Fonctions de plusieurs variables

Cours de mathématiques
Deuxième année

Exo7

Sommaire

1 Fonctions de plusieurs variables **1**

1 Introduction 1
2 Graphe 5
3 Limites 14
4 Coordonnées polaires 20

2 Calcul différentiel **25**

1 Dérivées partielles 25
2 Différentielle 29
3 Fonctions de classe $\mathscr{C}^1$ 34

3 Matrice jacobienne **39**

1 Matrice jacobienne 39
2 Matrice jacobienne d'une composée 44

4 Gradient – Théorème des accroissements finis **49**

1 Gradient 49
2 Calcul d'incertitudes 58
3 Théorème des accroissements finis 61

5 Difféomorphismes **65**

1 Difféomorphismes 65
2 Théorème d'inversion locale 67
3 Théorème des fonctions implicites 70

6 Extremums **77**

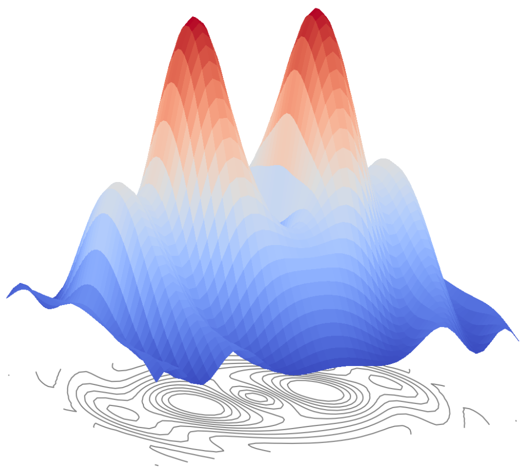
1 Motivation – Cas d'une variable 77
2 Dérivées partielles d'ordre 2 79
3 Formule de Taylor à l'ordre 2 82
4 Différentielle seconde 85
5 Minimum et maximum : cas de deux variables 87
6 Minimum et maximum : cas général 94
7 Extremums liés 96

Chapitre
1

Fonctions de plusieurs variables

1. Introduction

En première année, vous avez étudié les fonctions d'une variable : par exemple, si $t \mapsto f(t)$ représente l'évolution d'une population en fonction du temps, vous savez déterminer ses caractéristiques (croissance, maximum, limite...). Mais de nombreux phénomènes dépendent de plusieurs paramètres : par exemple, le volume d'un gaz dépend de la température et de la pression, ou bien l'altitude d'un point à la surface de la Terre dépend de la latitude et de la longitude. Le but de ce cours est de faire le même travail que pour les fonctions d'une variable : étudier la croissance, les maximums, les limites... Bien sûr, la situation est plus délicate, mais aussi plus intéressante, du fait qu'il y a plusieurs variables !



1.1. Que sont les fonctions de plusieurs variables ?

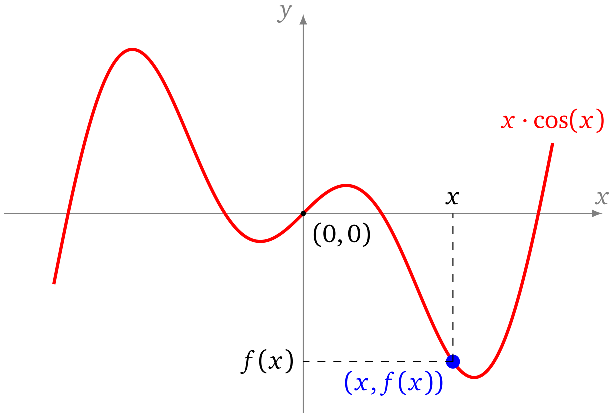
Dans ce chapitre, nous allons étudier les fonctions de plusieurs variables dans le cadre particulier de $\mathbb{R}^2$ ou $\mathbb{R}^3$, mais également dans le cadre général de $\mathbb{R}^n$. Ces fonctions seront donc de la forme

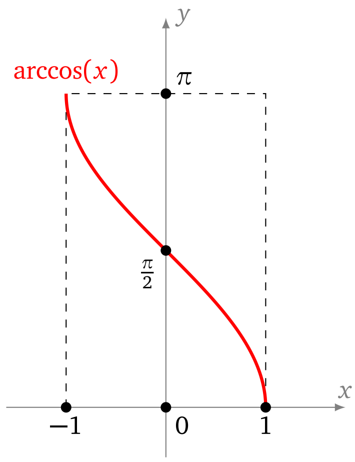
$$f : E \subset \mathbb{R}^n \to \mathbb{R},$$

où $n \geqslant 1$ est un entier naturel.
Autrement dit, les éléments de l'ensemble de départ E seront des n-uplets du type $(x_1, \ldots, x_n)$ que l'on peut considérer comme des vecteurs, et les éléments de l'ensemble d'arrivée seront des réels.

