

Fonctions de plusieurs variables

Jusqu'à maintenant, vous avez essentiellement appris à étudier les fonctions d'une variable. Cependant, pour modéliser de nombreux phénomènes, les fonctions d'une variable ne suffisent pas ; on a souvent besoin de fonctions de plusieurs variables.

Un exemple. Pour un échantillon d'une mole de gaz de Van der Waals, la pression P du gaz est une fonction de deux variables : sa température T , et le volume V occupé par cet échantillon. On a en effet :

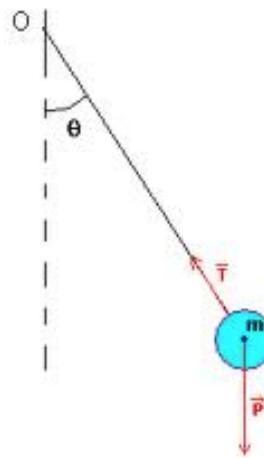
$$P(T, V) = \frac{RT}{V - b} - \frac{a}{V^2}$$

où a , b et R sont des constantes (a et b dépendent du gaz considéré, R est une constante universelle).

Un autre exemple. L'énergie totale E d'un pendule pesant est une fonction de deux variables : l'angle θ que fait le pendule avec la verticale, et sa vitesse angulaire $\dot{\theta}$. On a en effet :

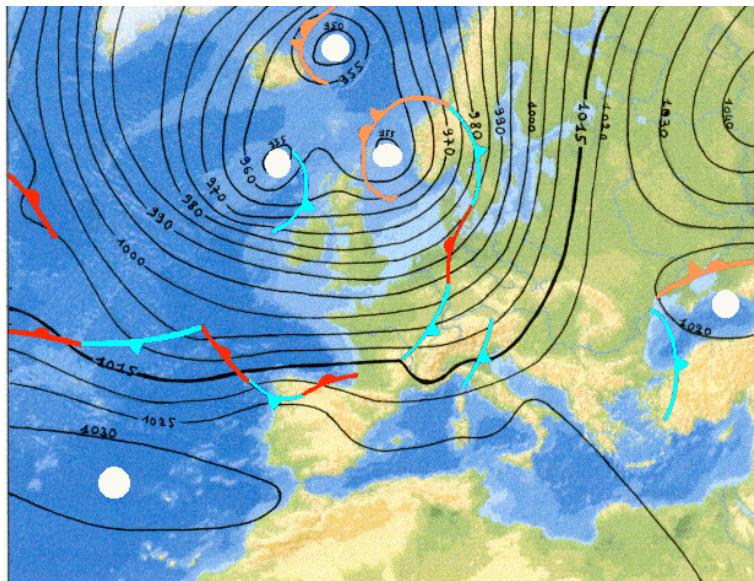
$$E(\theta, \dot{\theta}) = \frac{1}{2}m\ell^2\dot{\theta}^2 + mg\ell(1 - \cos(\theta))$$

où m , g et ℓ sont des constantes (la masse du pendule, la constante de gravitation universelle, et la longueur de la tige du pendule).



Encore un exemple. Si vous voulez faire décrire le temps qu'il fait, à un moment donnée, en Europe, vous allez modéliser la pression et la température par des fonctions de deux variables :

- un point sur la carte de l'Europe sera repéré par deux variables : sa longitude x et sa latitude y ;
 - la pression en ce point, notée $P(x, y)$, sera donc une fonction des variables x et y ;
 - de même, la température en ce point, notée $T(x, y)$, sera donc une fonction des variables x et y .



Bien sûr, si vous voulez être plus précis, il faudra introduire la variable altitude z , et considérer la pression P et la température T comme des fonctions des trois variables x, y, z . Si vous voulez décrire l'évolution de la pression P et la température T au cours du temps, vous devrez introduire une quatrième variable t , et considérer la pression P et la température T comme des fonctions des quatre variables x, y, z, t . *Etc.*

Une fois qu'on a modélisé un phénomène naturel par une fonction de plusieurs variables, il faut être capable — si on veut étudier le phénomène naturel en question — de décrire le comportement de cette fonction. Ceci permet de répondre à de nombreuses questions. Par exemple :

- Un pendule pesant possède-t-il des positions d'équilibre ? Ces positions sont-elles stables ou instables ?
 - Pour un gaz de Van der Waals, existe-t-il des valeurs de la température et du volume pour lesquelles la pression est minimale ?
 - Si on connaît le volume et la température d'un gaz de Van der Waals avec une certaine précision ϵ , avec quelle précision ceci peut-on en déduire la pression du gaz ?

Autrement dit, tout comme il est utile de savoir étudier les fonctions d'une variable, il est utile de savoir étudier les fonctions de plusieurs variables.

Nous allons apprendre à étudier les fonctions de plusieurs variables. La plupart du temps, nous considérerons des fonctions qui dépendent juste de deux — éventuellement trois — variables ; cela permet d'éviter des notations et des calculs trop lourds, et toutes les idées sont déjà présentes.

Vous connaissez de nombreuses notions permettant d'étudier les fonctions d'une variable :

- domaine de définition ;
- représentation graphique (graphe) ;
- limites ;
- fonction dérivée ; sens et tableau de variation ;
- tangente ;
- développement limité ;
- *etc.*

Pour chacune de ces notions, nous devrons nous demander comment elle change quand on passe aux fonctions de plusieurs variables.

À votre avis, qu'est-ce qui change ? Qu'est-ce qui reste essentiellement pareil ?

Pour simplifier, on ne se préoccupera pas de questions de continuité ou de dérivabilité. Autrement dit, on supposera tacitement que toutes les fonctions considérées dans la suite admettent des dérivées partielles dans toutes les directions et à tous les ordres. Ceci nous permettra de mieux nous concentrer sur le calcul et surtout l'utilisation de ces dérivées partielles.

Chapitre 4

Préliminaire : un peu de géométrie dans le plan et l'espace

4.1 Géométrie dans le plan \mathbb{R}^2

4.1.1 Produit scalaire

Si on se donne deux vecteurs $\vec{u}_1 = (x_1, y_1)$ et $\vec{u}_2 = (x_2, y_2)$ dans le plan, on peut définir leur *produit scalaire* par la formule très simple

$$\vec{u}_1 \cdot \vec{u}_2 = x_1 x_2 + y_1 y_2.$$

La *longueur* d'un vecteur (on dit aussi sa *norme*) est alors définie par la formule

$$\|\vec{u}_1\| = \sqrt{\vec{u}_1 \cdot \vec{u}_1} = \sqrt{x_1^2 + y_1^2}.$$

Il y a aussi une autre formule pour le produit scalaire : on a

$$\vec{u}_1 \cdot \vec{u}_2 = \|\vec{u}_2\| \|\vec{u}_1\| \cos \alpha$$

où α est l'angle entre les deux vecteurs. Un des intérêts de cette formule est qu'elle permet de voir quand le produit scalaire s'annule : on a (si aucun des deux vecteurs n'est nul)

$$\begin{aligned} \vec{u}_1 \cdot \vec{u}_2 = 0 &\Leftrightarrow \cos \alpha = 0 \\ &\Leftrightarrow \alpha = \frac{\pi}{2} (\text{ modulo } \pi) \\ &\Leftrightarrow \text{les deux vecteurs sont orthogonaux.} \end{aligned}$$

En pratique. La première formule permet de calculer très facilement le produit scalaire de deux vecteurs quand on a leurs coordonnées ; la deuxième permet de dire que si le produit scalaire est nul, alors les vecteurs sont orthogonaux.

Remarque. Pour le mathématicien, ce qui précéde n'a pas de sens : on n'a pas défini clairement ce qu'était l'angle α , ni ce que ça voulait dire que deux vecteurs sont orthogonaux. Pour s'en sortir, il suffit de procéder en sens inverse : l'angle α est *défini* par la deuxième formule ; de même, on peut *définir* l'orthogonalité en disant que deux vecteurs sont orthogonaux si leur produit scalaire est nul.

Cependant, tout le monde (y compris les mathématiciens !) sait bien que les architectes n'ont pas attendu l'invention du produit scalaire pour savoir mesurer des angles et tracer des perpendiculaires...

Équation d'une droite

Soit maintenant $M_0 = (x_0, y_0)$ un point du plan, et $\vec{n} = (a, b)$ un vecteur du plan. Quelle est l'équation de la droite D passant par M_0 et orthogonale à \vec{n} ? Le point clé est d'arriver à caractériser les points de (D) en terme de produit scalaire. Soit $M = (x, y)$ un point quelconque du plan. On a l'équivalence :

$$\begin{aligned} M \text{ appartient à la droite } D &\Leftrightarrow \text{les vecteurs } \overrightarrow{M_0M} \text{ et } \vec{n} \text{ sont orthogonaux} \\ &\Leftrightarrow \overrightarrow{M_0M} \cdot \vec{n} = 0 \\ &\Leftrightarrow a(x - x_0) + b(y - y_0) = 0. \end{aligned}$$

Résumons :

Dans le plan \mathbb{R}^2 , la droite passant par le point $M_0 = (x_0, y_0)$ et orthogonale au vecteur $\vec{n} = (a, b)$ admet pour équation

$$a(x - x_0) + b(y - y_0) = 0.$$

Exemple. Vous pouvez vérifier votre formule sur l'exemple suivant : l'équation de la droite passant par le point $M_0 = (1, 2)$ et orthogonal au vecteur $\vec{v} = (3, 4)$ est

$$3(x - 1) + 4(y - 2) = 0.$$

Attention ! Dans cet exemple, le vecteur $\vec{v} = (3, 4)$ se dessine à partir du point $M_0 = (1, 2)$: on dessine donc une flèche qui part de M_0 et qui aboutit au point $(4, 6)$ (et non pas au point $(3, 4)$).

4.2 Géométrie dans \mathbb{R}^3

4.2.1 Produit scalaire

La définition et les propriétés du produit scalaire dans \mathbb{R}^2 se généralise sans problème dans l'espace \mathbb{R}^3 , en rajoutant une coordonnée partout. Par exemple, le produit scalaire de deux vecteurs $\vec{u}_1 = (x_1, y_1, z_1)$ et $\vec{u}_2 = (x_2, y_2, z_2)$ est défini par

$$\vec{u}_1 \cdot \vec{u}_2 = x_1 x_2 + y_1 y_2 + z_1 z_2.$$

Et deux vecteurs sont orthogonaux si et seulement si leur produit scalaire est nul.

4.2.2 Équation d'un plan

Donnons-nous un point $M_0 = (x_0, y_0, z_0)$ de \mathbb{R}^3 , et $\vec{n} = (a, b, c)$ un vecteur. Dans ce cadre, l'ensemble des points M tels que le vecteur $\overrightarrow{M_0M}$ est orthogonal au vecteur \vec{n} est un plan P . La technique vue dans \mathbb{R}^2 conduit donc à l'équation de ce plan P .

Plus précisément, soit $M = (x, y, z)$ un point quelconque de l'espace. On a l'équivalence :

$$\begin{aligned} M \text{ appartient au plan } P &\Leftrightarrow \text{les vecteurs } \overrightarrow{M_0M} \text{ et } \vec{n} \text{ sont orthogonaux} \\ &\Leftrightarrow \overrightarrow{M_0M} \cdot \vec{n} = 0 \\ &\Leftrightarrow a(x - x_0) + b(y - y_0) + c(z - z_0) = 0. \end{aligned}$$

Résumons :

Dans l'espace \mathbb{R}^3 , la droite passant par le point $M_0 = (x_0, y_0, z_0)$ et orthogonale au vecteur $\vec{n} = (a, b, c)$ admet pour équation

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

Exemple. Vous pouvez vérifier votre formule sur l'exemple suivant : l'équation du plan passant par le point $M_0 = (1, 2, 3)$ et orthogonal au vecteur $\vec{v} = (4, 5, 6)$ est

$$4(x - 1) + 5(y - 2) + 6(z - 3) = 0.$$

Réiproquement, on peut montrer que toute équation du type $\alpha x + \beta y + \gamma z + \delta = 0$ est l'équation d'un plan, qui est orthogonal au vecteur (α, β, γ) .

Remarque. Nous verrons bientôt une autre manière de formuler ce fait : les graphes des fonctions affines de deux variables (c'est-à-dire les fonctions du type $f(x, y) = ax + by + c$) sont des plans.

Équations de droites

Comme nous venons de la voir, dans l'espace \mathbb{R}^3 , une équation du type $\alpha x + \beta y + \gamma z + \delta = 0$ définit un plan. Pour définir une droite, il faut deux équations. En effet, toute droite D peut être vue comme l'intersection de deux plans P_1 et P_2 . La droite D est alors l'ensemble des solutions du système à deux équations constitué par des équations de P_1 et P_2 .

