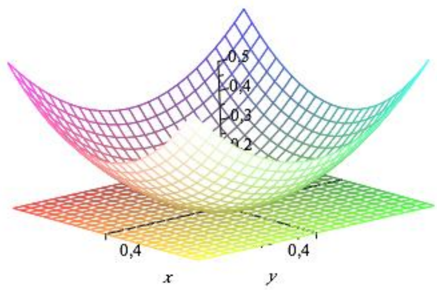


FIGURE 1. Minimum $f(x,y) = x^2 + y^2$

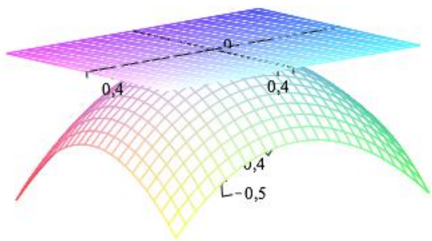


FIGURE 2. Maximum $f(x,y) = -x^2 - y^2$

On est dans le cas d'un point col : si $y = x$ ou bien $y = -x$, $f(x,y) = 0$, la fonction f prend donc la valeur $f(0,0)$ le long des deux droites d'équations

$$x - y = 0 \text{ ou bien } x + y = 0$$

et

$$f(x,y) > 0 \text{ si } x - y \text{ et } x + y \text{ sont de même signe}$$

$$f(x,y) < 0 \text{ si } x - y \text{ et } x + y \text{ sont de signe oppposé.}$$

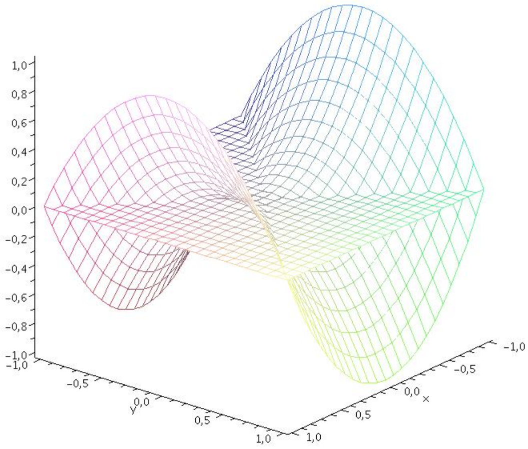


FIGURE 3. Col $f(x,y)=x^2-y^2$

3.1.4. *Cas d'un point dégénéré.*

$$f(x,y)=x^2+y^3.$$

On a

$$\nabla f(x,y)=(2x,3y^2)$$

et donc le seul point critique est le point $(0,0)$ et $f(0,0)=0$.On a

$$\frac{\partial^2 f}{\partial^2 x}(x,y)=2,\ \frac{\partial^2 f}{\partial x\partial y}(x,y)=0,\ \frac{\partial^2 f}{\partial^2 y}(x,y)=6y$$

et donc

$$\mathrm{Hess}(f)(0,0)=0.$$

Le point est donc dégénéré et comme le montre la figure, ce n'est ni un maximum, ni un minimum, ni un col. En fait, dans le cas des points dégénérés, on peut obtenir toute une variété de situations différentes. Dans ce cours, on ne les classifiera pas de maniére systématique.

3.2. La formule d'approximation quadratique. On va donner une idée de la preuve de ce théorème : pour cela nous aurons besoin de la formule suivante que nous admettrons

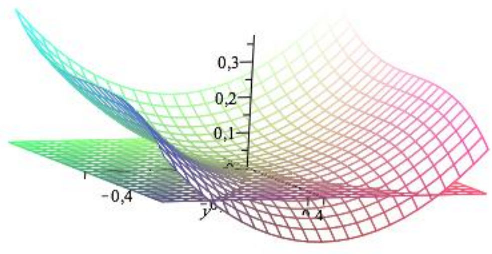


FIGURE 4. Point dégénéré $f(x,y) = x^2 + y^3$.

THÉORÈME 3.3 (Formule d'approximation quadratique). *Soit f une fonction différentiable à l'ordre 3 sur une boule $B(\vec{x}_0, r)$. Ecrivons*

$$\vec{x}_0 = (x_0, y_0),\ (x,y) = \vec{x}_0 + \vec{h} = (x_0 + h_1, y_0 + h_2)$$

$$\begin{aligned} f(x,y) = f(\vec{x}_0 + \vec{h}) &= f(\vec{x}_0) + \frac{\partial f}{\partial x}(\vec{x}_0).h_1 + \frac{\partial f}{\partial y}(\vec{x}_0).h_2 \\ &+ \frac{1}{2}\Big(\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0).h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0).h_1 h_2 + \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0).h_2^2\Big) + (h_1^2 + h_2^2)\varepsilon(\vec{h}) \end{aligned}$$

En particulier si $\vec{x}_0$ est un point critique, la formule devient

$$\begin{aligned} f(x,y) &= f(x_0 + h_1, y_0 + h_2) \\ &= f(\vec{x}_0) + \frac{1}{2}\Big(\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0).h_1^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0).h_1 h_2 + \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0).h_2^2\Big) + (h_1^2 + h_2^2)\varepsilon(\vec{h}) \\ &= f(\vec{x}_0) + Q(h_1, h_2) + (h_1^2 + h_2^2)\varepsilon(\vec{h}) \end{aligned}$$

avec

$$Q(U,V) = \frac{1}{2}\Big(\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0).U^2 + 2\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0).UV + \frac{\partial^2 f}{\partial^2 y}(\vec{x}_0).V^2\Big) = aU^2 + bUV + cV^2$$

avec

$$a = \frac{1}{2}\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0),\ b = \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0),\ c = \frac{1}{2}\frac{\partial^2 f}{\partial^2 y}(\vec{x}_0).$$