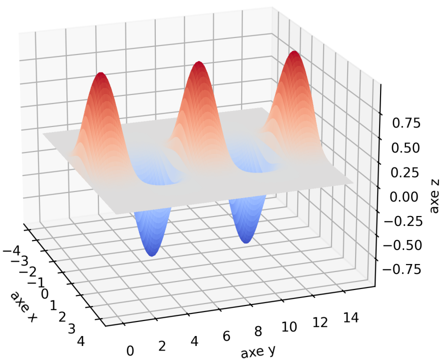
Exemple 1.

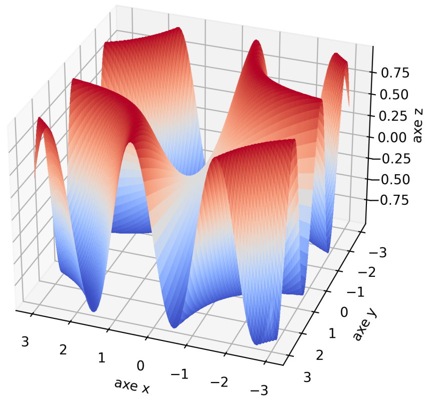
1. $n = 1$. $f : I \subset \mathbb{R} \to \mathbb{R}$: c'est le cas le plus simple, $x \mapsto f(x)$, celui qui est connu depuis le lycée. Voici les graphes des fonctions $x \mapsto x \cdot \cos(x)$ et $x \mapsto \arccos(x)$:





2. $n = 2$. $f : E \subset \mathbb{R}^2 \to \mathbb{R}$. On préfère noter les variables par (x, y) (au lieu de (x_1, x_2)). Ces fonctions $(x, y) \mapsto f(x, y)$ seront notre principal sujet d'étude et sont représentées par des surfaces. À gauche, le graphe de la fonction $(x, y) \mapsto \sin(y)e^{-x^2}$. À droite, le graphe de la fonction $(x, y) \mapsto \sin(xy)$.





Dès que $n > 2$, il est assez difficile d'avoir une vision graphique.

Nous allons aussi étudier des fonctions, dites fonctions vectorielles, dont l'ensemble d'arrivée n'est pas $\mathbb{R}$, mais $\mathbb{R}^p$, donc de la forme

$$f : E \subset \mathbb{R}^n \to \mathbb{R}^p,$$

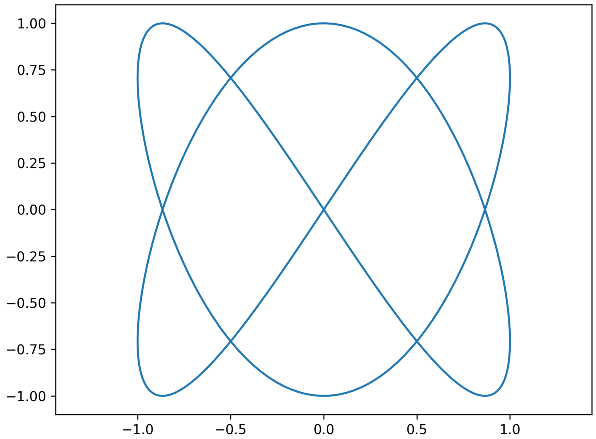
où $n \geqslant 1$ et $p \geqslant 1$ sont des entiers naturels.

Dans ce cas, si $x = (x_1, \ldots, x_n)$ est un vecteur de $\mathbb{R}^n$, alors $f(x)$ est un vecteur de $\mathbb{R}^p$, du type $f(x) = (f_1(x_1, \ldots, x_n), \ldots, f_p(x_1, \ldots, x_n))$. Attention, dans la suite, x désignera parfois le vecteur $x = (x_1, \ldots, x_n)$ et parfois x désignera un seul réel (comme par exemple pour une fonction de deux variables $f(x, y)$).

Exemple 2.

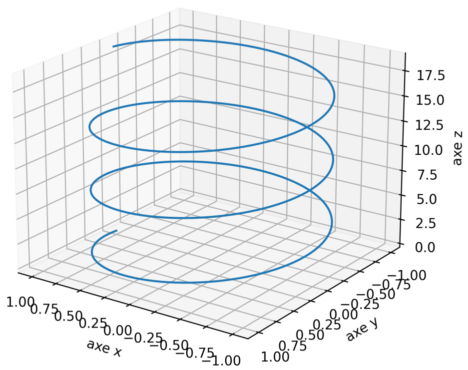
1. $n = 1$, $p = 2$. $f : I \subset \mathbb{R} \to \mathbb{R}^2$ est représentée par une courbe paramétrée du plan.

   Exemple : la fonction $f : \mathbb{R} \to \mathbb{R}^2$ définie par $t \mapsto (\sin(2t), \sin(3t))$. Cela correspond à la courbe paramétrée définie par $x(t) = \sin(2t)$ et $y(t) = \sin(3t)$.



On peut interpréter le dessin comme l'ensemble des positions $(x(t), y(t)) \in \mathbb{R}^2$ que prend une particule dans le plan en fonction du temps $t \in \mathbb{R}$.