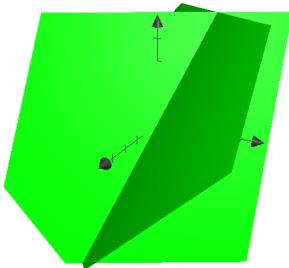
Exemple. Dans \mathbb{R}^3 , l'axe (Ox) est l'intersection des plans (Oxy) et (Oxz) , c'est-à-dire l'intersection des plans d'équations $z = 0$ et $y = 0$. L'axe (Ox) admet donc pour système d'équations

$$\begin{cases} z = 0 \\ y = 0 \end{cases}$$

Exemple. L'ensemble des points de \mathbb{R}^3 qui vérifie le système d'équations

$$\begin{cases} 2x + y + z - 3 = 0 \\ x - y + z - 1 = 0 \end{cases}$$

est une droite, intersection du plan P_1 d'équation $2x + y + z - 3 = 0$ et du plan P_2 d'équation $x - y + z - 1 = 0$ (figure ci-dessous).



Chapitre 5

Fonctions de plusieurs variables : généralités et représentations graphiques

5.1 Fonction, ensemble de définition

Définition (Fonction, ensemble de définition : définition formelle). *Soit E un sous-ensemble de \mathbb{R}^2 . Une fonction f définie sur E et à valeurs dans \mathbb{R} fait correspondre, à tout élément (x, y) de E , un unique élément $f(x, y)$ dans \mathbb{R} . L'ensemble E est appelé ensemble de définition de f . La fonction f est une fonction des deux variables x et y .*

En remplaçant E par un sous-ensemble de \mathbb{R}^3 , on définit de même une notion de fonction de trois variables. En remplaçant E par un sous-ensemble de \mathbb{R}^4 , on définit de même une notion de fonction de quatre variables. Etc.

Définition (Fonction, ensemble de définition : définition pratique). *D'un point strictement mathématique et formel, pour définir une fonction, il faut procéder comme ci-dessus : d'abord se donner un sous-ensemble de \mathbb{R}^2 , puis définir une fonction sur cet ensemble. En pratique (comme vous en avez l'habitude pour les fonctions d'une variable), une fonction de deux variables est souvent définie simplement par une formule. Dans ce cas, ce qu'on appelle ensemble de définition de la fonction est l'ensemble des points de \mathbb{R}^2 où la formule a un sens. Par exemple :*

1. $f(x, y) = x^2 + y^2$ est une fonction définie sur \mathbb{R}^2 tout entier; autrement dit, l'ensemble de définition de cette fonction est \mathbb{R}^2 tout entier;
2. $f(x, y) = \arctan \frac{y}{x}$ est une fonction définie pour $x \neq 0$, autrement dit, l'ensemble de définition de cette fonction est le plan \mathbb{R}^2 privé de l'axe des ordonnées.

Bien sûr, la même remarque s'applique aux fonctions de trois variables ; par exemple, $f(x, y, z) = \frac{1}{x^2 + y^2 + z^2}$ est une fonction des trois variables x, y, z définie pour $x^2 + y^2 + z^2 \neq 0$ (ce qui est équivalent à $(x, y, z) \neq (0, 0, 0)$) ; l'ensemble de définition de cette fonction est donc l'espace \mathbb{R}^3 privé de l'origine $(0, 0, 0)$.

Vous verrez en TD des fonctions définies par des formules simples, et ayant pourtant des ensembles de définitions compliqués. Souvent il y a deux étapes assez différentes pour la

compréhension de l'ensemble de définition d'une fonction f de deux variables :

- d'abord, on trouve une équation pour l'ensemble de définition de f (par exemple, l'équation $x \neq 0$ pour la fonction $\arctan \frac{y}{x}$) ;
- puis, il faut utiliser cette équation pour tracer, dans le plan \mathbb{R}^2 , l'ensemble de définition de f (pour la fonction $\arctan \frac{y}{x}$, c'est le plan \mathbb{R}^2 privé de l'axe des ordonnées). Cette deuxième étape est souvent la plus difficile.

Fonction de deux variables versus courbes paramétrées planes. Attention : fonctions de deux variables et courbes paramétrées planes sont des objets très différents. Une fonction de deux variables (sous-entendu “à valeurs réelles”) est une fonction qui part de \mathbb{R}^2 (ou une partie de \mathbb{R}^2) et arrive dans \mathbb{R} . Une courbe paramétrée plane est une fonction qui part de \mathbb{R} et arrive dans \mathbb{R}^2 . Une fonction de deux variables est utilisée pour modéliser une situation où une quantité réelle (par exemple la température de l'air) dépend de deux autres quantités réelles (par exemple la latitude et la longitude). Une courbe paramétrée plane est utilisée pour modéliser une situation où deux quantités réelles (typiquement l'abscisse et l'ordonnée d'un point) dépendent d'une seule autre quantité réelle typiquement le temps).

5.2 Graphe

Définition (Graphe). Soit f une fonction de deux variables, \mathcal{D}_f son ensemble de définition. On appelle *graphe de f* , ou *surface représentative de f* , l'ensemble des points (x, y, z) de l'espace \mathbb{R}^3 où (x, y) est dans \mathcal{D}_f et qui vérifient la relation $z = f(x, y)$.

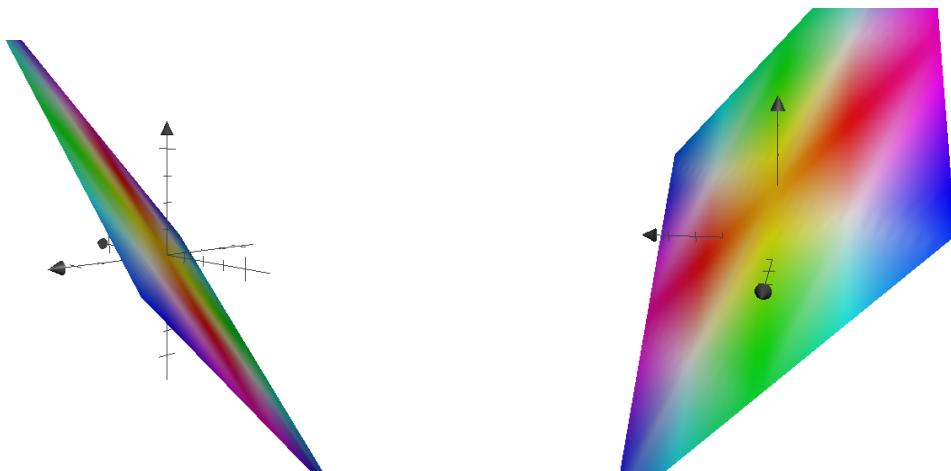
Pour comprendre cette notion, il est bon de se rappeler comment on faisait pour tracer le graphe d'une fonction d'une variable : pour chaque x sur son axe, on s'élèvait d'une hauteur $y = f(x)$. En deux variables, c'est similaire : pour chaque point (x, y) sur le plan des variables, on s'élève d'une hauteur $z = f(x, y)$.

Remarque. Bien sûr, on peut définir également le graphe d'une fonction de trois, quatre, etc. variables. Hélas, le graphe d'une fonction de p variables est un objet de dimension p qui vit dans \mathbb{R}^{p+1} . On ne peut donc pas le visualiser pour $p \geq 3$. C'est pourquoi on ne considérera que des graphes de fonctions de deux variables.

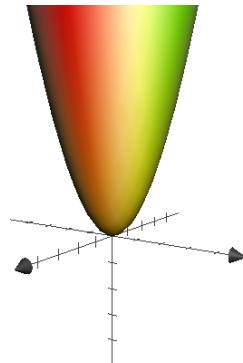
Exemple de graphes (I) : fonctions affines. Comme en une variable, les fonctions de deux variables les plus simples sont les *fonctions affines*

$$(x, y) \mapsto ax + by + c$$

où a, b, c sont des constantes. Leur graphes sont des plans. Réciproquement, tout plan (non vertical) est le graphe d'une fonction affine. Ci-dessous, les graphes des fonctions $f_1(x, y) = x + y$ et $f_2(x, y) = -x - y + 2$.

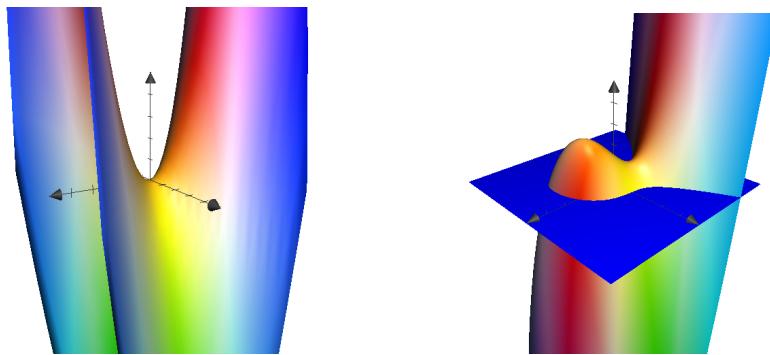


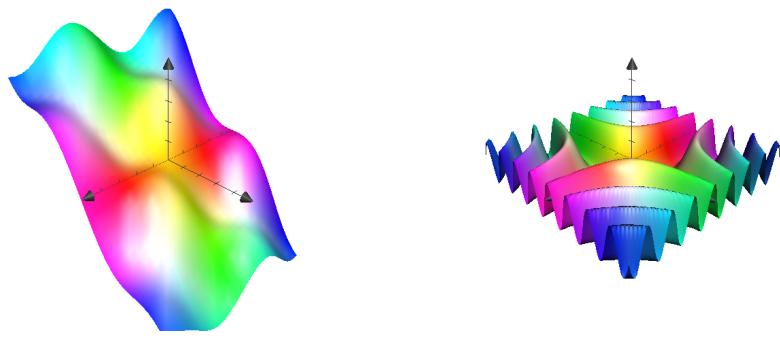
Exemple de graphe (II) : $x^2 + y^2$. La figure ci-contre montre une portion du graphe de la fonction $f(x, y) = x^2 + y^2$ (“paraboloïde à une nappe”). Sur la figure, on a dessiné le plan (Oxy) (“plan des variables”); au dessus de chaque point (x, y) dans ce plan, on place le point de hauteur $z = f(x, y)$. Autrement dit, si on imagine que le graphe est le dessin d’un relief, la fonction f est la fonction altitude (x, y étant latitude et longitude). Le plan (Oxy) est le plan d’altitude $z = 0$, graphe de la fonction $(x, y) \mapsto 0$; pour une carte du relief, il correspond au niveau de la mer.



Exemple de graphes (III). Les figures ci-dessous représentent successivement

- la fonction $f(x, y) = x^2 - y^2$ (selle de cheval, ou col de montagne ou hyperbololoïde à une nappe) ;
- la fonction $f(x, y) = -\frac{x^3}{3} - xy - y^2 + x + \frac{3}{2}$ (une presqu’île avec une montagne, et la mer d’équation $z = 0$) ;
- la fonction $f(x, y) = \sin(x) \sin(y) - y$ (un champ de bosse) ;
- la fonction $f(x, y) = \sin(xy)$.

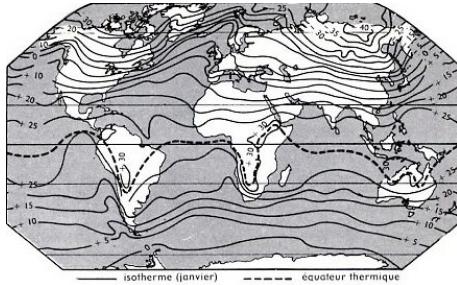




5.3 Lignes de niveau

Le graphe d'une fonction de deux (*a fortiori* trois, quatre, *etc.*) variables est beaucoup plus difficile à tracer que le graphe d'une fonction d'une variable : difficulté du dessin, et surtout, pas de notion équivalente au tableau de variation des fonctions d'une variable¹. C'est pourquoi on utilise souvent d'autres modes de représentation graphique. Par exemple, la carte de la première page offre une représentation graphique d'une fonction pression $P(x, y)$; on a dessiné des *isobares* : un isobare est une courbe sur laquelle f est constante.

Ce type de représentation est très utilisé ; par exemple, pour la température, on trace les isothermes, courbes sur lesquelles la température est constante. Pour les cartes géographique représentant le relief, à la place d'une carte en trois dimension (pas pratique à plier), on trace les courbes d'altitude constante.



1. Nous y reviendrons. Le problème vient du fait que les points de \mathbb{R}^2 ne sont pas ordonnés contrairement aux points de \mathbb{R} .



Définition (Lignes de niveau). Soit f une fonction de deux variables, et h un nombre réel. On appelle ligne de niveau h de la fonction f est l'ensemble des points (x, y) du plan (Oxy) en lesquels f prend la valeur h :

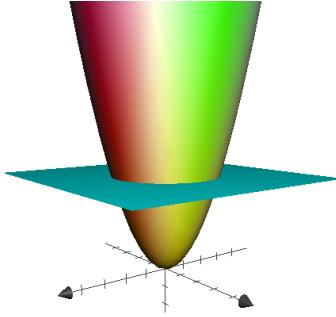
$$L_h = \{(x, y) \text{ tels que } f(x, y) = h\}.$$

Exemple. Déterminer la ligne de niveau 1 de la fonction $f(x, y) = x^2 + y^2$? La ligne de niveau 0? La ligne de niveau -1 ?