Considérons le polynôme quadratique

$$P(X) = aX^2 + bX + c;$$

mettant V en facteur, on a

$$Q(U, V) = V^2(a(U/V)^2 + b(U/V) + c) = V^2 P(U/V)$$

et donc si $h_2 \neq 0$, on a

$$f(x_0 + h_1, y_0 + h_2) = f(\vec{x}_0) + h_2^2 P(h_1/h_2) + (h_1^2 + h_2^2)\varepsilon(\vec{h});$$

cela nous raméne à étudier le signe de $P(X)$. Le discriminant de ce polynôme est $\Delta = b^2 - 4ac$ et vaut précisement

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

4. Extrema locaux

Supposons que le Héssien est négatif ; comme

$$\Delta = b^2 - 4ac < 0, \text{ on voit que } a, c \neq 0 \text{ et que } ac > 0\,; \text{ de plus}$$

le polynôme $P(X)$ ne s'annule pas sur $\mathbf{R}$ et reste de signe constant. Ce signe est donné par le signe du premier coefficient a c.a.d. le signe de $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$.

Supposons alors que $a = \frac{\partial^2 f}{\partial x^2}(\vec{x}_0) < 0$, alors pour $h_2 \neq 0$, la partie quadratique vaut

$$Q(h_1, h_2) = h_2 P(h_1/h_2) < 0$$

et si $h_2 = 0$, la partie quadratique vaut

$$ah_1^2 \leqslant 0$$

avec égalité ssi $h_1 = 0$. Ainsi compte-tenu du fait que le terme $(h_1^2 + h_2^2)\varepsilon(\vec{h})$ est un terme d'erreur quand (h_1, h_2) est proche de zero, on voit que

$$f(x, y) - f(\vec{x}_0) = ah_1^2 + bh_1h_2 + ch_2^2 + \text{Erreur}$$

a tendance à être négatif quand (x, y) est proche de (x_0, y_0) : $f(x_0, y_0)$ est maximal au voisinage de $\vec{x}_0$: $\vec{x}_0$ est un maximum local.

Le même raisonnement s'applique quand $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) > 0$ et fournit un minimum local.

5. Etude du point col

Supposons que le Héssien est positif.

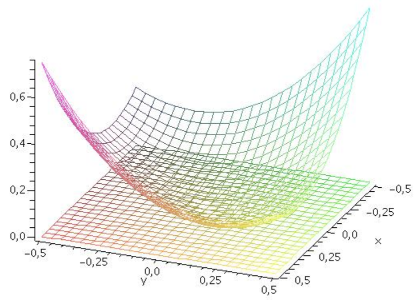


FIGURE 5. Minimum $f(x,y) = x^2 - xy + y^2$

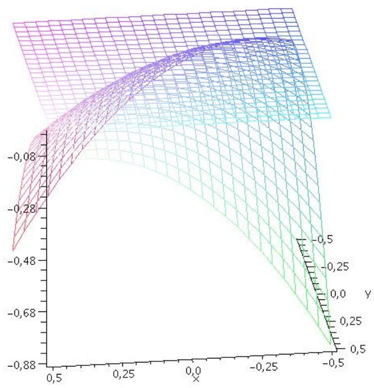


FIGURE 6. maximum $f(x,y) = -(x^2 - 3xy/2 + y^2)$

5.1. Cas où P est de degré 2. Supposons d'abord que $\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \neq 0$ (ie. P est un "vrai" polynôme de degré 2) alors le polynôme $P(X)$ a deux racines réelles distinctes

$$X_+ = \frac{-b+\sqrt{\Delta}}{2a} = \frac{-\frac{\partial^2 f}{\partial x \partial y} + \sqrt{\mathrm{Hess}(f)}}{\frac{\partial^2 f}{\partial^2 x}},$$

$$X_- = \frac{-b-\sqrt{\Delta}}{2a} = \frac{-\frac{\partial^2 f}{\partial x \partial y} - \sqrt{\mathrm{Hess}(f)}}{\frac{\partial^2 f}{\partial^2 x}}$$

et s'écrit

$$P(X) = a(X - X_+)(X - X_-).$$

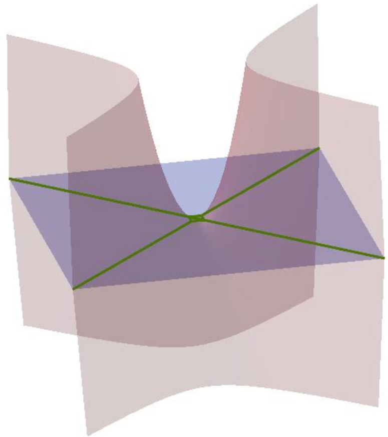


FIGURE 7. Un point col et ses droites isotropes

On en déduit que le terme quadratique vaut

$$Q(h_1,h_2) = h_2^2 P(h_1/h_2) = a(h_1 - X_+ h_2)(h_1 - X_- h_2) = \frac{1}{2}\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0)(h_1 - X_+ h_2)(h_1 - X_- h_2);$$

Ainsi on voit que la différence

$$f(x,y) - f(\vec{x}_0) = \frac{1}{2}\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0)(h_1 - X_+ h_2)(h_1 - X_- h_2) + \text{Erreur}.$$