2. $n = 1$, $p = 3$. $f : I \subset \mathbb{R} \to \mathbb{R}^3$ est représentée par une courbe paramétrée de l'espace. Exemple : la fonction $f : \mathbb{R} \to \mathbb{R}^3$ définie par $t \mapsto (\cos t, \sin t, t)$. Cela correspond au mouvement dans l'espace d'une particule $(x(t), y(t), z(t))$, qui ici parcourt une hélice.

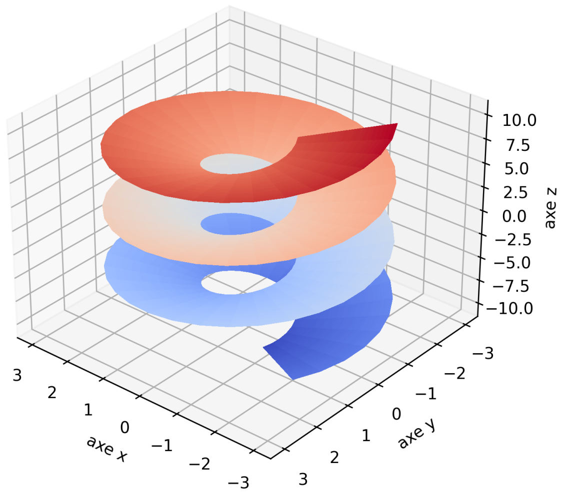


3. $n = 2$, $p = 3$. $f : E \subset \mathbb{R}^2 \to \mathbb{R}^3$: elles sont représentées par exemple par des surfaces paramétrées.

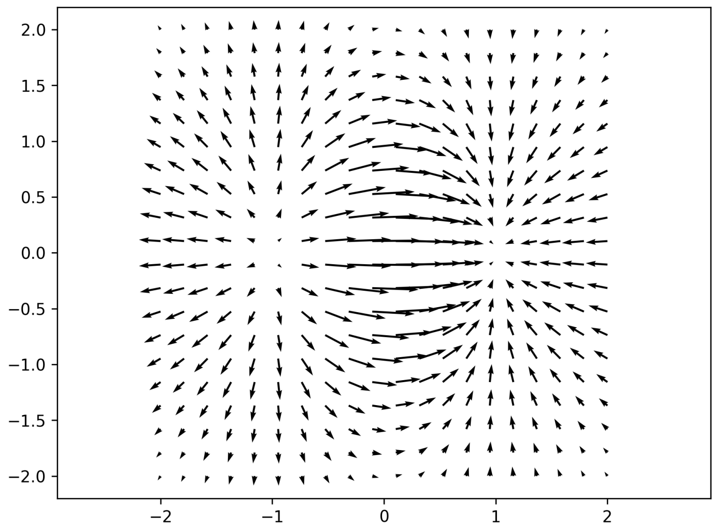
Voici l'exemple de la fonction

$$(u, v) \mapsto \big((2 + \sin v)\cos u, (2 + \sin v)\sin u, u + \cos v\big).$$

Chaque paramètre $(u, v) \in \mathbb{R}^2$ correspond à un point $(x(u, v), y(u, v), z(u, v)) \in \mathbb{R}^3$ de la surface hélicoïdale.



4. $n = 2$, $p = 2$. $f : E \subset \mathbb{R}^2 \to \mathbb{R}^2$: elles sont représentées par exemple par des champs de vecteurs. À chaque point du plan (x, y) on associe le vecteur $\vec{v} = f(x, y)$. (Sur la figure ci-dessous, seuls certains vecteurs sont représentés.)



Dans ce chapitre, nous étudions surtout les fonctions $f : \mathbb{R}^n \to \mathbb{R}$, et plus particulièrement les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$ ou $f : \mathbb{R}^3 \to \mathbb{R}$. Nous reviendrons plus tard sur les fonctions vectorielles $f : \mathbb{R}^n \to \mathbb{R}^p$.

1.2. Topologie de $\mathbb{R}^n$ (rappels/compléments)

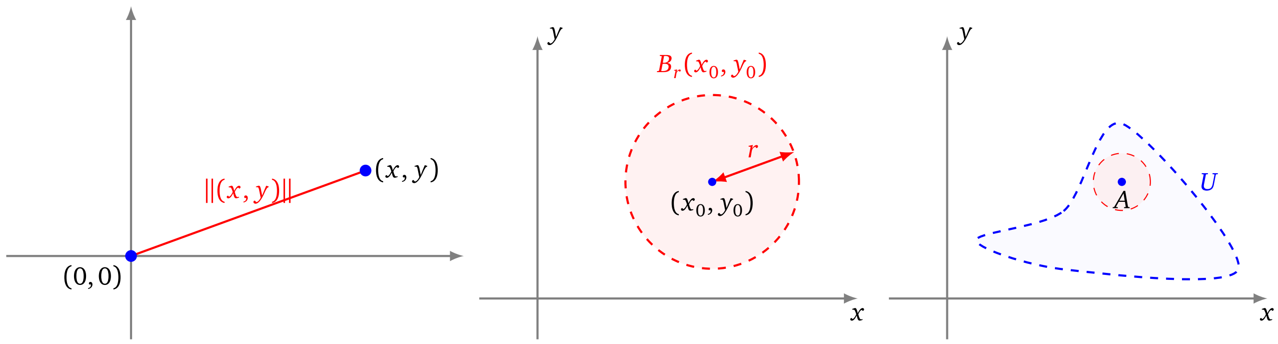
Voici quelques rappels de topologie dans l'espace vectoriel $\mathbb{R}^n$.