Lien avec le graphe. Le graphe de f est l'ensemble des points (x, y, z) tels que $z = f(x, y)$. La ligne de niveau h de f est l'ensemble des points (x, y) tels que $f(x, y) = h$. Pour passer du graphe de f à la ligne de niveau h , il faut donc se restreindre aux points du graphe qui vérifient $z = h$, puis oublier la coordonnée z . L'ensemble des points qui vérifient $z = h$ est le plan horizontal d'altitude h . Oublier la coordonnée z , c'est projeter sur le plan (Oxy) . On voit donc que :

La ligne de niveau h de f est obtenue en projetant dans le plan (Oxy) l'intersection du graphe de f avec le plan horizontal $z = h$.

Exemple. Ci-dessous, la ligne de niveau 5 de la fonction $f(x, y) = x^2 + y^2$, vue comme intersection du graphe de cette fonction (qui est un paraboloïde) et du plan d'équation $z = 5$.



5.4 Fonctions partielles

Beaucoup de problèmes concernant les fonctions de plusieurs variables peuvent se ramener à des problèmes concernant les fonctions d'une seule variable. Pour cela, on utilise les *fonctions partielles*, qui sont obtenues en fixant la valeur de l'une des variables.

Définition (Fonctions partielles). *Soit f une fonction de deux variables x, y , et (x_0, y_0) un point du domaine de définition de f . On appelle fonctions partielles au point (x_0, y_0) les deux fonctions*

$$f_1 : x \mapsto f(x, y_0) \quad \text{et} \quad f_2 : y \mapsto f(x_0, y)$$

Exemple. Si l'on reprend notre fonction définie par $f(x, y) = x^2 + y^2$, qu'on se place au point $(0, 0)$, alors :

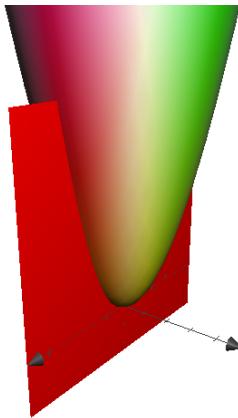
- la première fonction partielle est $f_1 : x \mapsto f(x, 0) = x^2$ (on a fixé $y = 0$) ;
- la seconde fonction partielle est $f_2 : y \mapsto f(0, y) = y^2$ (on a fixé $x = 0$).

Test. Donner les fonctions partielles de $f(x, y) = x^2 + y^2$ au point $(x_0, y_0) = (1, 2)$.

Lien avec le graphe. Fixer $y = y_0$ revient à se placer dans le plan vertical d'équation $y = y_0$. De même, fixer $x = x_0$ revient à se placer dans le plan vertical d'équation $x = x_0$. On voit donc que :

- Le graphe de la fonction partielle $f_1 : x \mapsto f(x, y_0)$ est la projection dans le plan (Oxz) de l'intersection du graphe de f avec le plan vertical d'équation $y = y_0$.
- Le graphe de la fonction partielle $f_2 : y \mapsto f(x_0, y)$ est la projection dans le plan (Oyz) de l'intersection du graphe de f avec le plan vertical d'équation $x = x_0$.

Exemple. Sur la figure ci-dessous, on a représenté le graphe de la fonction $f(x, y) = x^2 + y^2$ (un paraboloïde) et le plan d'équation $x = 0$. Leur intersection est le graphe de la fonction partielle $f_2 : y \mapsto f(0, y) = y^2$ (ce graphe est bien sûr une parabole).



Exercice. Donner les fonctions partielles de $f(x, y) = \sin(x)\sin(y)$, et expliquer l'allure du graphe de f .

Cas des fonctions de trois variables. Si f est maintenant une fonction de deux variables x, y, z , et (x_0, y_0, z_0) est un point du domaine de définition de f , on peut aussi définir des notions de fonctions partielles de f en (x_0, y_0, z_0) . Mais la situation est plus compliquée : on a trois fonctions partielles de deux variables

- $(x, y) \mapsto f(x, y, z_0)$ (on fixe la valeur de z) ;
- $(y, z) \mapsto f(x_0, y, z)$ (on fixe la valeur de x) ;
- $(x, z) \mapsto f(x, y_0, z)$ (on fixe la valeur de y) .

et trois fonctions partielles d'une variables :

- $x \mapsto f(x, y_0, z_0)$ (on fixe les valeurs de y et z) ;
- $y \mapsto f(x_0, y, z_0)$ (on fixe les valeurs de x et z) ;
- $z \mapsto f(x_0, y_0, z)$ (on fixe les valeurs de x et y) .

Bien sûr en quatre variables, c'est encore plus compliqué...

5.5 Un danger

Attention, les différents objets (ensemble de définition, graphe, lignes de niveaux, graphes des fonctions partielles) permettant de représenter graphiquement une fonction vivent dans des espaces différents.

Ainsi, pour une fonction de deux variables $(x, y) \mapsto f(x, y)$:

- le graphe de f est un sous-ensemble de l'espace \mathbb{R}^3 muni des coordonnées (x, y, z) ;
- l'ensemble de définition de f est un sous-ensemble du plan horizontal muni des coordonnées (x, y) ;
- le dessin des lignes de niveau de f se situe lui-aussi dans le plan horizontal muni des coordonnées (x, y) ;
- le dessin du graphe d'une fonction partielle $f_1 : x \mapsto f(x, y_0)$ est un dessin dans un plan vertical muni des coordonnées (x, z) ;
- le dessin du graphe d'une fonction partielle $f_2 : y \mapsto f(x_0, y)$ est un dessin dans un plan vertical muni des coordonnées (y, z) .

Pour chacun des objets graphiques qu'on considère, il ne faut pas se tromper de cadre !

Chapitre 6

Fonctions de plusieurs variables : dérivées

Lorsqu'on veut des informations sur le comportement d'une fonction *d'une variable* $x \mapsto f(x)$ au voisinage d'un point x_0 , on peut calculer sa dérivée, qui nous une approximation de f par une fonction affine. Cette fonction affine n'est autre que la partie principale du développement limité à l'ordre 1 de f . Graphiquement, cela revient à approcher le graphe de f par sa tangente. On en déduit en particulier le sens de variation de f au voisinage de x_0 .

Peut-on décrire une théorie similaire pour une fonction de plusieurs variables ?

La réponse est affirmative. Par exemple, $f(x, y)$ est une fonction de deux variable, on va encore pouvoir approcher $f(x, y)$, au voisinage d'un point (x_0, y_0) , par une fonction affine ; on obtiendra cette approximation affine en calculant les *dérivées partielles* de f au point considéré. Le graphe de l'approximation affine sera le plan tangent au graphe de f . On pourra aussi en déduire le “sens de variation” de f au voisinage de (x_0, y_0) , mais la situation est plus compliquée qu'en dimension 1 : en effet, f peut être croissante dans certaines directions, et décroissante dans d'autres.

Rappelons que l'on suppose tacitement que toutes les fonctions considérées dans la suite admettent des dérivées partielles dans toutes les directions et à tous les ordres.

6.1 Dérivées partielles

Exemple. Considérons la fonction de deux variables $(x, y) \mapsto f(x, y)$ définie par

$$f(x, y) = -\frac{x^3}{3} - xy - y^3 + \frac{3}{2}.$$

Si l'on fixe la valeur de l'une des deux variables, on obtient une fonction d'une variable (c'est une fonction partielle de f). On peut alors dériver cette fonction d'une variable. Par exemple,

- si on fixe $y = 0$, alors on obtient $f(x, 0) = -\frac{x^3}{3} + \frac{3}{2}$. La dérivée de cette fonction de x est $x \mapsto -x^2$. La valeur de cette dérivée en $x = 1$ est -1 . On notera $\frac{\partial f}{\partial x}(1, 0) = -1$.
- si on fixe $x = 1$, alors on obtient $f(1, y) = y - y^3 + \frac{7}{6}$. La dérivée de cette fonction de y est $y \mapsto 1 - 3y^2$. La valeur de cette dérivée en $y = 0$ est 1 . On notera $\frac{\partial f}{\partial y}(1, 0) = 1$.

Passons aux définitions formelles. Commençons par le cas d'une fonction de deux variables.

Définition (Dérivées partielles en un point). *Soit f une fonction de deux variables (x, y) , définie au voisinage de (x_0, y_0) . Notons $f_1 : x \mapsto f(x, y_0)$ et $f_2 : y \mapsto f(x_0, y)$ les deux fonctions partielles de f au point (x_0, y_0) . On appelle dérivées partielles de f au point (x_0, y_0) les nombres*

$$\frac{\partial f}{\partial x}(x_0, y_0) := f'_1(x_0) \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) := f'_2(y_0).$$

Remarque. Souvenez-vous que la dérivée en un point z_0 d'une fonction d'une variable $z \mapsto g(z)$ est définie comme la limite quand h tend vers 0 du taux d'accroissement de g entre z_0 et $z_0 + h$:

$$g'(z_0) = \lim_{h \rightarrow 0} \frac{g(z_0 + h) - g(z_0)}{h}.$$

Par conséquent, si f est une fonction de deux variables (x, y) , les dérivées partielles de f en un point (x_0, y_0) peuvent elles-aussi s'écrire comme des limites de taux d'accroissements ; plus précisément, on a :

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \quad \text{et} \quad \frac{\partial f}{\partial y}(x_0, y_0) = \lim_{k \rightarrow 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k}.$$

Exemple. Calculer les dérivées partielles de la fonction $g(x, y) = \sin(x + y) + y$ au point $(\pi/2, \pi/2)$.

Définition (Fonctions dérivées partielles). *Les fonctions dérivées partielles de f sont les fonctions de deux variables*

$$(x, y) \mapsto \frac{\partial f}{\partial x}(x, y) \quad \text{et} \quad (x, y) \mapsto \frac{\partial f}{\partial y}(x, y).$$

En pratique, on calcule directement les fonctions $(x, y) \mapsto \frac{\partial f}{\partial x}(x, y)$ et $(x, y) \mapsto \frac{\partial f}{\partial y}(x, y)$, sans passer par les fonctions partielles : on dérive par rapport à une variable en considérant l'autre comme une constante.

Exemple. Calculer les fonctions dérivées partielles des fonctions $f(x, y) = x^2 + xy + y^3$ et $g(x, y) = \sin(x + y) + y$

Remarque. Attention au sens des différents symboles ! Dans les définitions ci-dessus :

- x_0, y_0 sont des nombres, (x_0, y_0) est un point du plan ;
- x est juste un symbole : $\frac{\partial}{\partial x}$ est un symbole qui signifie qu'on dérive suivant la première variable, qui est souvent notée x (mais pas toujours : par exemple en physique, les variables sont parfois le temps t , la pression p , etc.).
- $\frac{\partial f}{\partial x}(x_0, y_0)$ est un nombre ;
- $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont des fonctions de deux variables.

Bien sûr, on a des définitions identiques pour des fonctions de trois, quatre, *etc.* variables ; une fonction de n variable possède n dérivées partielles ; pour les obtenir, on fixe les valeurs de toutes les variables sauf une et on dérive par rapport à la variable restante. Par exemple, pour une fonction de trois variables (que je note volontairement p, T, V pour changer) :

Définition. Soit f une fonction de deux variables (p, T, V) , définie au voisinage de (p_0, T_0, V_0) .
On note

$$f_1 : p \mapsto f(p, T_0, V_0) \quad f_2 : T \mapsto f(p_0, T, V_0) \quad f_3 : V \mapsto f(p_0, T_0, V)$$

les trois fonctions partielles d'une variable de f au point (p_0, T_0, V_0) . Les dérivées partielles de f au point (p_0, T_0, V_0) sont les nombres

$$\frac{\partial f}{\partial p}(p_0, T_0, V_0) = f'_1(p_0) \quad \frac{\partial f}{\partial T}(p_0, T_0, V_0) = f'_2(T_0) \quad \frac{\partial f}{\partial V}(p_0, T_0, V_0) = f'_3(V_0).$$

Les fonctions dérivées partielles de f sont les fonctions de trois variables

$$(p, T, V) \mapsto \frac{\partial f}{\partial p}(p, T, V) \quad (p, T, V) \mapsto \frac{\partial f}{\partial T}(p, T, V) \quad (p, T, V) \mapsto \frac{\partial f}{\partial V}(p, T, V).$$

6.2 Développement limité à l'ordre 1

6.2.1 Rappel concernant les fonctions d'une variable

Soit f est une fonction d'une variable x , et x_0 est un point du domaine de définition de f .