à tendance à être très petite quand (h_1,h_2) est proche de $(0,0)$ et varie le long d'une des deux droites d'équation

$$h_1 - X_+ h_2 = 0,\ h_1 - X_- h_2 = 0;$$

en revanche quand (h_1,h_2) est en dehors de ces droites et proche de $(0,0)$, la différence change de signe en fonction du quadrant (définit par ces deux droites) dans lequel (h_1,h_2) se trouve. Rappelons que $h_1 = x - x_0$, $h_2 = y - y_0$ et les équations de ces droites s'écrivent

$$x - x_0 - X_+(y - y_0) = 0,\ x - x_0 - X_-(y - y_0) = 0.$$

5.2. Cas où P est de degré 1. Si maintenant $a = \frac{1}{2}\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) = 0$, P est un polynôme de degré 1 et la formule précédente s'écrit

$$\begin{aligned} f(x,y) - f(\vec{x}_0) &= b.h_1h_2 + c.h_2^2 + \text{Erreur} \\ &= \frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0)(h_1 - Xh_2)h_2 + \text{Erreur}, \quad X = -\frac{1}{2}\frac{\partial^2 f}{\partial^2 y}(\vec{x}_0)/\frac{\partial^2 f}{\partial x\partial y}(\vec{x}_0). \end{aligned}$$

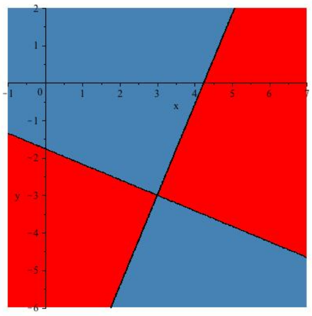


FIGURE 8. Un point col et ses droites isotropes vue du dessus

A nouveau on obtient deux droites caractéristiques, le long desquelles $f(x, y) - f(\vec{x}_0)$ est très petit, et en dehors desquelles le signe de la différence d'équation

$$x - x_0 - X(y - y_0),\ y - y_0 = 0.$$

DÉFINITION 3.3. *Supposons que* $\vec{x}_0 = (x_0, y_0)$ *soit un point col : ie.* $\nabla f(\vec{x}_0) = 0$, $\mathrm{Hess}(f)(\vec{x}_0) > 0$. *Les droites d'équation*

$$x - x_0 - X_+(y - y_0) = 0,\ x - x_0 - X_-(y - y_0) = 0$$

(si $\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \neq 0$*) ou d'équation*

$$x - x_0 - X(y - y_0),\ y - y_0 = 0$$

(si $\frac{\partial^2 f}{\partial^2 x}(\vec{x}_0) \neq 0$*) sont appellées les* droites caractéristiques (ou isotropes) associées au point col $\vec{x}_0$.

6. Cas des points dégénérés (facultatif)

Si $\mathrm{Hess}(f)(\vec{x}_0) = 0$ le discriminant de P est nul.

Il y a plusieurs possibilités : si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) \neq 0$ alors P est un "vrai" polynôme de degré 2 et il admet une racine double $X_0 = -b/2a$. Ainsi

$$f(x, y) - f(\vec{x}_0) = \frac{1}{2}(h_1 - X_0 h_2)^2 + \mathrm{Erreur},\ X_0 = -\frac{\partial^2 f}{\partial x \partial y}(\vec{x}_0) / \frac{\partial^2 f}{\partial^2 x}(\vec{x}_0)$$

Ainsi quand (x, y) varie au voisinage de $\vec{x}_0$ le long de la droite d'équation

$$x - x_0 - X_0(y - y_0) = 0$$

la différence $f(x, y) - f(\vec{x}_0)$ a tendance à être très petite. En dehors de cette droite (encore appelée droite caractéristique ou isotrope) on voit que au voisinage de $\vec{x}_0$, cette différence à tendance à être du même signe que $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0)$. Ainsi le graphe de f ressemble prés de $\vec{x}_0$, soit à une "ligne de créte" (si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) < 0$), soit à une "crevasse" (si $\frac{\partial^2 f}{\partial x^2}(\vec{x}_0) > 0$). Il y a bien d'autre configurations possibles : P peut être un polynôme

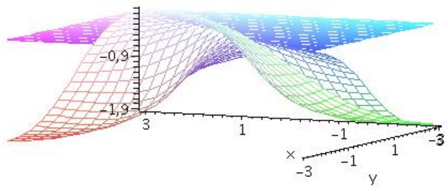


Figure 9. Créte $f(x, y) = -(x - y)^2/(1 + x^2 + y^2)$

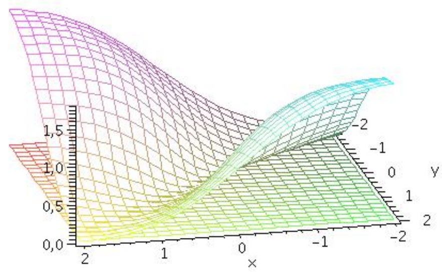


Figure 10. Crevasse $f(x, y) = (x - y)^2/(1 + x^2 + y^2)$

constant, nul ou non nul. Dans le dernier cas on obtient a nouveau une droite isotrope (verticale) d'équation

$$y - y_0 = 0,$$

et une configuration de ligne de créte ou de crevasse.