- Le ***produit scalaire*** usuel de $x = (x_1, \dots, x_n)$ et $y = (y_1, \dots, y_n)$, noté $\langle x \mid y \rangle$ (ou bien parfois $x \cdot y$), est défini par
$$\langle x \mid y \rangle = x_1 y_1 + \cdots + x_n y_n.$$
- La ***norme euclidienne*** sur $\mathbb{R}^n$ est la norme associée à ce produit scalaire. Pour $x \in \mathbb{R}^n$, la norme euclidienne de x, notée $\|x\|$, est définie par
$$\|x\| = \sqrt{\langle x \mid x \rangle} = \sqrt{x_1^2 + \cdots + x_n^2}.$$
- La ***distance*** entre le point $A = (a_1, \dots, a_n)$ et le point $M = (x_1, \dots, x_n)$ est
$$\|M - A\| = \sqrt{(x_1 - a_1)^2 + \cdots + (x_n - a_n)^2}.$$

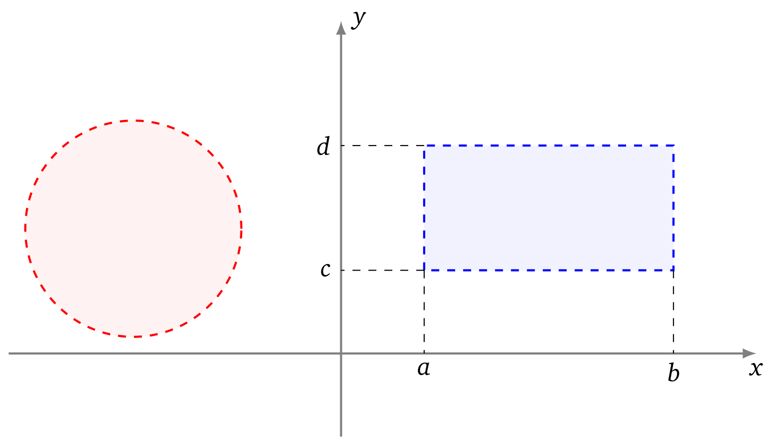
- La ***boule ouverte*** de centre $A = (a_1, \dots, a_n) \in \mathbb{R}^n$ et de rayon $r > 0$, notée $B_r(A)$, est l'ensemble suivant :
$$B_r(A) = \{M \in \mathbb{R}^n \mid \|M - A\| < r\}.$$
- Soient U une partie de $\mathbb{R}^n$ et $A \in U$. On dit que U est un ***voisinage*** de A si U contient une boule ouverte centrée en A.
- On dit que U est un ***ouvert*** de $\mathbb{R}^n$ si, pour tout point $A \in U$, U contient une boule ouverte centrée en A.

Dans le cas de $\mathbb{R}^2$, on note plutôt les coordonnées d'un point par (x, y). Alors :
- $\langle (x,y) \mid (x',y') \rangle = xx' + yy'$
- $\|(x,y)\| = \sqrt{x^2 + y^2}$
- $B_r(x_0, y_0) = \left\{(x,y) \in \mathbb{R}^2 \mid (x - x_0)^2 + (y - y_0)^2 < r^2\right\}$ (on parle de ***disque*** plutôt que de boule).

De gauche à droite : la norme euclidienne, un disque ouvert, un ouvert U.



Exemples :
- tout rectangle ouvert $]a, b[\times]c, d[$ est un ouvert de $\mathbb{R}^2$ (à droite sur la figure),
- tout disque ouvert de $\mathbb{R}^2$ est un ouvert de $\mathbb{R}^2$ (à gauche sur la figure).



2. Graphe

2.1. Fonctions

Définition 1.
Soit E une partie de $\mathbb{R}^n$. Une ***fonction*** $f : E \to \mathbb{R}$ associe à tout $(x_1, \dots, x_n)$ de E un seul nombre réel $f(x_1, \dots, x_n)$.

Exemple 3.