Un développement limité à l'ordre 1 de la fonction f au point x_0 est une expression de la forme

$$f(x) = a + b(x - x_0) + (x - x_0)\epsilon(x - x_0)$$

où a, b sont des nombres réels et ϵ est une fonction qui tend vers 0 en 0. Cette condition sur ϵ assure que $(x - x_0)\epsilon(x - x_0)$ très petit (en fait, négligeable devant $a + b(x - x_0)$ si a ou b est non-nul) lorsque x s'approche de x_0 . On écrit aussi

$$f(x) = a + b(x - x_0) + o(x - x_0)$$

ou encore, en posant $x = x_0 + h$,

$$f(x_0 + h) = a + b.h + o(h).$$

Si la fonction f est dérivable, à dérivée continue au point x_0 , alors on sait calculer facilement un développement limité de f en x_0 : on a

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

ou de manière équivalente

$$f(x_0 + h) = f(x_0) + f'(x_0)h + o(h).$$

C'est la formule de Taylor.

On dit que la fonction $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ est la meilleure approximation affine de f au voisinage du point x_0 . En effet, les formules ci-dessus indique que, pour x proche de x_0 , la fonction $x \mapsto f(x)$ est bien approchée par la fonction affine $x \mapsto f(x_0) + f'(x_0)(x - x_0)$.

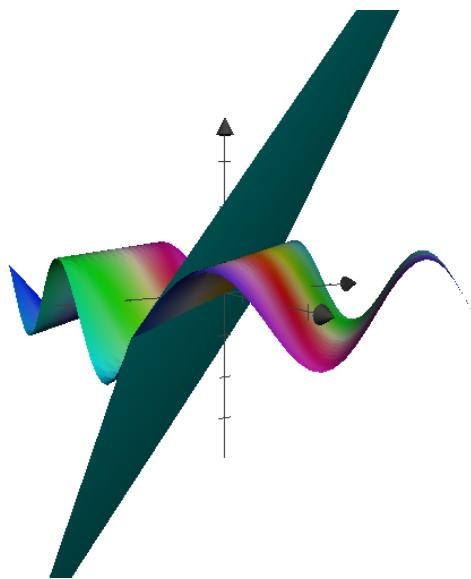
Graphiquement, la droite d'équation $y = f(x_0) + f'(x_0)(x - x_0)$ est la tangente au graphe de f au point x_0 .

6.2.2 Un exemple avec une fonction de deux variables

Le développement limité (habituel, en une variable) de la fonction sinus en 0 conduit à un développement limité de la fonction $(x, y) \mapsto \sin(x + y)$:

$$\begin{aligned}\sin(x + y) &= x + y + (x + y)\varepsilon_1(x + y) \\ &= x + y + \|(x, y)\| \left(\frac{x + y}{\|(x, y)\|} \varepsilon_1(x + y) \right) \\ &= x + y + \|(x, y)\| \varepsilon_2(x, y)\end{aligned}$$

avec $\varepsilon_2(x, y)$ qui tend vers 0 quand le couple (x, y) tend vers $(0, 0)$ (ce qui signifie que la norme $\|(x, y)\|$ du couple (x, y) tend vers 0). On dira que $(x, y) \mapsto x + y$ est une approximation affine de $(x, y) \mapsto \sin(x + y)$ au voisinage du point $(0, 0)$. On constate graphiquement que le plan d'équation $z = x + y$ est le plan tangent en $(0, 0)$ au graphe de f .



6.2.3 En général pour les fonctions de deux variables

Développement limité. Soit f une fonction de deux variables x et y , et (x_0, y_0) un point du domaine de définition de f .

Définition (Développement limité à l'ordre 1). *Un développement limité à l'ordre 1 de la fonction au point (x_0, y_0) est une expression de la forme*

$$(1) \quad f(x, y) = a + b(x - x_0) + c(y - y_0) + \|(x - x_0, y - y_0)\| \varepsilon(x - x_0, y - y_0)$$

où a, b, c sont des nombres et ε est une fonction de deux variables qui tend vers 0 en $(0, 0)$.

La condition “ ε tend vers 0 en $(0, 0)$ ” signifie que le nombre $\varepsilon(x - x_0, y - y_0) = \varepsilon(h, k)$ devient aussi petit que l'on veut quand $(x - x_0, y - y_0) = (h, k)$ se rapproche de $(0, 0)$, autrement dit quand (x, y) se rapproche de (x_0, y_0) .

Remarque. En posant $h = x - x_0$ et $k = y - y_0$, l'expression 1 peut se réécrire sous la forme :

$$f(x_0 + h, y_0 + k) = a + bh + ck + \|(h, k)\|\varepsilon(h, k).$$

Proposition 6.2.1 (Formule de Taylor-Young à l'ordre 1, en deux variables). *On suppose (comme toujours dans ce texte) que les dérivées partielles de f existent et sont continues au voisinage de (x_0, y_0) . Alors le développement limité à l'ordre 1 de f en (x_0, y_0) s'écrit :*

$$f(x, y) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0).(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0).(y - y_0) + \|(x - x_0, y - y_0)\|\varepsilon(x - x_0, y - y_0),$$

ou encore, de manière équivalente, :

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0).h + \frac{\partial f}{\partial y}(x_0, y_0).k + \|(h, k)\|\varepsilon(h, k)$$

Meilleure approximation affine. Écrire un développement limité à l'ordre 1 de f au point (x_0, y_0) revient à écrire f sous la forme de la somme d'une fonction affine et d'un reste qui devient négligeable quand on s'approche de (x_0, y_0) . Ceci justifie la définition suivante :

Définition (Approximation affine). *On dit que la fonction affine*

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

est la meilleure approximation affine de f au voisinage du point (x_0, y_0) .

Typiquement, la meilleure approximation affine d'une fonction sert à estimer l'incertitude que l'on a sur la valeur de $f(x, y)$ si on connaît les valeurs x et y qu'approximativement.

Supposons par exemple qu'une usine fabrique des cylindres d'aluminium de diverses tailles. Le diamètre d de chaque cylindre est mesuré à une précision de 0,05% près. la longueur ℓ de chaque cylindre est mesuré à une précision de 0,1% près. Quelle est alors l'incertitude sur le volume des cylindres produits par cette usine ?

On peut estimer cette incertitude en utilisant une approximation affine de la fonction volume $V(d, \ell) = \frac{\pi d^2 \ell}{4}$. En effet, d'après les formules ci-dessus, on a

$$V(d_0 + \Delta d, \ell_0 + \Delta \ell) - V(d_0, \ell_0) \simeq \frac{\partial V}{\partial d}(d_0, \ell_0).\Delta d + \frac{\partial V}{\partial \ell}(d_0, \ell_0).\Delta \ell$$

si Δd et $\Delta \ell$ sont petits. On calcule

$$\frac{\partial V}{\partial d}(d_0, \ell_0) = \frac{\pi d_0 \ell_0}{2} = 2 \frac{V(d_0, \ell_0)}{d_0} \quad \text{et} \quad \frac{\partial V}{\partial \ell}(d_0, \ell_0) = \frac{\pi d_0^2}{4} = \frac{V(d_0, \ell_0)}{\ell_0},$$

et on applique la formule ci-dessus avec $\Delta d = \frac{5}{10000}d_0$ (c'est-à-dire Δd inférieur en valeur absolue à 0,05% de d_0) et $|\Delta \ell| \leq \frac{1}{1000}\ell_0$ (c'est-à-dire $\Delta \ell$ inférieur en valeur absolue à 0,1% de ℓ_0). On obtient

$$|V(d_0 + \Delta d, \ell_0 + \Delta \ell) - V(d_0, \ell_0)| \lesssim 2 \frac{V(d_0, \ell_0)}{d_0} \frac{5}{10000}d_0 + \frac{V(d_0, \ell_0)}{\ell_0} \frac{1}{1000}\ell_0 = \frac{2}{1000}V(d_0, \ell_0).$$

Ainsi, si on connaît d à 0,05% près, et ℓ à 0,1% près, alors on connaît donc $V(d, \ell)$ à 0,2% près. L'usine peut garantir le volume de ces cylindres à environ 0,2% près. Notons que ce raisonnement donne une très bonne estimation de l'incertitude : la véritable incertitude est de 0,20001% près.

Plan tangent. Le graphe d'une fonction affine est un plan. Le plan tangent au graphe d'une fonction en un point (x_0, y_0) est (plus ou moins par définition) le plan qui approche le mieux le graphe de f au voisinage de (x_0, y_0) . On a donc naturellement la proposition suivante :

Proposition 6.2.2 (Équation du plan tangent). *Le plan d'équation*

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

(c'est-à-dire le graphe de la meilleure approximation affine de f au voisinage de (x_0, y_0)) est le plan tangent au graphe de f au point $(x_0, y_0, f(x_0, y_0))$.

Dans \mathbb{R}^3 , le plan d'équation $z = ax + by + c$ est horizontal si et seulement si $a = b = 0$ (pour le voir, le plus simple est de se rappeler que le vecteur de coordonnées $(a, b, 1)$ est orthogonal au plan considéré (voir le chapitre 4), et de noter que ce vecteur est vertical si et seulement si $a = b = 0$). En combinant, la proposition 6.2.2 avec cette observation, on voit que :

Proposition 6.2.3 (Condition pour avoir un plan tangent horizontal). *Pour que le plan tangent au graphe de f en un point $(x_0, y_0, f(x_0, y_0))$ soit horizontal, il faut et il suffit que les deux dérivées partielles de f s'annulent au point (x_0, y_0) . On dit alors que (x_0, y_0) est un point critique de f .*

Remarque. Dans la formule de Taylor-Young (voir ci-dessus), si l'on fixe la valeur de la variable y égale à y_0 , on retrouve le DL de la fonction partielle :

$$f(x, y_0) = f(x_0, y_0) + (x - x_0) \frac{\partial f}{\partial x}(x_0, y_0) + |x - x_0| \varepsilon(x - x_0, 0).$$

Ceci peut s'interpréter graphiquement : on dessine le graphe et son plan tangent en un point ; l'intersection de ce dessin avec le plan vertical (Oxz) donne le graphe de la fonction partielle et sa tangente. On a bien sûr un phénomène similaire si on fixe la valeur de la variable x égale à x_0

6.2.4 Pour les fonctions de trois variables

On peut définir des notions de *développement limité à l'ordre 1*, de *meilleure approximation affine*, etc. pour une fonction de trois, quatre, ... variables. Par exemple, si f une fonction de trois variables (x, y, z) , un *développement limité de f à l'ordre 1 en (x_0, y_0, z_0)* , z_0 est une expression de la forme

$$f(x, y, z) = a + b(x - x_0) + c(y - y_0) + d(z - z_0) + \varepsilon(x - x_0, y - y_0, z - z_0)$$

où a, b, c, d sont des nombres et ε est une fonction de trois variables qui tend vers 0 en $(0, 0, 0)$. La seule différence avec le cas à deux variables est qu'on ne peut pas visualiser le graphe d'une

fonction à trois variables. En particulier, on ne peut pas dessiner le “plan” tangent à ce graphe (qui n'est d'ailleurs pas un plan, mais ce qu'on appelle un *hyperplan* de dimension 3).

On a également une formule de Taylor pour une fonction de trois, quatre,... variables. Pour une fonction de trois variables, cette formule s'écrit

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

ou encore

$$f(x_0 + h, y_0 + k, z_0 + \ell) = f(x_0, y_0, z_0) + \frac{\partial f}{\partial x}(x_0, y_0, z_0).h + \frac{\partial f}{\partial y}(x_0, y_0, z_0).k + \frac{\partial f}{\partial z}(x_0, y_0, z_0).\ell \\ + \|(h, k, \ell)\|\varepsilon(h, k, \ell).$$

6.3 Dérivation composée

Considérons une fonction f de plusieurs variables (disons deux variables pour fixer les idées). Il y a deux façons de composer f avec une autre fonction : à l'arrivée, ou bien au départ.

6.3.1 Composition à l'arrivée

Considérons tout d'abord le cas où l'on compose f à l'arrivée. Puisque f arrive dans \mathbb{R} , on doit la composer par une fonction g qui part de \mathbb{R} , c'est-à-dire par une fonction d'une seule variable. On ne s'intéressera qu'au seul cas où g arrive aussi dans \mathbb{R} . On obtient ainsi une fonction de deux variables $g \circ f$. Pour calculer les dérivées partielles de $g \circ f$, il suffit d'appliquer la formule de dérivation des fonctions d'une variable.