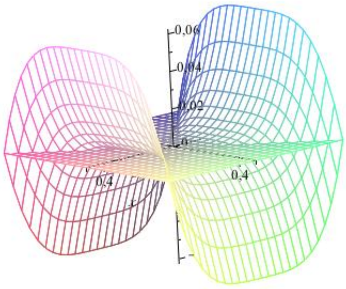


FIGURE 11. Col $f(x,y) = x^4 - y^4$

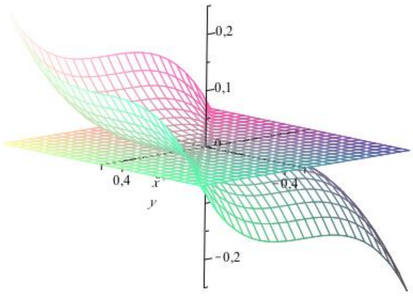


FIGURE 12. Point dégénéré $f(x,y) = x^3 + y^3$

Enfin si P est le polynôem nul, la formula d'approximation quadratique prend la forme

$$f(x,y) = f(x_0 + h_1, y_0 + h_2) = f(\vec{x}_0) + (h_1^2 + h_2^2)\varepsilon(\vec{h})$$

et il est nécéssaire de regarder de plus prés la forme du terme d'erreur $(h_1^2 + h_2^2)\varepsilon(\vec{h})$: on peut obtenir un extrema local ($f(x,y) = x^4 + y^4$), un point col ($f(x,y) = x^4 - y^4$) ou quelque chose de tout à fait différent.

CHAPITRE 4

Courbes paramétrées

1. Courbe dépendant d'un paramètre

Une courbe paramétrée (dans $\mathbf{R}^3$) est une fonction d'un intervalle I de $\mathbf{R}$ à valeurs dans $\mathbf{R}^3$

$$\varphi : \begin{array}{ccc} I = (a,b) \subset \mathbf{R} & \mapsto & \mathbf{R}^3 \\ t & \mapsto & \varphi(t) = (x(t), y(t), z(t)) \end{array}$$

En d'autre termes a tout point t de l'intervalle I on associe un point repéré par ses trois coordonnées $x(t), y(t), z(t)$ définies par trois fonctions d'une variable

$$x(t), y(t), z(t) : I \mapsto \mathbf{R}.$$

Remarque 1.1. Si $t \to z(t) = 0$ est la fonction identiquement nulle, on obtient une *courbe* dans le plan $z = 0$: *une courbe plane.*

1.1. Courbe paramétrée vs. courbe géométrique. Une interprétation possible est que la fonction φ décrit la trajectoire d'un point en fonction du temps : $\varphi(t)$ est la position du point au temps t.

L'image $\varphi(I)$ de l'intervalle I,

$$\varphi(I) = \{\varphi(t),\ t \ dans\ I\}$$

s'appelle la *courbe géométrique* associée à φ. Il est important de bien distinguer entre courbe géométrique et courbe paramétrée : cette derniére donne plus d'information puisque on sait ou se trouve le point en chaque temps t donné. En revanche, la *courbe géométrique* ne donne que l'ensemble de toutes les positions prises par le point quand le temps varie. Ainsi deux courbes paramétrées différentes peuvent donner la même courbe géométrique. Pour aller dans l'autre sens (de la courbe géométrique vers une courbe paramétrès) on fait la définition suivante :

Définition 4.1. *Soit C une courbe géométrique ; une paramétrisation (ou un paramétrage) de C est une courbe paramétrée*

$$\varphi : I \subset \mathbf{R} \to \mathbf{R}^3$$

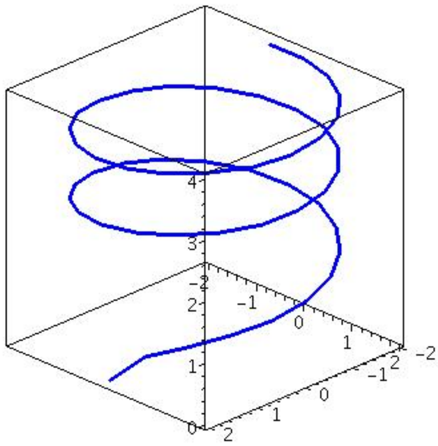


FIGURE 1. $t \mapsto (2\cos(t), 2\sin(t), t^{1/2})$

dont l'image (la courbe géométrique associée) $\varphi(I)$ est C et qui est une bijection entre I et C (à tout point $P \in C$ correspond un unique $t_P \in I$ tel que $\varphi(t_P) = P$).

Ainsi si $C = \varphi(I)$ est la courbe géométrique, on dit que φ fournit une *paramétrisation* de C.

1.2. Exemples.

1.2.1. *Segment de droite.* Soient $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ deux points de l'espace. Soit

$$I = [0,1], \ \varphi(t) = (x(t), y(t), z(t)), \ \begin{array}{l} x(t) = (1-t)x_A + tx_b \\ y(t) = (1-t)y_A + ty_b \\ z(t) = (1-t)z_A + tz_b \end{array}$$

Ainsi $\varphi(0) = A$, $\varphi(1) = B$ et $\varphi(t)$ est le point du segment $[A, B]$ dont le rapport de la distance à A par rapport à la longueur de $[A, B]$ est t. Ainsi la courbe géométrique $\varphi([0, 1])$ est le segment de droite $[A, B]$.

Notons que si on définit

$$I = [0,1], \ \varphi_2(t) = (x_2(t), y_2(t), z_2(t)), \ \begin{array}{l} x_2(t) = tx_A + (1-t)x_b \\ y_2(t) = ty_A + (1-t)y_b \\ z_2(t) = tz_A + (1-t)z_b \end{array}$$

on obtient le même segment de droite $[A, B]$ mais parcouru en sens inverse (de B à A) ; ainsi la courbe géométrique (le segment $[A, B]$) est la même.

1.2.2. *Cercle.* Soit P_0 le point de coordonnées (x_0, y_0, z_0), $R > 0$ et $I = [0, 2\pi]$ et

$$\varphi(t) = (x_0 + R\cos(t), y_0 + R\sin(t), z_0).$$

On obtient comme courbe géométrique un cercle centré en P_0 de rayon R. Notons que ce cercle est contenu dans le plan d'équation $z = z_0$: on a une courbe plane horizontale. Cette courbe n'est pas exactement un paramétrage du cercle car le point

(x_0+R, y_0, z_0) correspond a deux t, 0 et 2π. En revanche la restriction de φ à l'intervalle $[0, 2\pi[$ est un paramétrage.