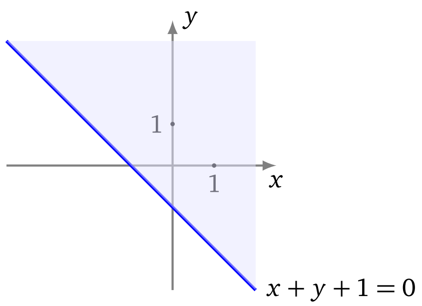
1. Distance d'un point à l'origine en fonction de ses coordonnées :
$$f \;:\; \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \sqrt{x^2+y^2}. \end{array}$$

2. Aire d'un rectangle en fonction de sa longueur et sa largeur :
$$\begin{array}{rccl} f & : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & & (x,y) & \longmapsto & xy. \end{array}$$
3. Aire d'un parallélépipède en fonction de ses trois dimensions :
$$\begin{array}{rccl} f & : & \mathbb{R}^3 & \longrightarrow & \mathbb{R} \\ & & (x,y,z) & \longmapsto & 2(xy+yz+xz). \end{array}$$

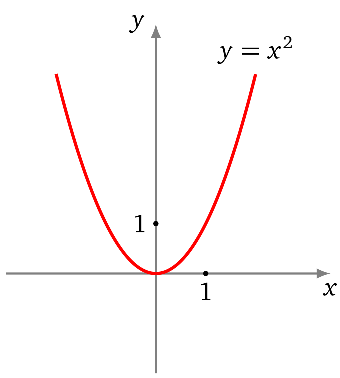
Définition 2.
Si on nous donne d'abord une expression pour $f(x_1,\dots,x_n)$, alors le ***domaine de définition*** de f est le plus grand sous-ensemble $D_f \subset \mathbb{R}^n$ tel que, pour chaque $(x_1,\dots,x_n)$ de D_f, $f(x_1,\dots,x_n)$ soit bien définie. La fonction est alors $f : D_f \to \mathbb{R}$.

Exemple 4.

1. $f(x,y) = \ln(1+x+y)$
   Il faut que $1+x+y$ soit strictement positif, afin de pouvoir calculer son logarithme. Donc :
   $$D_f = \left\{(x,y) \in \mathbb{R}^2 \mid 1+x+y > 0\right\}$$
   Pour tracer cet ensemble, on trace d'abord la droite d'équation $1+x+y=0$. On détermine ensuite de quel côté de la droite est l'ensemble $1+x+y>0$. Ici, c'est au-dessus de la droite.

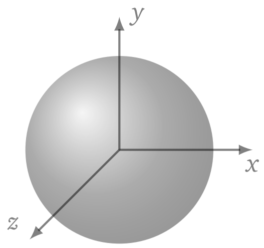


2. $f(x,y) = \exp\left(\frac{x+y}{x^2-y}\right)$
   Le dénominateur ne doit pas s'annuler :
   $$D_f = \left\{(x,y) \in \mathbb{R}^2 \mid x^2-y \neq 0\right\}$$
   Les points de l'ensemble de définition sont tous les points du plan qui ne sont pas sur la parabole d'équation $(y=x^2)$.



3. $f(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2-2}}$
   L'expression sous la racine doit être positive (pour pouvoir prendre la racine carrée) et ne doit pas s'annuler (pour pouvoir prendre l'inverse). Donc :
   $$D_f = \left\{(x,y,z) \in \mathbb{R}^3 \mid x^2+y^2+z^2 > 2\right\}$$

Autrement dit, ce sont tous les points en dehors de la boule fermée centrée en $(0,0,0)$ et de rayon $\sqrt{2}$.



Définition 3.
L'*image* d'une fonction $f : E \to \mathbb{R}$ est l'ensemble des valeurs $f(x_1,\ldots,x_n)$ prises par f lorsque $(x_1,\ldots,x_n)$ parcourt E :
$$\operatorname{Im} f = \big\{ f(x_1,\ldots,x_n) \mid (x_1,\ldots,x_n) \in E \big\} \subset \mathbb{R}$$

Exemple 5.

1. $f(x,y) = \ln(1+x+y)$
   L'image de f est $\mathbb{R}$ tout entier : $\operatorname{Im} f = \mathbb{R}$.
   Preuve : soit $z \in \mathbb{R}$. Alors, pour $(x,y) = (e^z,-1) \in D_f$, on a
   $$f(x,y) = f(e^z,-1) = \ln(e^z) = z.$$
   Donc tout $z \in \mathbb{R}$ est dans l'image de f.
2. $f(x,y) = \exp\left(\frac{x+y}{x^2-y}\right)$
   $\operatorname{Im} f =]0,+\infty[$.
   Preuve : On a bien sûr $f(x,y) > 0$ pour tout $(x,y) \in D_f$. Réciproquement, soit $z \in]0,+\infty[$. Si $z \neq 1$ alors, pour $(x,y) = (\frac{1}{\ln z},0) \in D_f$, on a
   $$f(x,y) = f(\tfrac{1}{\ln z},0) = \exp\left(\frac{\frac{1}{\ln z}}{(\frac{1}{\ln z})^2}\right) = \exp(\ln z) = z.$$
   Si $z = 1$ alors, pour $(x,y) = (1,-1) \in D_f$, on a $f(x,y) = f(1,-1) = \exp(0) = 1$.
3. Pour $f(x,y,z) = \frac{1}{\sqrt{x^2+y^2+z^2-2}}$, on a $\operatorname{Im} f =]0,+\infty[$. À vous de faire la preuve !