Exemple. Prenons $f(x, y) = x^2y^3$, soit g une fonction dérivable de \mathbb{R} dans \mathbb{R} , calculons les dérivées partielles de la fonction $g \circ f(x, y) = g(x^2y^3)$. Lorsque y est fixé égal à y_0 , on a la fonction partielle $x \mapsto g(x^2y_0^3)$, qui est une fonction d'une seule variable. D'après la formule de dérivation composée en une variable, on a

$$\frac{\partial g \circ f}{\partial x}(x_0, y_0) = 2x_0g'(x_0^2y_0^3).$$

La formule générale, si f est une fonction de deux variables x, y et g est une fonction d'une variable, est

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

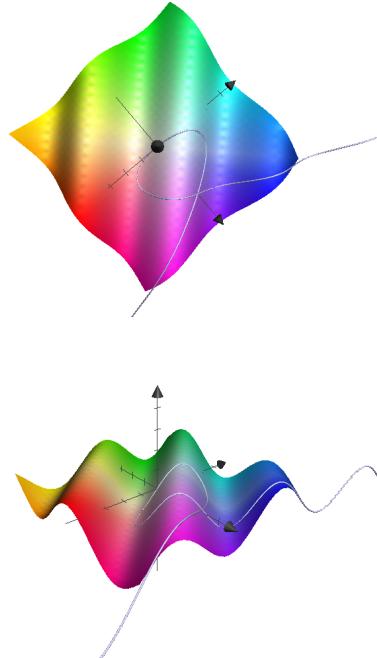
On n'a pas besoin d'apprendre cette formule, puisque c'est juste une application de la formule de dérivation composée en une variable, comme on vient de le voir sur un exemple. Évidemment, on a une formule similaire pour la seconde dérivée partielle de $g \circ f$. Et on a aussi une formule similaire dans le cas où f est une fonction de trois variables x, y, z :

$$\frac{\partial g \circ f}{\partial x}(x_0, y_0, z_0) = g'(f(x_0, y_0)) \frac{\partial f}{\partial x}(x_0, y_0)$$

6.3.2 Composition au départ

Le cas où l'on compose f au départ est un peu plus compliqué. Puisque f part de \mathbb{R}^2 (rappelons que c'est une fonction de deux variables), on va devoir composer par une fonction qui arrive dans \mathbb{R}^2 , c'est-à-dire par une courbe paramétrée plane $M : t \mapsto M(t) = (x(t), y(t))$.

Exemple. Un surfeur se déplace sur les vagues dont la surface suit l'équation $z = f(x, y) = \sin(x+y)$. La position du surfeur en fonction du temps (longitude x et latitude y), est donnée par la courbe paramétrée $t \mapsto M(t) = (x(t), y(t)) = (t^2, t^3 - 3t)$. L'altitude du surfeur en fonction du temps est alors donnée par la fonction composée $f \circ M : t \mapsto f \circ M(t) = f(x(t), y(t))$.



Soit $M : t \mapsto M(t) = (x(t), y(t))$ une courbe paramétrée plane. La fonction composée $f \circ M : t \mapsto f(x(t), y(t))$ est alors une fonction de \mathbb{R} dans \mathbb{R} .

Proposition 6.3.1. *On a*

$$(f \circ M)'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t))$$

Remarque. Le cas particulier le plus simple est $(x(t), y(t)) = (t, 0)$. Alors $f \circ M(t) = f(t, 0) = f_1(t)$ est la première fonction partielle de f en $(0, 0)$, et on retrouve simplement la définition

de la première dérivée partielle :

$$f'_1(t) = \frac{\partial f}{\partial x}(t, 0).$$

Supposons maintenant que $y(t)$ est la fonction nulle ; on a alors $f \circ M(t) = f(x(t), 0) = f_1(x(t))$, qui est une fonction composée de deux fonctions *d'une seule variable* ; on retrouve alors la formule de dérivation en une variable,

$$(f \circ M)'(t) = (f_1 \circ x)'(t) = x'(t) \cdot f'_1(x(t)) = x'(t) \frac{\partial f}{\partial x}(x(t), 0).$$

Exercice. On reprend le surfeur de tout à l'heure.

1. Quelle est l'altitude du surfeur en fonction du temps ?
2. Quel est le vecteur vitesse au temps t ?
3. Quel angle fait, au temps t , l'axe de sa planche avec l'horizontale ? (on admet que la planche est tangente aux vagues).

Solution. Puisque le surfeur se déplace sur les vagues, on a $z(t) = f(x(t), y(t)) = \sin(t^2 + t^3 - 3t)$. Le vecteur vitesse a trois composantes, $(x'(t), y'(t), z'(t))$, avec $x'(t) = 2t$, $y'(t) = 3t^2 - 3$ et $z'(t) = (2t + 3t^2 - 3) \cos(t^2 + t^3 - 3t)$ (on peut calculer $z'(t)$ directement, ou bien en appliquant la formule de dérivation composée que l'on vient de voir). L'angle est donné par $\arctan\left(\frac{z'(t)}{\sqrt{x'(t)^2 + y'(t)^2}}\right)$.

Démonstration de la proposition ci-dessus. On va obtenir un développement limité de la fonction $(f \circ M)(t)$ à l'aide de développements limités des fonctions $x(t)$, $y(t)$ et $f(x, y)$. Comme la $f \circ M$ est une brave fonction de \mathbb{R} dans \mathbb{R} , on sait que si elle admet un développement limité en un t_0 , alors elle est dérivable, et le développement limité donne la dérivée. On l'écrit en $t = 0$:

$$\begin{aligned} f(x(t), y(t)) &= f(x(0) + \underbrace{tx'(0) + t\varepsilon_1(t)}_h, y(0) + \underbrace{ty'(0) + t\varepsilon_2(t)}_k) \\ &= f(x(0), y(0)) + h \frac{\partial f}{\partial x}(x(0), y(0)) + k \frac{\partial f}{\partial y}(x(0), y(0)) + \|(h, k)\| \varepsilon(h, k) \\ &= \dots (\text{remplacer } h \text{ et } k \text{ par leur expression ; puis regrouper les termes}) \\ &= f(x(0), y(0)) + t \left(x'(0) \frac{\partial f}{\partial x}(x(0), y(0)) + y'(0) \frac{\partial f}{\partial y}(x(0), y(0)) \right) + \text{reste} \end{aligned}$$

et on vérifie que le reste est quelquechose de petit devant t : il s'écrit $t\varepsilon_3(t)$ avec $\lim_{t \rightarrow 0} \varepsilon_3(t) = 0$. Conclusion de ce calcul : la fonction $t \mapsto f(x(t), y(t))$ admet un DL1 en $t = 0$, par conséquent elle est dérivable, et sa dérivée est le facteur de ' t ', c'est-à-dire ce qu'on voulait montrer. \square

Test. Réécrire la preuve de la proposition dans le cas particulier d'une fonction $g(t) = f(e^t, \sqrt{1+t})$, où f est une fonction de deux variables (de classe C^1).

6.4 Application : gradient et lignes de niveau

Dans tout ce paragraphe, f est une fonction de deux variables (x, y) définie au voisinage d'un point (x_0, y_0) . La formule de dérivation composée, vue au paragraphe précédent, peut s'écrire comme un produit scalaire : en effet, $x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t))$ est le produit scalaire des deux vecteurs : le vecteur de coordonnées $(x'(t), y'(t))$ (c'est le vecteur vitesse de

la courbe $t \mapsto M(t)$) et le vecteur de coordonnées $(\frac{\partial f}{\partial x}(x(t), y(t)), \frac{\partial f}{\partial y}(x(t), y(t)))$. Ce qui nous conduit à la définition du gradient.

Définition (Gradient). *Le gradient de f en (x_0, y_0) , noté $\overrightarrow{\text{Grad}}_{(x_0, y_0)} f$ ou $\vec{\nabla}_{x_0, y_0} f$, est le vecteur dont les coordonnées sont les dérivées partielles de f :*

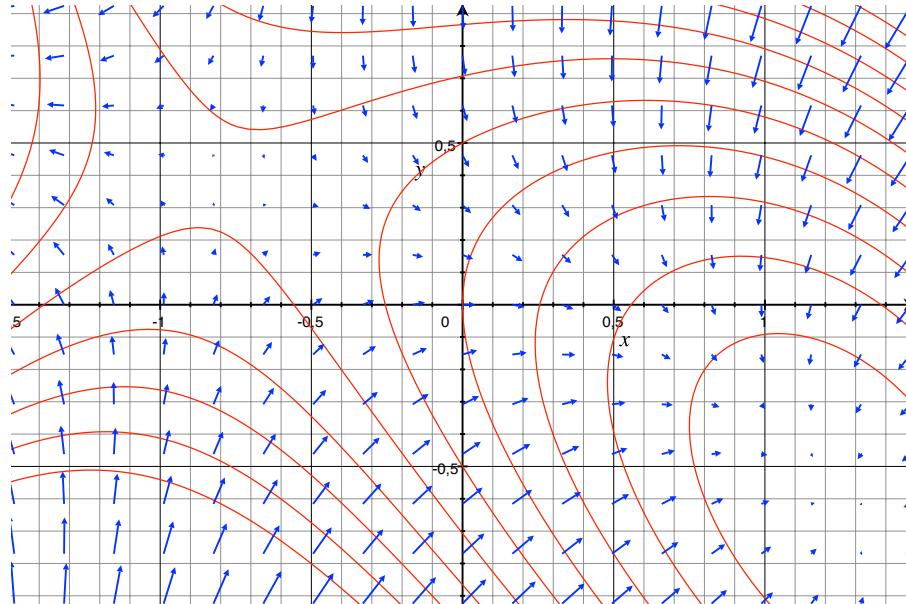
$$\overrightarrow{\text{Grad}}_{(x_0, y_0)} f = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$$

Proposition 6.4.1. *Soit $M : \mathbb{R} \mapsto \mathbb{R}^2$ une courbe paramétrée plane. On note $\overrightarrow{OM}'(t)$ le vecteur vitesse de M à l'instant t . La formule de dérivation composée de la section précédente s'écrit*

$$(f \circ M)'(t) = \overrightarrow{\text{Grad}}_{M(t)} f \cdot \overrightarrow{OM}'(t).$$

On constate expérimentalement que le gradient est orthogonal aux lignes de niveau : ci-dessous, on a tracé simultanément, dans le plan (Oxy) (plan des variables), des lignes de niveau et des vecteurs gradient de la fonction “montagne”

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



Pour exprimer précisément ceci, il faut supposer que la ligne de niveau peut être décrite comme une courbe paramétrée.

Proposition 6.4.2 (Gradient et lignes de niveaux). *Soit $t \mapsto M(t) = (x(t), y(t))$ un paramétrage d'une partie d'une ligne de niveau de f . Alors le vecteur vitesse $\overrightarrow{OM}'(t)$ est orthogonal au gradient de f au point $M(t)$.*

Autrement dit, le gradient de f est orthogonal aux lignes de niveaux de f .

Démonstration. On s'intéresse à la fonction $f \circ M : t \mapsto f(M(t)) = f(x(t), y(t))$. Dire que $M(t)$ est contenu dans une ligne de niveau de f , c'est dire que cette fonction est constante. Par suite, la dérivée de la fonction $f \circ M$ est identiquement nulle. Par ailleurs, la formule de dérivation composée nous donne

$$(f \circ M)'(t) = x'(t) \frac{\partial f}{\partial x}(x(t), y(t)) + y'(t) \frac{\partial f}{\partial y}(x(t), y(t)) = \overrightarrow{\text{Grad}}_{M(t)} f \cdot \overrightarrow{OM}'(t).$$

Par conséquent, le produit scalaire $\overrightarrow{\text{Grad}}_{M(t)} f \cdot \overrightarrow{OM}'(t)$ est nul, ce qui signifie exactement que le vecteur vitesse $\overrightarrow{OM}'(t)$ est orthogonal au gradient de f au point $M(t)$. La proposition est démontrée \square

Cette proposition nous permet de trouver une équation de la tangente T à la ligne de niveau de f passant par un point (x_0, y_0) donné, pour peu que le gradient de f au point (x_0, y_0) ne soit pas nul :

Proposition 6.4.3 (Équation de la tangente à une ligne de niveau). *Soit (x_0, y_0) un point du domaine de définition de f . On suppose que le gradient de f au point (x_0, y_0) n'est pas nul. Alors la tangente à la ligne de niveau de f passant par le point (x_0, y_0) (existe et) admet pour équation*

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

Démonstration. Soit T la tangente en (x_0, y_0) à la ligne de niveau de f passant par ce point. D'après la proposition 6.4.2, le vecteur $\overrightarrow{\text{Grad}}_{(x_0, y_0)} f = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0) \right)$ est orthogonal à T . Par ailleurs, on sait que T passe par le point (x_0, y_0) . Pour trouver une équation de la droite T , il suffit donc d'appliquer la formule (re)vue au chapitre 4. \square

La proposition 6.4.2 nous dit (en gros) que le gradient de f ne pointe jamais dans une direction où f reste constante. En fait, on a beaucoup mieux : le gradient de f pointe dans la direction où f augmente le plus vite. C'est ce qu'affirme la proposition suivante :

Proposition 6.4.4 (Gradient et plus grande pente). *Parmi les courbes paramétrées $t \mapsto M(t)$ qui passent à $t = 0$ par le point $M_0 = (x_0, y_0)$ avec une vitesse 1 en ce point, les courbes pour lesquelles dérivée de $z(t) = f(M(t))$ en $t = 0$ est la plus grande sont les courbes dont le vecteur vitesse en $\overrightarrow{OM}'(0)$ est parallèle au gradient de f et de même sens.*

Autrement dit, le gradient $\overrightarrow{\text{Grad}}_{M_0} f$ indique la direction de plus grande pente du graphe de f , c'est-à-dire la direction dans laquelle f augmente le plus vite.