On étudiera également les deux situation particuliéres suivantes.

1.2.3. *Courbe contenue dans une graphe.* Soit $f : \mathbf{R}^2 \mapsto \mathbf{R}$ une fonction de deux variables. Donnons nous deux fonctions sur I, $x(t), y(t) : I \mapsto \mathbf{R}$ et posons

$$z(t) = f(x(t), y(t));$$

on obtient une courbe paramétrée

$$\varphi(t) : t \mapsto (x(t), y(t), z(t)) = (x(t), y(t), f(x(t), y(t))).$$

En particulier la courbe géométrique associée est entiérement contenue dans le *graphe* de f,

$$\mathscr{G}_f = \{(x, y, z) \in \mathbf{R}^3,\ z = f(x, y).\} = \{(x, y, f(x, y)),\ x, y \in I \times I\}.$$

1.2.4. *Courbe contenue dans une variété de niveau.* La situation précédent se généralise au cas des surfaces de niveau : étant donné une fonction de trois variables

$$F : \mathbf{R}^3 \to \mathbf{R},$$

et $C \in \mathbf{R}$ une constante, on a la variété de niveau C

$$V_F(C) = \{(x, y, z) \in \mathbf{R}^3,\ F(x, y, z) = C\}.$$

Considérons une courbe paramétrée

$$\varphi(t) : t \in I \mapsto (x(t), y(t), z(t)),$$

telle que pour tout $t \in I$ on ait

$$F(\varphi(t)) = F(x(t), y(t), z(t)) = C.$$

Alors la courbe géométrique $\varphi(I)$ est contenue dans la variété de niveau $V_F(C)$.

Remarque. Ce cas généralise le précédent en prenant

$$F(x, y, z) = f(x, y) - z,\ C = 0.$$

2. Vecteur tangent-vecteur vitesse

Définition 4.2. *Soit $\varphi(t)$ une courbe paramétrée ; le vecteur tangent (où vecteur vitesse) en $t = t_0 \in I$ est la limite (quand elle existe)*

$$\lim_{t \to t_0, t \neq t_0} \frac{\varphi(t) - \varphi(t_0)}{t - t_0} = \lim_{h \to 0, h \neq 0} \frac{\varphi(t_0 + h) - \varphi(t_0)}{h}$$

$$= \lim_{h \to 0, h \neq 0} (\frac{x(t_0 + h) - x(t_0)}{h}, \frac{y(t_0 + h) - y(t_0)}{h}, \frac{z(t_0 + h) - z(t_0)}{h}).$$

Ainsi cette limite existe ssi les fonctions $x(t), y(t), z(t)$ sont dérivables en t_0. On note ce vecteur $\varphi'(t_0)$ ou encore $\dot{\varphi}(t_0)$

$$\varphi'(t_0) = \dot{\varphi}(t_0) = (x'(t_0), y'(t_0), z'(t_0))$$

En particulier on a une formule d'approximation linéaire (posant $t = t_0 + h$)

$$\varphi(t) = \varphi(t_0 + h) = \varphi(t_0) + h\varphi'(t_0) + h\vec{\varepsilon}(h) \tag{2.1}$$

avec $\vec{\varepsilon}(h)$ un vecteur tel que

$$\|\vec{\varepsilon}(h)\| \to 0,\ \ h \to 0.$$

2.1. Tangente à une courbe. Considérons la fonction linéaire qui apparait dans la formule d'approximation linéaire ci-dessus

$$h \mapsto \varphi(t_0) + h\varphi'(t_0) \in \mathbf{R}^3$$

ou encore en posant $t = t_0 + h$

$$t \mapsto \varphi(t_0) + (t - t_0)\varphi'(t_0) \in \mathbf{R}^3.$$

Ce sont deux courbes paramétrées qui valent $\varphi(t_0)$ en $t = t_0$ ou $h = 0$ et dont la courbe géométrique associée est une *droite* : la droite passant par $\varphi(t_0)$ et qui est portée par le vecteur $\varphi'(t_0)$.

DÉFINITION 4.3. *Cette droite est la tangente à la courbe φ au point $\varphi(t_0)$.*

La formule d'approximation linéaire (2.1) s'écrit

$$\varphi(t_0 + h) - \varphi(t_0) + h\varphi'(t_0) = h\vec{\varepsilon}(h),\ \ h \to 0$$

ce qui s'interprete en disant que au voisinage de $\varphi(t_0)$ la courbe paramétrée φ est proche de sa tangente.

2.2. Exemples.

2.2.1. *Segment de droite.* Soient $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ deux points de l'espace. Soit

$$I = [0,1],\ \ \varphi(t) = (x(t), y(t), z(t)),\ \ \begin{aligned} x(t) &= (1-t)x_A + tx_b \\ y(t) &= (1-t)y_A + ty_b \\ z(t) &= (1-t)z_A + tz_b \end{aligned}$$

On a

$$\varphi'(t) = (x'(t), y'(t), z'(t)),\ \ \begin{aligned} x'(t) &= -x_A + x_b \\ y'(t) &= -y_A + y_b \\ z'(t) &= -z_A + z_b \end{aligned}$$

Ainsi $\varphi'(t_0)$ est égal au vecteur constant $\vec{AB}$. De même pour

$$\varphi_2(t) = (x_2(t), y_2(t), z_2(t)),\ \ \begin{aligned} x_2(t) &= tx_A + (1-t)x_b \\ y_2(t) &= ty_A + (1-t)y_b \\ z_2(t) &= tz_A + (1-t)z_b \end{aligned}$$

on obtient que $\varphi_2'(t_0) = (x_A - x_B, y_A - y_B, z_A - z_B)$ est égal au vecteur constant $\vec{BA}$.