Définition 4.
Soit E une partie de $\mathbb{R}^n$. Une *fonction vectorielle* $f : E \to \mathbb{R}^p$ associe à tout $(x_1,\ldots,x_n)$ de E un p-uplet de nombres réels. On la note
$$\begin{array}{rccl} f : & \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ & x & \longmapsto & \big(f_1(x), f_2(x), \ldots, f_p(x)\big). \end{array}$$

Nous nous limiterons souvent aux dimensions inférieures ou égales à 3 pour n et p, la généralisation aux dimensions supérieures ne posant pas de problème particulier, sauf pour faire des dessins. Voici quelques exemples simples.

Exemple 6.

1. Aire et volume d'un parallélépipède rectangle en fonction de ses trois dimensions :
$$\begin{array}{rccl} f : & \mathbb{R}^3 & \longrightarrow & \mathbb{R}^2 \\ & (x,y,z) & \longmapsto & \big(2(xy+yz+xz), xyz\big). \end{array}$$
2. Coordonnées polaires d'un point du plan :
$$\begin{array}{rccl} f : & \mathbb{R}_+ \times [0,2\pi[& \longrightarrow & \mathbb{R}^2 \\ & (r,\theta) & \longmapsto & (r\cos\theta, r\sin\theta). \end{array}$$

2.2. Graphe et lignes de niveau

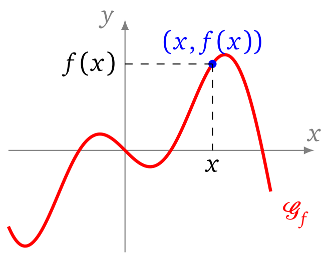
Définition 5.
Soit $f : D_f \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de 2 variables. Le ***graphe*** $\mathscr{G}_f$ de f est le sous-ensemble de $\mathbb{R}^3$ formé des points de coordonnées $(x, y, f(x, y))$ avec (x, y) dans l'ensemble de définition. Le graphe est donc :

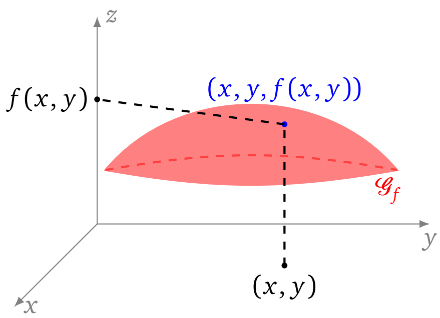
$$\mathscr{G}_f = \left\{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in D_f \text{ et } z = f(x, y)\right\}.$$

Représenter graphiquement le graphe n'est possible que pour les fonctions d'une seule variable ou de deux variables. Pour les fonctions d'une variable $f : D_f \subset \mathbb{R} \to \mathbb{R}$, on rappelle que le graphe est

$$\mathscr{G}_f = \left\{(x, y) \in \mathbb{R}^2 \mid x \in D_f \text{ et } y = f(x)\right\}.$$

Dans le cas d'une variable (à gauche), le graphe est une courbe ; dans le cas de deux variables qui nous intéresse ici, c'est une surface.



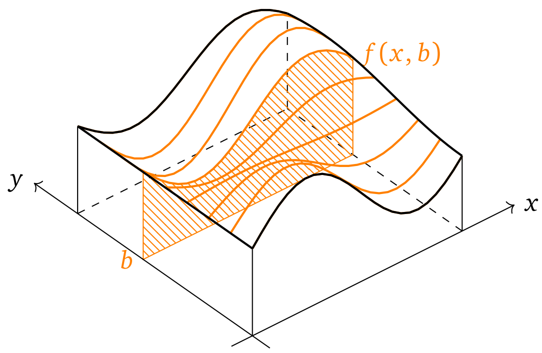


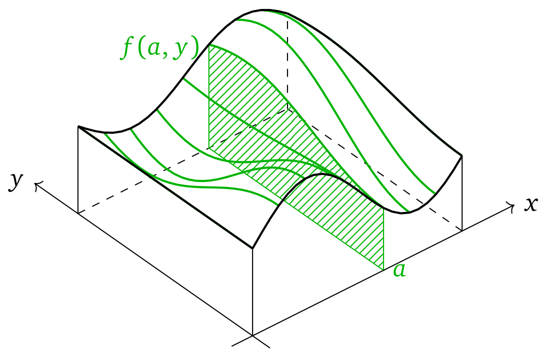
Afin de tracer le graphe d'une fonction de deux variables, on découpe la surface en morceaux.

Tranches.
Une première façon de faire : tracer, pour quelques valeurs de a et b, les graphes des fonctions partielles

$$f_1 : x \mapsto f(x, b) \qquad \text{et} \qquad f_2 : y \mapsto f(a, y).$$

La première représente l'intersection du graphe $\mathscr{G}_f$ avec le plan $(y = b)$ (en orange) et la seconde représente l'intersection du graphe avec le plan $(x = a)$ (en vert).





Lignes de niveau.
On va aussi s'intéresser à d'autres courbes tracées sur la surface : les courbes de niveau.

Définition 6.
Soit $f : D_f \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de deux variables.