Démonstration. Soit $t \mapsto M(t) = (x(t), y(t))$ une courbe paramétrée qui passe à l'instant $t = 0$ par le point $M_0 = (x_0, y_0)$ avec une vitesse 1 en ce point. Soit $\overrightarrow{OM}'(0) = (x'(0), y'(0))$ le vecteur vitesse de cette courbe en $t = 0$ (c'est-à-dire au point M_0) ; on a donc $\|\overrightarrow{OM}'(0)\| = 1$.

D'après la formule de dérivation composée, on a

$$\begin{aligned}
z'(t) &= x'(t) \frac{\partial f}{\partial x}(x_0, y_0) + y'(t) \frac{\partial f}{\partial y}(x_0, y_0) \\
&= \overrightarrow{OM}'(0) \cdot \overrightarrow{\text{Grad}}_{M_0} f \\
&= \left\| \overrightarrow{OM}'(0) \right\| \left\| \overrightarrow{\text{Grad}}_{(x_0, y_0)} f \right\| \cos \alpha \\
&= \left\| \overrightarrow{OM}'(0) \right\| \left\| \overrightarrow{\text{Grad}}_{(x_0, y_0)} f \right\| \cos \alpha
\end{aligned}$$

où α est l'angle entre le vecteur vitesse $\overrightarrow{OM}'(0)$ et $\overrightarrow{\text{Grad}}_{M_0} f$. La valeur de l'expression considérée est donc maximale quand $\cos(\alpha)$ l'est, c'est-à-dire lorsque $\cos \alpha = 1$, c'est-à-dire $\alpha = 0$, c'est-à-dire quand $\overrightarrow{OM}'(0)$ est colinéaire au gradient, et de même sens. Ceci termine la preuve. \square

Exemple. On reprend la montagne d'équation $f(x, y) = -\frac{x^3}{3} - xy - y^2 + x + \frac{3}{2}$. En exemple, on avait calculé les dérivées partielles de f au point $(1, 1/2)$. On a donc $\overrightarrow{\text{Grad}}_{(1, \frac{1}{2})} f = (-1/2, -2)$. La ligne de plus grande pente est donc vers le Sud-Sud-Ouest. Une bille posée en ce point roulerait vers le Nord-Nord-Ouest.

Remarque. En physique, de nombreuses forces s'obtiennent comme gradients d'une fonction appelée *potentiel* : la force d'attraction gravitationnelle qui s'exerce sur un objet massif est proportionnelle au gradient du potentiel gravitationnel (calculé au point où se trouve l'objet), la force électrique qui s'exerce sur un objet chargé est proportionnelle au gradient du potentiel électrique.... La proposition ci-dessus indique que de telles forces attirent toujours les objets dans les directions où le potentiel décroît le plus vite : les objets sont ainsi attirés vers les zones où les potentiels sont minimaux. Par exemple, vous savez que la valeur du potentiel gravitationnel terrestre en un point M donné est proportionnel à la distance entre M et le centre de la Terre, et vous savez également fort bien que les objets massifs tendent à "tomber", c'est-à-dire à se rapprocher du centre de la Terre.

Chapitre 7

Extrema d'une fonction de plusieurs variables

7.1 Retour sur les fonctions d'une variable

7.1.1 Extrema globaux et locaux

On considère une fonction d'une variable f de domaine de définition D , et un point x_0 de D

Définition (Maximum global, maximum local).
– *On dit que f atteint son maximum absolu au point x_0 si $f(x) \leq f(x_0)$ pour tout $x \in D$*
– *On dit que f admet un maximum local au point x_0 si l'inégalité $f(x) \leq f(x_0)$ est vérifiée pour tout $x \in D$ assez proche de x_0 .*

Remarques.

- L'expression “pour tout $x \in D$ assez proche de x_0 ” signifie plus précisément “il existe $\varepsilon > 0$ tel que $f(x) \leq f(x_0)$ pour tout $x \in [x_0 - \varepsilon, x_0 + \varepsilon] \cap D$ ”.
- En pratique, f admet un maximum local au point x_0 si, dans le tableau de variation de f , on voit une flèche qui monte jusqu'en $f(x_0)$, puis une flèche qui descend. Cependant, cette caractérisation qui nous intéresse pas ici, car elle n'a aucune chance de se généraliser aux fonctions de plusieurs variables (on rappelle que parler d'une fonction de deux variables croissante ou décroissante n'a aucun sens, puisque les points de \mathbb{R}^2 ne sont pas ordonnés).
- Un maximum absolu est *a fortiori* un maximum local.

On définit de même les notions de minimum absolu et de minimum local (il suffit de changer les “ $f(x) \leq f(x_0)$ ” en “ $f(x) \geq f(x_0)$ ”). Extremum signifie minimum ou maximum.

7.1.2 Points critiques (les suspects !)

On va se concentrer sur la recherche des extrema *locaux* d'une fonction donnée. Ceci va se faire en deux étapes. On suppose maintenant que f est définie au voisinage¹ de x_0 . On suppose aussi (comme toujours) que f est dérivable.

1. Ceci signifie qu'il existe un nombre $\epsilon > 0$ — que l'on doit imaginer petit — tel que l'intervalle $]x_0 - \epsilon, x_0 + \epsilon[$ est contenu dans l'ensemble de définition D . Moins formellement, ceci signifie que x_0 n'est pas une borne de D . On veut par exemple éviter la situation où $D =]-1, 0] \cup]1, +\infty[$ et $x_0 = 0$.

Définition (Point critique). *On dit que x_0 est un point critique de f si $f'(x_0) = 0$.*

Remarque. Géométriquement, x_0 est un point critique si la tangente au graphe de f au point $(x_0, f(x_0))$ est horizontale.

Proposition 7.1.1 (Principe de Fermat). *Si f admet un minimum local ou un maximum local au point x_0 alors x_0 est un point critique de f*



Pierre de Fermat (??-1655)

Preuve de la proposition 7.1.1. Notons

$$T(h) = \frac{f(x_0 + h) - f(x_0)}{h}.$$

Alors, par définition de la dérivée, on a :

$$f'(x_0) = \lim_{h \rightarrow 0} T(h).$$

Supposons maintenant que x_0 est un minimum local de f . Alors on a $f(x) \geq f(x_0)$ pour tout x assez proche de x_0 . Autrement dit $f(x_0 + h) \geq f(x_0)$ pour tout h assez petit. Cette inégalité montre que $T(h)$ a le même signe que h . Donc

$$\lim_{h \rightarrow 0, h < 0} T(h) \leq 0 \quad \text{et} \quad \lim_{h \rightarrow 0, h > 0} T(h) \geq 0.$$

Mais on sait que la limite de $T(h)$ en 0 existe, donc

$$\lim_{h \rightarrow 0, h < 0} T(h) = \lim_{h \rightarrow 0, h > 0} T(h).$$

La seule possibilité est donc que

$$\lim_{h \rightarrow 0, h < 0} T(h) = \lim_{h \rightarrow 0, h > 0} T(h) = 0$$

Et donc $f'(x_0) = 0$. Le cas où x_0 est un maximum local se traite de même. \square

Autrement dit, les extrema locaux sont à chercher parmi les points critiques, c'est-à-dire les points où la dérivée s'annule. Quand on cherche des extrema, les "suspects" sont les points critiques. Attention : tous les extrema sont des points critiques, mais (comme le montre le dernier exemple ci-dessous) tous les points critiques ne sont pas des extrema !

Exemples à avoir en tête :

- 0 est un point critique de $1 + x^2$ (et un min local) ;
- 0 est un point critique de $1 - x^2$ (et un max local) ;
- 0 est un point critique de $1 + x^3$ (et ni min ni max local : en particulier, la réciproque de la proposition est fausse, tous les suspects ne sont pas forcément coupables !).

7.1.3 Condition d'ordre 2 (les coupables !)

Une fois qu'on a un point critique x_0 (suspecté, sans preuve, d'être un extremum local), comment faire pour savoir s'il s'agit d'un minimum local (comme $1 + x^2$), d'un maximum local (comme $1 - x^2$) ou d'un innocent (comme $1 + x^3$) ? On fait un développement limité à l'ordre 2 (via la formule de Taylor-Young). On écrit :

$$f(x_0 + h) = f(x_0) + 0 + h^2 \frac{f''(x_0)}{2} + h^2 \varepsilon(h) = f(x_0) + h^2 \left(\frac{f''(x_0)}{2} + \varepsilon(h) \right).$$

- si $f''(x_0) > 0$, on voit que $f(x_0 + h) > f(x_0)$ pour h assez proche de 0 (en effet, puisque $\varepsilon(h)$ tend vers 0 quand h tend vers 0, la quantité $\frac{f''(x_0)}{2} + \varepsilon(h)$ tend vers $\frac{f''(x_0)}{2}$; en particulier, $\frac{f''(x_0)}{2} + \varepsilon(h)$ est strictement positif pour h assez petit ; en utilisant le développement limité ci-dessus, on en déduit que $f(x_0 + h) > f(x_0)$) ;
- si $f''(x_0) < 0$, on voit que $f(x_0 + h) < f(x_0)$ pour h assez proche de 0 ;
- si $f''(x_0) = 0$, on ne peut rien dire (il faudrait faire un développement limité à l'ordre 3 ou plus...). On dit que le point critique est *dégénéré*.

En résumé :

Proposition 7.1.2 (Condition d'ordre 2). *Soit x_0 un point critique de f .*

- *Si $f''(x_0) > 0$, alors c'est un minimum local ;*
- *si $f''(x_0) < 0$, alors c'est un maximum local ;*
- *si $f''(x_0) = 0$, alors on ne peut rien dire sans plus d'information.*

Exemple. On considère la fonction $f : \mathbb{R} \rightarrow \mathbb{R}$ définie par $f(x) = x^3 (1 - \frac{3}{5}x^2)$. Calculer les deux premières dérivées f' et f'' de f . Vérifier que f a trois points critiques : un minimum local, un maximum local et un point critique dégénéré.

7.2 Notion d'extrema pour les fonctions de plusieurs variables

Pour les fonctions de plusieurs variables, on va suivre la même méthode : il y a encore une notion de point critique, et une condition d'ordre 2 (un peu plus compliquée). Il y aura un nouveau type de point qui n'existe pas en une variable (le point selle).

Définition (Maximum global, maximum local). *Soit f une fonction de deux variables x, y , d'ensemble de définition D . Soit (x_0, y_0) un point de D .*

- *On dira que f atteint son maximum au point (x_0, y_0) si, pour tout $(x, y) \in D$, on a $f(x, y) \leq f(x_0, y_0)$.*
- *On dira que f admet un maximum local au point (x_0, y_0) si l'inégalité $f(x, y) \leq f(x_0, y_0)$ est vérifiée pour tout $(x, y) \in D$ assez proche de (x_0, y_0) , c'est-à-dire pour tout (x, y) dans un petit disque centré en (x_0, y_0) .*

Bien sûr, on définit de façon analogue les minima et minima locaux. Et bien sûr, on a une définition strictement analogue pour une fonction de trois, quatre, etc. variables

7.3 Points critiques des fonctions de plusieurs variables

La notion de point critique définie ci-dessus pour les fonctions d'une variable se généralise facilement aux fonctions de plusieurs variables.

Définition (Point critique d'une fonction de deux variables). *Soit f une fonction de deux variables x, y , et (x_0, y_0) un point du domaine de définition de f . On dit que (x_0, y_0) est un point critique de f si les deux dérivées partielles de f s'annulent en ce point :*

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

Géométriquement, ceci signifie que le plan tangent au graphe de f au point $(x_0, y_0, f(x_0, y_0))$ est horizontal.