La tangente à la courbe est exactement la droite AB.

2.2.2. *Cercle.* Soit $C_0 = (x_0, y_0, z_0)$, $R > 0$ et $I = [0, 2\pi]$ et

$$\varphi(t) = (x_0 + R\cos(t), y_0 + R\sin(t), z_0).$$

On obtient

$$\varphi'(t) = (-R\sin(t), R\cos(t), 0).$$

On obtient un vecteur contenu dans le plan horizontal d'équation $z = 0$.

REMARQUE 2.1. Notons que le vecteur $\varphi'(t_0)$ est perpendiculaire au vecteur

$$\overrightarrow{C_0\varphi(t_0)} = (R\cos(t_0), R\sin(t_0), 0)$$

et donc la tangente en $\varphi(t_0)$ est précisement la tangent au centre en ce point.

2.2.3. *Courbe contenue dans un graphe.* Soit $f : \mathbf{R}^2 \mapsto \mathbf{R}$ une fonction de deux variables et deux fonctions sur I, $x(t), y(t) : I \mapsto \mathbf{R}$ et posons

$$z(t) = f(x(t), y(t));$$

$$\varphi(t) : t \mapsto (x(t), y(t), z(t)) = (x(t), y(t), f(x(t), y(t))).$$

En particulier la courbe géométrique associée est entiérement contenue dans le *graphe* de f,

$$\mathscr{G}_f = \{(x, y, z) \in \mathbf{R}^3,\ z = f(x, y).\} = \{(x, y, f(x, y),\ x, y \in I \times I\}.$$

La régle de dérivation des fonctions composées dit que (si $x(t), y(t)$ sont dérivables et $f(x, y)$ est différentiable)

$$\varphi'(t) = (x'(t), y'(t), z'(t))$$

avec

$$z'(t) = \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t) = \nabla f(x(t), y(t)).(x'(t), y'(t)).$$

Etant donné (x_0, y_0) ; rappelons que le vecteur normal au graphe de f au point $(x_0, y_0, f(x_0, y_0))$ est donné par

$$\vec{n}(f)(x_0, y_0) = (\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1)$$

Ainsi, notant $z_0 = f(x(t_0), y(t_0))$ on obtient

$$\vec{n}(f)(x_0, y_0).\varphi'(t_0) = \frac{\partial f}{\partial x}(x(t_0), y(t_0)).x'(t_0) + \frac{\partial f}{\partial y}(x(t_0), y(t_0)).y'(t_0) + (-1).z'(t_0) = 0.$$

Rappelons également que l'équation du plan tangent à $\mathscr{G}_f$ au point

$$\vec{x}_0 = (x_0, y_0, f(x_0, y_0)) = (x_0, y_0, z_0)$$

est donnée par (posant $\vec{x} := (x, y, z)$)

$$(\vec{x}-\vec{x}_0).\vec{n}(f)(x_0, y_0) = (x-x_0)\frac{\partial f}{\partial x}(x(t_0), y(t_0))+(y-y_0)\frac{\partial f}{\partial y}(x(t_0), y(t_0))-(z-z_0) = 0.$$

Si $\vec{x}$ appartient à la tangente à la courbe en $\varphi(t_0) = \vec{x}_0$, il est de la forme

$$\vec{x} = \vec{x}_0 + (t - t_0)\varphi'(t_0),$$

pour $t \in \mathbf{R}$; il vérifie donc

$$(\vec{x} - \vec{x}_0).\vec{n}(f)(x_0, y_0) = (t - t_0)\vec{n}(f)(x_0, y_0).\varphi'(t_0) = 0$$

et appartient donc au plan tangent.

On a donc la

PROPOSITION 4.1. *Le vecteur tangent $\varphi'(t_0)$ est perpendiculaire au vecteur normal au graphe de f au point $\vec{x}_0(x_0, y_0, f(x_0, y_0))$, $\vec{n}_{(x_0,y_0)}(f)$; en d'autre termes, le vecteur tangent est paralléle au plan tangent au graphe de f en $(x_0, y_0, f(x_0, y_0))$ et la droite tangente à la courbe en $\vec{x}_0$ appartient au plan tangent au graphe en ce point.*

2.3. Vecteur tangent à une courbe contenue dans une variété de niveau. En fait l'exemple précédent est un cas particulier du résultat général suivant (appliquer à la fonction

$$F(x, y, z) = f(x, y) - z,\ C = 0.\ )$$

THÉORÈME 4.1. *Soit $F(x, y, z)$ une fonction de trois variables et C une constante et $V_F(C)$ la variété de niveau associée. Soit $\varphi : I \to \mathbf{R}^3$ une courbe paramétrée entiérement contenue dans $V_F(C)$: telle que pour tout t on ait*

$$F(x(t), y(t), z(t)) = C.$$

Soit t dans I tel que la plan tangent à $V_F(C)$ au point $(x(t), y(t), z(t))$ existe (ie. si

$$\nabla F(x(t), y(t), z(t)) \neq \vec{0}.)$$

Alors, le vecteur tangent $\varphi'(t)$ est toujours paralléle à ce plan et la tangente à la courbe au point $\varphi(t)$ est contenue dans le plan tangent.