- La ***ligne de niveau*** $z = c \in \mathbb{R}$ est

$$L_c = \left\{(x, y) \in D_f \mid f(x, y) = c\right\}.$$

- La *courbe de niveau* $z = c$ est la trace de $\mathscr{G}_f$ dans le plan $(z = c)$:
$$\mathscr{G}_f \cap (z = c) = \left\{(x, y, c) \in \mathbb{R}^3 \mid f(x, y) = c\right\}.$$

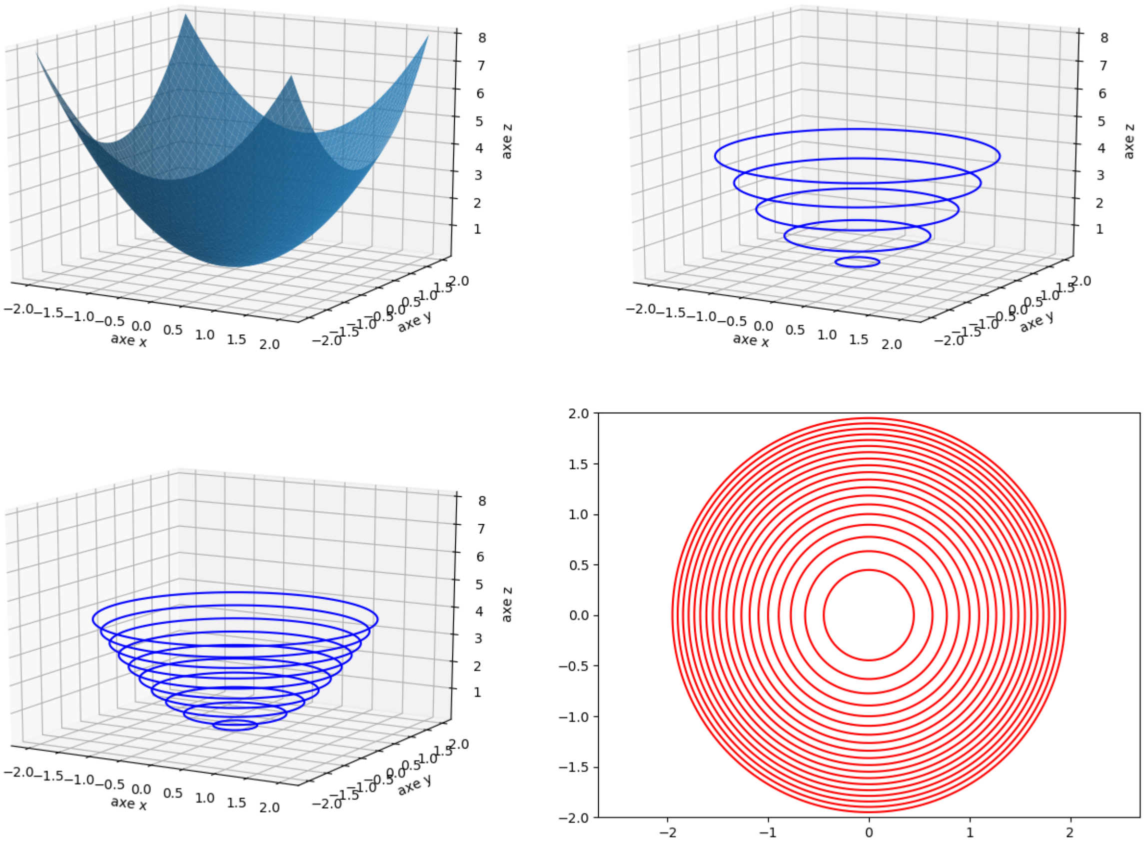
La ligne de niveau c est une courbe du plan $\mathbb{R}^2$, la courbe de niveau c est une courbe de l'espace $\mathbb{R}^3$. On obtient la courbe de niveau c en partant de la ligne de niveau c et en remontant à l'altitude c.

Exemple 7.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = x^2 + y^2$.