Définition (Point critique d'une fonction de trois variables). *Soit f une fonction de trois variables x, y, z , et (x_0, y_0, z_0) un point du domaine de définition de f . On dit que (x_0, y_0, z_0) est un point critique de f si les trois dérivées partielles de f s'annulent en ce point :*

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

Comme pour les fonctions d'une variable, tout extremum d'une fonction de plusieurs variables est un point critique ; c'est le *principe de Fermat* :

Proposition 7.3.1 (Principe de Fermat). *Soit f une fonction de plusieurs variables, et M_0 un point du domaine de définition de f . Si f admet un minimum local ou un maximum local au point M_0 , alors ce point est un point critique.*

Démonstration. Pour simplifier, on considère une fonction de deux variables, et on se place au point $M_0 = (0, 0)$. Supposons que ce point soit un maximum local de f : on a $f(x, y) \leq f(0, 0)$ pour tout (x, y) assez proche de $(0, 0)$. Alors ceci est *a fortiori* vrai lorsqu'on se restreint aux valeurs nulles de y : pour tout x assez proche de 0, on a $f(x, 0) \leq f(0, 0)$. Autrement dit, la fonction partielle $\varphi : x \mapsto f(x, 0)$ admet un maximum local en $x = 0$. Le théorème correspondant en une variable nous dit que $\varphi'(0) = 0$, mais par définition $\varphi'(0)$ est la dérivée partielle de f en $(0, 0)$ par rapport à la variable x . Bien sûr, le même raisonnement marche pour la dérivée par rapport à y . \square

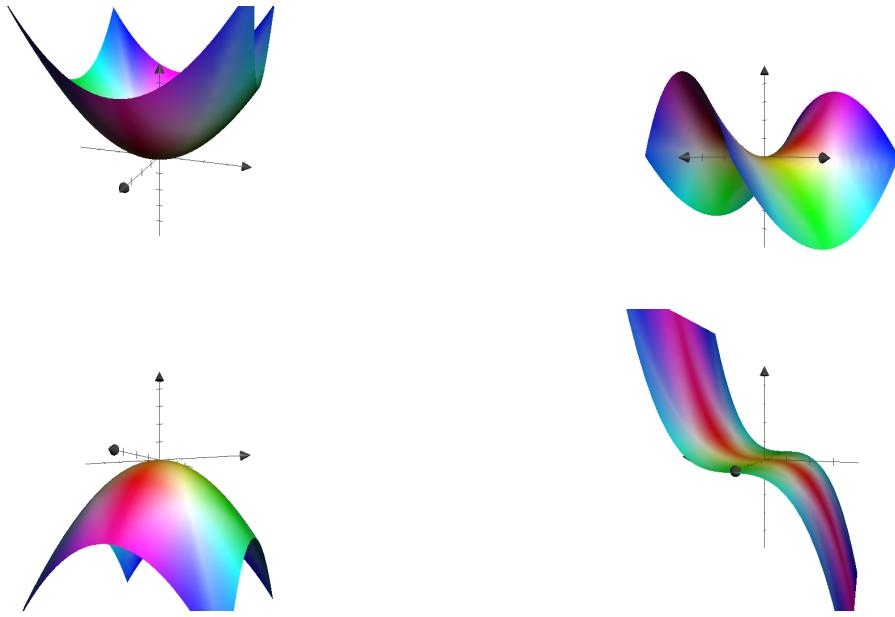
Ainsi, les **extrema d'une fonction de plusieurs variables sont à rechercher parmi les points critiques**. Mais attention, comme pour les fonctions d'une variable, tout extremum est un point critique, mais tout point critique n'est pas un extremum.

Exemple. Il est bon d'avoir les exemples suivants en tête.

- Le point $(0, 0)$ est un point critique de la fonction $f(x, y) = x^2 + y^2$. En fait, c'est un minimum local, et même un minimum absolu de f (en effet, $f(0, 0) = 0$ et $f(x, y) > 0$ pour tout $(x, y) \neq (0, 0)$). Le graphe de f est représenté ci-dessous en haut à gauche.
- Le point $(0, 0)$ est un point critique de la fonction $g(x, y) = -x^2 - y^2$. En fait, c'est un maximum local, et même un maximum absolu de g . Le graphe de g est représenté ci-dessous en bas à gauche.

— Le point $(0, 0)$ est un point critique de la fonction $h(x, y) = x^2 - y^2$. Cependant, ce n'est ni un minimum local, ni un maximum local. En fait, $(0, 0)$ est un "minimum dans la direction des x " (plus précisément, 0 est un minimum pour la fonction partielle $x \mapsto h(x, 0)$) et un "maximum dans la direction des y " (plus précisément, 0 est un maximum pour la fonction partielle $y \mapsto h(0, y)$); quand on tient compte des deux directions, ce n'est donc ni un minimum, ni un maximum. On dira que $(0, 0)$ est un *point selle* pour h . Le graphe de h est représenté ci-dessous en haut à droite.

— Le point $(0, 0)$ est un point critique de la fonction $i(x, y) = x^3$. Cependant, ce n'est ni un minimum local, ni un maximum local (c'est exactement la même situation que celle du point 0 pour la fonction d'une variable $x \mapsto x^3$). On dira que $(0, 0)$ est un *point critique dégénéré* pour la fonction i . Le graphe de i est représenté ci-dessous en bas à droite.



Exemple de recherche de points critiques. Soit $f(x, y) = x^4 + y^4 - 4xy$. On a

$$\frac{\partial f}{\partial x}(x, y) = 4x^3 - 4y \text{ et } \frac{\partial f}{\partial y}(x, y) = 4y^3 - 4x.$$

Pour que ces deux quantités s'annulent, il faut : $x^3 = y$ et $y^3 = x$. Ces égalités impliquent $y^9 = y$ et $x^9 = x$, et donc

$$x = -1, 0 \text{ ou } 1 \text{ et } y = -1, 0 \text{ ou } 1.$$

En tenant compte des contraintes $x^3 = y$ et $y^3 = x$, on voit que les seules possibilités sont $(x, y) = (0, 0)$ ou $(1, 1)$ ou $(-1, -1)$. On vérifie que les deux dérivées partielles s'annulent pour ces valeurs de (x, y) . Il y a donc trois points critiques : $(0, 0)$, $(1, 1)$, $(-1, -1)$.

7.4 Formule de Taylor-Young d'ordre 2 pour les fonctions de plusieurs variables

On sait maintenant que les extrema d'une fonction f de plusieurs variables sont à rechercher parmi les points critiques de f . Il nous reste à trouver une méthode pour déterminer si un point critique donné est un minimum, un maximum, ou ni l'un ni l'autre. Comme pour les fonctions d'une variable, on va utiliser la formule de Taylor-Young à l'ordre deux. Nous commençons par traiter le cas des fonctions de deux variables, avant de généraliser aux fonctions de trois variables ou plus.

7.4.1 Pour les fonctions de deux variables

Soit f une fonction de deux variables x, y . Une telle fonction admet *a priori* quatre dérivées partielles d'ordre 2,

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

Comme toujours dans ces notes, on suppose que ces dérivées partielles existent et sont continues (on dit que f est *de classe C^2*).

La proposition suivante affirme que deux des quatre dérivées partielles d'ordre 2 de f coïncident.

Proposition 7.4.1 (Lemme de Schwarz). *On a l'égalité :*

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

Remarque. Cette proposition est vraie *quelle que soit la fonction f* (de classe C^2). Quand vous calculez les dérivées partielles d'ordre 2 d'une fonction, si vous ne trouvez pas le même résultat pour $\frac{\partial^2 f}{\partial y \partial x}$ et pour $\frac{\partial^2 f}{\partial x \partial y}$, c'est que vous vous êtes trompé quelque part !

Nous pouvons maintenant énoncer la formule de Taylor pour f :

Théorème 7.4.2 (Formule de Taylor à l'ordre 2 pour les fonctions de deux variables). *Soit (x_0, y_0) un point au voisinage duquel f est définie². Notons*

$$\begin{aligned}a &= \frac{\partial f}{\partial x}(x_0, y_0) & b &= \frac{\partial f}{\partial y}(x_0, y_0) \\ r &= \frac{\partial^2 f}{\partial x^2}(x_0, y_0) & s &= \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) & t &= \frac{\partial^2 f}{\partial y^2}(x_0, y_0).\end{aligned}$$



Hermann Amandus Schwarz (1843-1921)

Alors f admet un développement limité à l'ordre 2 au point (x_0, y_0) qui s'écrit :

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + ah + bk + \frac{1}{2} (rh^2 + 2shk + tk^2) + \|(h, k)\|^2 \varepsilon(h, k)$$

avec $\varepsilon(h, k) \rightarrow 0$ quand $(h, k) \rightarrow (0, 0)$ (on rappelle que $\|(h, k)\|^2 = h^2 + k^2$).

7.4.2 Pour les fonctions de trois variables ou plus

Tout ce qui précède s'adapte au fonctions de trois variables ou plus... le seul problème est qu'il y a beaucoup de dérivées partielles à calculer et que la formule de Taylor est très lourde à écrire.

Considérons une fonction f de n variables x_1, \dots, x_n . Une telle fonction admet a priori n^2 dérivées partielles d'ordre 2 (il faut considérer toutes les dérivées du type $\frac{\partial^2 f}{\partial x_i \partial x_j}$, on a n choix pour i et n choix pour j). Mais on a un analogue de la proposition 7.4.1 qui affirme que

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Il n'y a donc que $n(n+1)/2$ dérivées partielles réellement différentes. On peut alors écrire la formule de Taylor :

Théoreme 7.4.3 (Formule de Taylor à l'ordre 2 pour les fonctions de trois variables). *Soit $\underline{x}^0 = (x_1^0, \dots, x_n^0)$ un point au voisinage duquel f est définie. Alors f admet un développement limité à l'ordre 2 en p . En notant $\underline{h} = (h_1, \dots, h_n)$ et $\underline{x}^0 + \underline{h} = (x_1^0 + h_1, \dots, x_n^0 + h_n)$, ce développement limité s'écrit :*

$$f(\underline{x}^0 + \underline{h}) = f(\underline{x}^0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\underline{x}^0) h_i + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\underline{x}^0) h_i h_j + \|\underline{h}\|^2 \varepsilon(\underline{h})$$

avec $\varepsilon(\underline{h}) \rightarrow 0$ quand $\underline{h} \rightarrow 0$.

7.5 Types de points critiques pour les fonctions de plusieurs variables

Nous allons maintenant expliquer comment déterminer la nature (minimum, maximum, ou ni l'un ni l'autre) d'un point critique en utilisant les formules de Taylor vues au paragraphe précédent. Pour les fonctions de *deux* variables, il existe un critère simple ; que nous allons énoncer et démontrer. Pour les fonctions de *trois variables ou plus*, la situation est beaucoup plus complexe, et nous ne traiterons explicitement que des cas particuliers.

7.5.1 Pour les fonctions de deux variables

Soit f une fonction de deux variables x, y , et (x_0, y_0) un point critique de cette fonction. Notons

$$r = \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \quad s = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \quad t = \frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

Puisque (x_0, y_0) est un point critique, les dérivées partielles d'ordre 1 de f au point (x_0, y_0) sont nulles. La formule de Taylor vue au paragraphe précédent s'écrit donc

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{1}{2} (rh^2 + 2shk + tk^2) + \|(h, k)\|^2 \varepsilon(h, k)$$

$\varepsilon(h, k) \rightarrow 0$ quand $(h, k) \rightarrow (0, 0)$. Rappelons que (x_0, y_0) est un maximum (respectivement minimum) local si la quantité $f(x_0 + h, y_0 + k) - f(x_0, y_0)$ est négative (respectivement positive) lorsque le vecteur (h, k) est assez petit. D'après la formule ci-dessus, le signe de cette quantité dépend essentiellement du signe de $q(h, k) := rh^2 + 2shk + tk^2$. Pour comprendre ce qui ce passe commençons par examiner le cas particulier où $s = 0$.

Le cas particulier $s = 0$. On a alors $q(h, k) := rh^2 + tk^2$, et le signe de cette quantité est facile à déterminer :

- si $r > 0$ et $t > 0$, alors $q(h, k) > 0$ quel que soit $(h, k) \neq (0, 0)$. Le point (x_0, y_0) sera donc³ un minimum local.
- si $r < 0$ et $t < 0$, alors $q(h, k) < 0$ quel que soit $(h, k) \neq (0, 0)$. Le point (x_0, y_0) sera donc un maximum local.
- si r et t sont de signes différents, alors le signe de $q(h, k)$ dépend des valeurs relatives de h et k . Autrement dit, $q(h, k)$ prend à la fois des valeurs positives, et des valeurs négatives⁴. Le point (x_0, y_0) ne sera donc ni un minimum local, ni un maximum local.
On dira que c'est un point *selle*
- si r ou t est nul, alors $q(h, k)$ prend la valeur 0 pour certains vecteurs (h, k) non-nuls. On ne peut donc plus négliger le terme $\|(h, k)\|^2 \varepsilon(h, k)$, et on ne sait pas si (x_0, y_0) est un minimum, un maximum, ou ni l'un ni l'autre (pour le savoir, il faudrait faire un développement limité à l'ordre 3 ou plus...). On dira que (x_0, y_0) est un *point critique dégénéré*.