Preuve: Considérons la fonction

$$h(t) :\ t \to F(x(t), y(t), z(t)).$$

Par hypothése, cette fonction est constante égale à C et sa dérivée est nulle : par la régle de dérivation des fonctions composées (voir ci-dessous), on a pour tout $t_0 \in I$

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

Ainsi le vecteur vitesse $\varphi'(t)$ est perpendiculaire au vecteur normal à la surface $V_F(C)$ au point $\varphi(t) = (x(t), y(t), z(t))$ et donc est paralléle au plan tangent en ce point.

Soit $\vec{x}$ appartenant à la tangente à la courbe en $\vec{x}_0\varphi(t_0)$, $\vec{x}$ est de la forme

$$\vec{x} = \varphi(t_0) + h\varphi'(t_0),$$

pour $h \in \mathbf{R}$; il vérifie donc

$$(\vec{x} - \vec{x}_0).\nabla F(x(t_0), y(t_0), z(t_0)) = h\nabla F(x(t_0), y(t_0), z(t_0)).\varphi'(t_0) = 0$$

et appartient donc au plan tangent. $\square$

2.4. Appendice : formule de dérivation des fonctions composées. Soit

$$F : \begin{array}{ccc} D \subset \mathbf{R}^n & \to & \mathbf{R} \\ (x_1, \cdots, x_n) & \to & F(x_1, \cdots, x_n) \end{array}$$

une fonction de n-variables qu'on suppose différentiable sur D ; soient $x_1(t), \cdots x_n(t)$, n fonctions d'une variable réelle

$$x_i : \begin{array}{ccc} I \subset \mathbf{R} & \to & \mathbf{R} \\ t & \to & x_i(t) \end{array}$$

qu'on suppose dérivables. On définit

$$\varphi : \begin{array}{ccc} I & \to & \mathbf{R}^n \\ t & \to & \varphi(t) = (x_1(t), \cdots, x_n(t)), \end{array}$$

et on suppose que $\varphi(I) \subset D$. On définit alors la fonction d'une variable

$$g = F \circ \varphi : t \in I \to F(\varphi(t)) = F(x_1(t), \cdots x_n(t)).$$

On a le

Théorème 4.2 (Dérivation des fonctions composées). *La fonction $t \to g(t)$ est dérivable sur I et sa dérivée en $t_0 \in I$ vaut*

$$\begin{aligned} g'(t_0) &= x_1'(t_0)\frac{\partial F}{\partial x_1}(x_1(t_0), \cdots x_n(t_0)) + \cdots + x_n'(t_0)\frac{\partial F}{\partial x_n}(x_1(t_0), \cdots x_n(t_0)) \\ &= \nabla F(\varphi(t_0)).\varphi'(t_0) \end{aligned}$$

avec

$$\varphi'(t_0) = (x_1'(t_0), \cdots x_n'(t_0)).$$

Preuve: □

3. Longueur d'une courbe paramétrée

Dans cette section on veut définir et calculer la longueur d'une courbe paramétrée.

3.1. Ligne polygonales. On part du cas le plus simple, celui d'un segment de droite $[A, B]$ avec $A = (x_A, y_A, z_A)$, $B = (x_B, y_B, z_B)$ dans ce cas, la longueur est simplement la norme euclidienne du vecteur $\vec{AB}$

$$\ell([A, B]) = \|\vec{AB}\| = ((x_B - x_A)^2 + (y_B - y_A)^2 + (z_B - z_A)^2)^{1/2}.$$

Le cas suivant est celui d'une *ligne polygonale* : ie. une réunion de segments

$$L = [P_1, P_2] \cup [P_2, P_3] \cup \cdots \cup [P_n, P_{n+1}], \;\; P_i = (x_i, y_i, z_i).$$

dans ce cas la longueur de L est simplement la somme des longueur des segments

$$\ell(L) = \ell([P_1, P_2]) + \cdots + [P_n, P_{n+1}] = \sum_{i=1}^{n} \ell([P_i, P_{i+1}]).$$

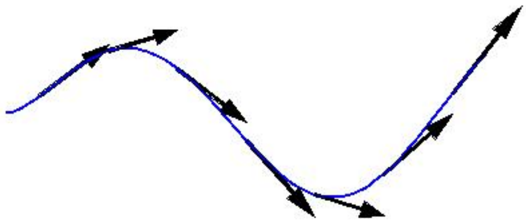


FIGURE 2. Approximation par une ligne polygonale

3.2. Courbes paramétrées. Etant donnée une courbe paramétrée

$$\varphi(t) : [a, b] \to \mathbf{R}^3$$

suffisament *réguliére* (ie. les coordonnées $x(t), y(t), z(t)$ sont continues et dérivables) on peut déduire de la formule d'approximation linéaire (2.1) que dans une sens convenable la courbe sera bien "approximée" par une ligne polygonale de la forme

$$L_k = [\varphi(t_1), \varphi(t_2)] \cup [\varphi(t_2), \varphi(t_3)] \cup \cdots \cup [\varphi(t_n), \varphi(t_{n+1})]$$

avec

$$t_1 = a < t_2 < \cdots < t_k < t_{n+1} = b$$

formant une subdivision de l'intervalle $[a, b]$ qui devient de plus en plus "fine" : k devient de plus en plus grand et les différences succéssives $h_i = t_{i+1} - t_i$ tendent vers 0 de manière *uniforme*. Par exemple, si on subdivise l'intervalle $[a, b]$ en n sous-intevalles de même longueur, n devennant de plus en plus grand : pour $0 \leqslant i \leqslant n$

$$t_i = a + (i-1)\frac{b-a}{n}, \;\; h_i = (b-a)/n.$$

En particulier la longueur $\ell(L_n)$ devrait bien approximer la longueur de la courbe $\ell(\varphi([a, b]))$. C'est effet le cas : supposons qu'on ait choisit une subdivision en n intervalles de même longueur $h = (b-a)/n$. Pour $i = 1 \ldots n$

$$\ell(\varphi([t_i, t_{i+1}])) \sim \ell([\varphi(t_i), \varphi(t_{i+1})]) = \|\overrightarrow{\varphi(t_i)\varphi(t_{i+1})}\|$$