3. En fait, il faudrait démontrer que l'on peut négliger le terme $\|(h, k)\|^2 \varepsilon(h, k)$.

4. Même si on impose à (h, k) d'être petit, puisque ce sont les valeurs relatives de h et k qui comptent.

Le cas général. Nous avons considéré ci-dessus le cas particulier où $s = 0$. Ce cas se traite facilement parce que $q(h, k)$ est alors une combinaison linéaire des termes h^2 et k^2 , et qu'un carré est toujours positif. Pour traiter le cas général $s \neq 0$, nous allons essayer d'écrire $q(h, k)$ sous la forme d'une combinaison linéaire de carrés. C'est la méthode de Gauss.

- Si $r \neq 0$, on écrit

$$\begin{aligned} q(h, k) &= r \left(h^2 + \frac{2s}{r} hk + \frac{t}{r} k^2 \right) \\ &= r \left(\left(h + \frac{s}{r} k \right)^2 - \frac{s^2}{r^2} k^2 + \frac{t}{r} k^2 \right) \\ &= r \left(\left(h + \frac{s}{r} k \right)^2 - \frac{s^2 - rt}{r^2} k^2 \right) \end{aligned}$$

On en déduit que :

- si $r > 0$ et $s^2 - rt < 0$, alors $q(h, k) > 0$ quel que soit $(h, k) \neq (0, 0)$. Le point (x_0, y_0) sera donc un minimum local.
- si $r < 0$ et $s^2 - rt < 0$, alors $q(h, k) < 0$ quel que soit $(h, k) \neq (0, 0)$. Le point (x_0, y_0) sera donc un maximum local.
- si $s^2 - rt > 0$ alors le signe de $q(h, k)$ dépend des valeurs relatives de h et k . Le point (x_0, y_0) ne sera donc ni un minimum local, ni un maximum local. On dira que c'est un point *selle*.
- si $s^2 - rt = 0$, alors on ne peut pas conclure. On dira que (x_0, y_0) est un *point critique dégénéré*.

- Si $r = 0$ et $t \neq 0$, alors on écrit

$$q(h, k) = 2shk + tk^2 = t \left(\left(k + \frac{s}{t} h \right)^2 - \frac{s^2}{t^2} h^2 \right).$$

On en déduit que le signe de $q(h, k)$ dépend des valeurs relatives de h et k . Le point (x_0, y_0) ne sera donc ni un minimum local, ni un maximum local. On dira que c'est un point *selle*. On remarque que $r = 0$ implique $s^2 - rt > 0$.

- Enfin, si $r = t = 0$, alors on écrit

$$q(h, k) = shk = \frac{s}{2} ((h+k)^2 - (h-k)^2)$$

On en déduit que le signe de $q(h, k)$ dépend des valeurs relatives de h et k . Le point (x_0, y_0) ne sera donc ni un minimum local, ni un maximum local. On dira que c'est un point *selle*. On remarque que $r = t = 0$ implique $s^2 - rt > 0$.

On peut résumer tout cela dans la proposition suivante :

Proposition 7.5.1. *Rappelons que l'on a noté*

$$r = \frac{\partial^2 f}{\partial x^2}(x_0, y_0) \quad s = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) \quad t = \frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

- Si $s^2 - rt > 0$, alors le point (x_0, y_0) n'est ni un maximum, ni un minimum. On dit que c'est un point *selle*. Au voisinage de f , le graphe de f “ressemble” au graphe de la figure page 31, en haut à droite.

- Si $s^2 - rt < 0$ et $r < 0$, alors le point (x_0, y_0) est un maximum local. Au voisinage de f , le graphe de f “ressemble” au graphe de la figure page 31, en bas à gauche.
- Si $s^2 - rt < 0$ et $r > 0$, alors le point (x_0, y_0) est un minimum local. Au voisinage de f , le graphe de f “ressemble” au graphe de la figure page 31, en haut à gauche.
- Si $s^2 - rt = 0$, on ne peut pas conclure. On dit que (x_0, y_0) est un point critique dégénéré. Attention, on ne sait pas à quoi ressemble l'allure du graphe de f au voisinage de (x_0, y_0) .

Bien sûr, pour que toutes les assertions de l'énoncé ci-dessus aient un sens précis, il faudrait définir le mot “ressemble” ; nous ne le ferons pas.

Bilan. Pour trouver les extrema locaux d'une fonction f de deux variables, on procède comme suit.

1. On calcule les dérivées partielles d'ordre 1 de f , et on trouve les points critiques de f : on sait que tous les extrema locaux de f font partie des points critiques (remarque : en pratique, on n'y arrive pas toujours) ;
2. pour chaque point critique (x_0, y_0) ,
 - (a) on calcule les dérivées secondees r, s, t en (x_0, y_0) ,
 - (b) on calcule $s^2 - rt$; si cette quantité est nulle, on ne peut rien dire ; sinon, on conclut sur le type (minimum local, maximum local, ou point selle) du point critique (x_0, y_0) via la proposition 7.5.1.

7.5.2 Pour les fonctions de trois variables

Les travaux du mathématicien C. F. Gauss fournissent une procédure pour déterminer de manière systématique la nature (minimum, maximum, selle, dégénéré) d'un point critique d'une fonction de n variables pour n quelconque. Cette procédure est cependant assez complexe, et nous n'allons pas la décrire ici. Nous nous contenterons de traiter un cas particulier pour les fonctions de trois variables.



Carl Friedrich Gauss (1777-1855)

On considère donc une fonction f de trois variables x, y, z , et un point critique (x_0, y_0, z_0) de

f . On a donc

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

Nous allons nous restreindre au cas particulier où

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0, z_0) = \frac{\partial^2 f}{\partial x \partial z}(x_0, y_0, z_0) = \frac{\partial^2 f}{\partial y \partial z}(x_0, y_0, z_0) = 0.$$

Ainsi (en tenant compte du lemme de Schwarz), il n'y a que trois dérivées partielles d'ordre 2 de f qui sont (éventuellement) non-nulles au point (x_0, y_0, z_0) . On les note :

$$a := \frac{\partial^2 f}{\partial x^2}(x_0, y_0, z_0) = , \quad b := \frac{\partial^2 f}{\partial y^2}(x_0, y_0, z_0) = , \quad c := \frac{\partial^2 f}{\partial z^2}(x_0, y_0, z_0).$$

Le développement de Taylor à l'ordre 2 de f au point (x_0, y_0, z_0) s'écrit donc :

$$f(x_0 + h, y_0 + k, z_0 + \ell) = f(x_0, y_0, z_0) + \frac{1}{2} (ah^2 + bk^2 + cl^2) + \|(h, k, \ell)\|^2 \varepsilon(h, k, \ell)$$

avec $\varepsilon(h, k, \ell) \rightarrow 0$ quand $(h, k, \ell) \rightarrow (0, 0, 0)$. La nature du point critique (x_0, y_0, z_0) dépend donc du signe de la quantité $ah^2 + bk^2 + cl^2$. Ce signe est facile à déterminer puisque h^2, k^2 et ℓ^2 sont des quantités positives. Plus précisément, on distingue les cas suivants :

- si les trois nombres a, b, c sont strictement positifs, alors $ah^2 + bk^2 + cl^2 > 0$ pour tout vecteur $(h, k, \ell) \neq (0, 0, 0)$. On en déduit que $f(x_0 + h, y_0 + k, z_0 + \ell) - f(x_0, y_0, z_0) > 0$ dès que (h, k, ℓ) est proche de $(0, 0, 0)$. Par conséquent, le point critique (x_0, y_0, z_0) est un minimum local.
- si les trois nombres a, b, c sont strictement négatifs, alors $ah^2 + bk^2 + cl^2 < 0$ pour tout vecteur $(h, k, \ell) \neq (0, 0, 0)$. On en déduit que $f(x_0 + h, y_0 + k, z_0 + \ell) - f(x_0, y_0, z_0) < 0$ dès que (h, k, ℓ) est proche de $(0, 0, 0)$. Par conséquent, le point critique (x_0, y_0, z_0) est un maximum local.
- si les nombres a, b, c sont non-nuls, mais ne sont pas tous du même signe, alors le signe de la quantité $ah^2 + bk^2 + cl^2$ dépend des valeurs relatives de h, k et ℓ . On en déduit que le point critique (x_0, y_0, z_0) n'est ni un minimum, ni un maximum ; on dit que c'est un point selle.
- si l'un des nombres a, b, c est nul, alors $ah^2 + bk^2 + cl^2$ s'annule pour certains vecteurs (h, k, ℓ) , ce qui empêche déterminer le signe de $f(x_0 + h, y_0 + k, z_0 + \ell) - f(x_0, y_0, z_0)$ (car on ne peut plus négliger $\|(h, k, \ell)\|^2 \varepsilon(h, k, \ell)$). On ne peut donc pas conclure

La proposition suivante résume tout cela :

Proposition 7.5.2. *Rappelons que l'on se place dans le cas particulier où*

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0, z_0) = \frac{\partial^2 f}{\partial x \partial z}(x_0, y_0, z_0) = \frac{\partial^2 f}{\partial y \partial z}(x_0, y_0, z_0) = 0.$$

et que l'on note

$$a := \frac{\partial^2 f}{\partial x^2}(x_0, y_0, z_0) = , \quad b := \frac{\partial^2 f}{\partial y^2}(x_0, y_0, z_0) = , \quad c := \frac{\partial^2 f}{\partial z^2}(x_0, y_0, z_0).$$

Quatre cas peuvent alors se produire :

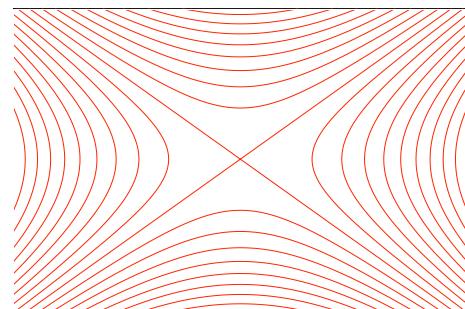
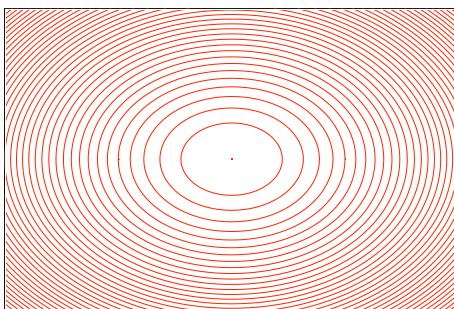
- si les trois nombres a, b, c sont strictement positifs, alors le point critique (x_0, y_0, z_0) est un minimum local ;
- si les trois nombres a, b, c sont strictement négatifs, alors le point critique (x_0, y_0, z_0) est un maximum local ;
- si les nombres a, b, c sont non-nuls, mais ne sont pas tous du même signe, alors le point critique (x_0, y_0, z_0) n'est ni un minimum, ni un maximum ; on dit que c'est un point selle ;
- si l'un des nombres a, b, c est nul, alors on ne peut pas conclure à la nature du point critique (x_0, y_0, z_0) ; on dit que c'est un point fixe dégénéré.

7.6 Allure des lignes de niveaux au voisinage d'un point critique

Le résultat suivant, qui est difficile à démontrer et que nous admettrons, est très utile pour repérer les points critiques d'une fonction de deux variables f et déterminer leur nature, lorsqu'on a un dessin des lignes de niveaux de f .

Théorème 7.6.1 (Lemme de Morse). *Soit f une fonction de deux variables et (x_0, y_0) un point critique de f , que l'on suppose non-dégénéré.*

- si (x_0, y_0) est un point minimum local ou un maximum local, alors, au voisinage de (x_0, y_0) , les lignes de niveaux de f “ressemblent” à celles de la figure ci-dessous à gauche. En particulier, les lignes de niveaux (au voisinage de (x_0, y_0)) sont des courbes fermées, emboîtées les unes dans les autres, qui entourent le point critique (x_0, y_0) .
- si (x_0, y_0) est un point selle, alors, au voisinage de (x_0, y_0) , les lignes de niveaux de f “ressemblent” à celles de la figure ci-dessous à droite. En particulier, la ligne de niveau $h_0 := f(x_0, y_0)$ est constituée (au voisinage de (x_0, y_0)) de deux courbes qui se croisent au point (x_0, y_0) , et divisent le plan en quatre quadrants. Les autres lignes de niveaux sont constituées de deux courbes non-fermées situées dans des quadrants opposés.



Bien sûr, pour que toutes les assertions de l'énoncé ci-dessus aient un sens précis, il faudrait définir le mot “ressemble” ; nous ne le ferons pas.

Exemple. La figure suivante représente quelques lignes de niveaux d'une fonction de deux variables. En utilisant le théorème ci-dessus, on voit facilement que la fonction possède deux points critiques (au moins car il peut en avoir en dehors de la figure) : un point critique selle situé à peu près en $(-0.3, 0.55)$, et un extremum (on ne peut pas savoir s'il s'agit d'un minimum ou d'un maximum avec seulement les lignes de niveaux) situé à peu près en $(1.2, -1.1)$.