On a par la formule d'approximation linéaire (2.1)

$\varphi(t_{i+1}) - \varphi(t_i) = h\varphi'(t_i) + h\vec{\varepsilon}(h)$ et donc

$$\ell(\varphi([t_i, t_{i+1}])) = h\|\varphi'(t_i)\| + Erreur$$

Sommant les différent termes on obtient que $\ell(\varphi([a, b]))$ est bien approximée par

$$\ell(\varphi([a, b])) \sim \sum_{i=1}^{n} \frac{b-a}{n}\|\varphi'(a + (i-1)\frac{b-a}{n})\|.$$

On reconnait une somme de Riemann qui converge vers l'integrale

$$\sum_{i=1}^{n} \frac{b-a}{n} \|\varphi'(a+(i-1)\frac{b-a}{n})\| \to \int_a^b \|\varphi'(t)\| dt.$$

On en déduit donc la formule

$$\ell(\varphi([a,b])) = \int_a^b \|\varphi'(t)\| dt.$$

3.2.1. *Exemple : segment de droite.*

$$\varphi(t) = ((1-t)x_A + tx_B, (1-t)y_A + ty_B, (1-t)z_A + tz_B).$$

On a

$$\varphi'(t) = (x_B - x_A, y_B - y_A, z_B - z_A) = \overrightarrow{BA}$$

et donc

$$\ell(\varphi([0,1])) = \int_0^1 \|\overrightarrow{BA}\| dt = \|\overrightarrow{BA}\|.$$

On retrouve donc la longueur attendue.

3.2.2. *Exemple : cercle.*

$$\varphi(t) = (x_0 + R\cos(t), y_0 + R\sin(t), z_0)$$

et

$$\varphi'(t) = (-R\sin(t), R\cos(t), 0).$$

Ainsi

$$\ell(\varphi([0,2\pi])) = \int_0^{2\pi} \sqrt{(-R(\sin(t))^2 + (R\cos(t))^2 + 0^2} dt = \int_0^{2\pi} R dt = 2\pi R.$$

3.2.3. *Exemple : courbe contenue dans un graphe.*

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

$$\varphi'(t) = (x'(t), y'(t), z'(t))$$

avec

$$z'(t) = \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t) = \nabla f(x(t), y(t)).(x'(t), y'(t)).$$

$$\ell(\varphi([a,b])) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} dt.$$

3.3. Propriétés de la longueur.

3.3.1. *Additivité par juxtaposition.* Soit

$$\varphi_1 : t \in [a,b] \mapsto (x_1(t), y_1(t), z_1(t))$$

et

$$\varphi_2 : t \in [b,c] \mapsto (x_2(t), y_2(t), z_2(t))$$

deux courbes paramétrées telles que $\varphi_2(b) = \varphi_1(b)$; on forme une nouvelle courbe

$$\varphi_3 : [a,c] \mapsto \mathbf{R}^3$$

en posant $\varphi_3(t) = \varphi_1(t)$ si $a \leqslant t \leqslant b$ et $\varphi_3(t) = \varphi_2(t)$ si $b \leqslant t \leqslant c$. On a alors

$$\ell(\varphi_3([a,c])) = \ell(\varphi_3([a,b])) + \ell(\varphi_3([b,c])) = \ell(\varphi_1([a,b])) + \ell(\varphi_2([b,c])).$$

3.3.2. *Invariance par changement de paramètre.* Soit

$$\varphi : t \in [a,b] \mapsto (x(t), y(t), z(t))$$

une courbe paramétrée ; *un changement de paramètre* est une fonction dérivable

$$u : s \in [a', b'] \mapsto [a,b]$$

telle que

(1) u est strictement croissante
(2) $u(a') = a$ et $u(b') = b$.

Ainsi quand s varie entre a' et b' $u(s)$ parcourt donc l'intégralité du segment $[a,b]$. On obtient donc une nouvelle courbe paramétrée

$$\varphi_2 : s \in [a', b'] \mapsto \varphi_2(s) = \varphi(u(s))$$

avec la même courbe géométrique associée.

Théorème 4.3. *La longueur ne change pas avec le changement de paramétrage :*

$$\ell(\varphi_2([a', b'])) = \int_{a'}^{b'} \|\varphi_2(s)\| ds = \int_a^b \|\varphi(t)\| dt = \ell(\varphi([a,b])).$$

Preuve: On a

$$\ell(\varphi_2([a', b'])) = \int_{a'}^{b'} \|\varphi_2(s)\| ds = \int_{a'}^{b'} \|\varphi_2'(u(s))\| ds$$

on a par dérivation des fonctions composées

$$\varphi_2'(u(s)) = u'(s)\varphi'(u(s)), \ \|\varphi_2'(u(s))\| = |u'(s)| \|\varphi'(u(s))\| = u'(s)\|\varphi'(u(s))\|$$

car $u'(s) \geqslant 0$ et donc

$$\ell(\varphi_2([a', b'])) = \int_{a'}^{b'} \|\varphi'(u(s))\| u'(s) ds.$$

On fait le changement de variable $t = u(s)$, $dt = u'(s)ds$ et donc comme $u(a') = a, \ u(b') = b$

$$\ell(\varphi_2([a', b'])) = \int_{a'}^{b'} \|\varphi'(u(s))\| u'(s) ds = \int_a^b \|\varphi'(t)\| dt.$$

□