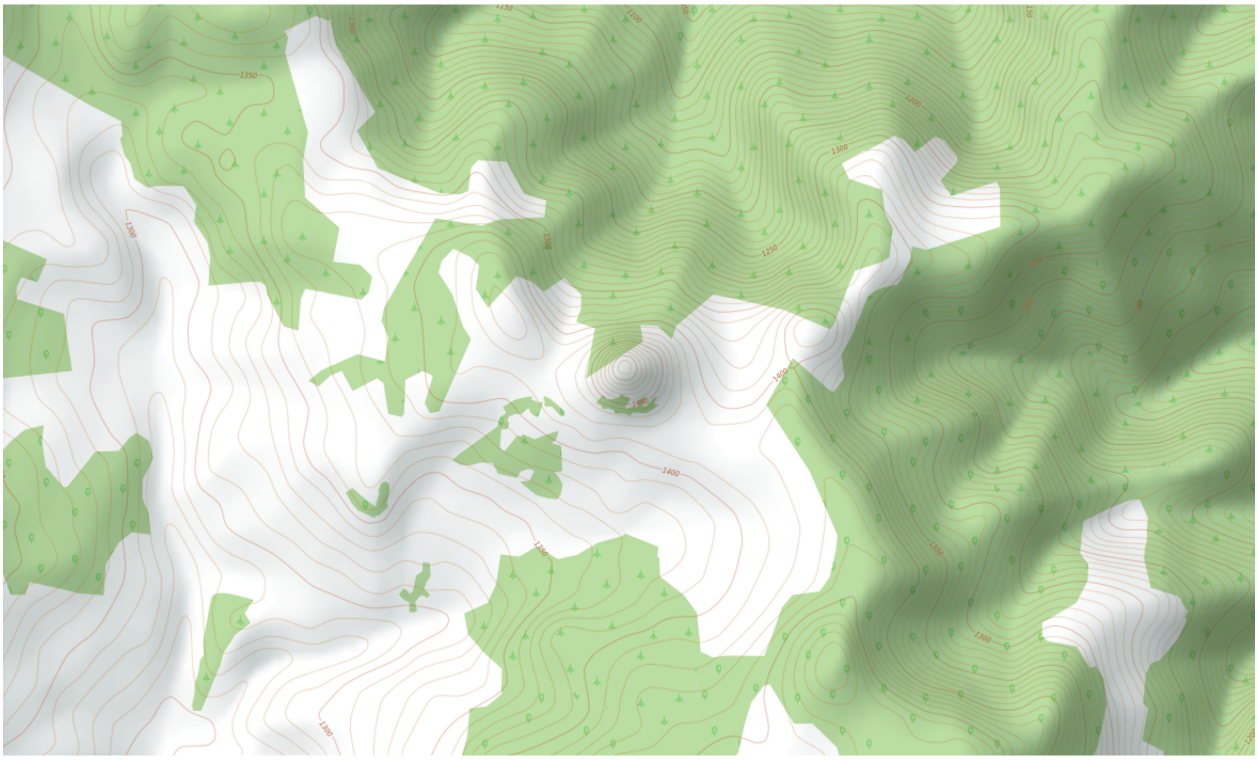
- Si $c < 0$, la ligne de niveau L_c est vide (aucun point n'a d'altitude négative).
- Si $c = 0$, la ligne de niveau L_0 se réduit à $\{(0, 0)\}$.
- Si $c > 0$, la ligne de niveau L_c est le cercle du plan de centre $(0, 0)$ et de rayon $\sqrt{c}$. On remonte L_c à l'altitude $z = c$: la courbe de niveau est alors le cercle horizontal de l'espace de centre $(0, 0, c)$ et de rayon $\sqrt{c}$.

Le graphe est alors une superposition de cercles horizontaux de l'espace de centre $(0, 0, c)$ et de rayon $\sqrt{c}$ avec $c \geqslant 0$.

Ci-dessous : (a) le graphe, appelé paraboloïde de révolution, (b) 5 courbes de niveau, (c) 10 courbes de niveau, (d) des lignes de niveau dans le plan.



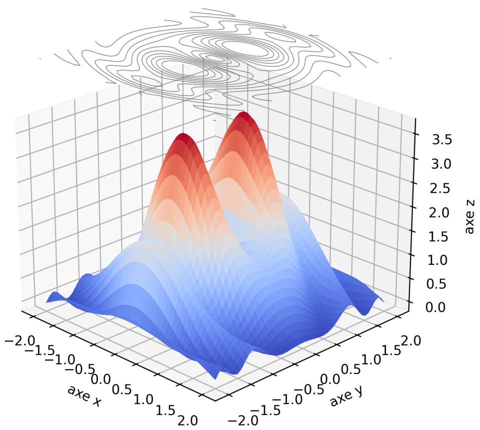
Exemple 8.
Sur une carte topographique, les lignes de niveau représentent les courbes ayant la même altitude.

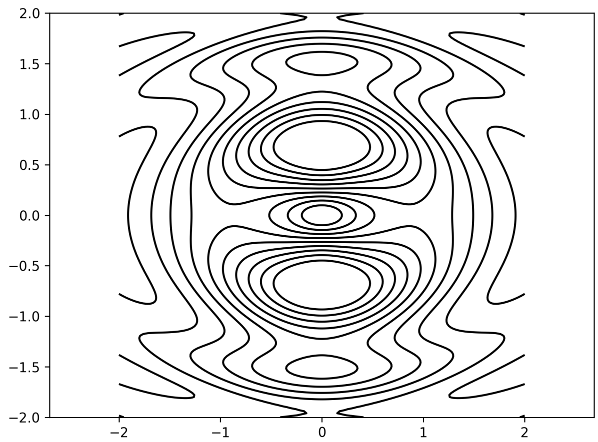


- Ici, une carte *Open Street Map* avec, au centre, le mont Gerbier-de-Jonc (source de la Loire, 1551 m).
- Les lignes de niveau correspondent à des altitudes par cran de 10 m (par exemple, pour $c = 1400$, $c = 1410$, $c = 1420\ldots$).
- Lorsque les lignes de niveau sont très espacées, le terrain est plutôt plat ; lorsque les lignes sont rapprochées, le terrain est pentu.
- Par définition, si on se promène en suivant une ligne de niveau, on reste toujours à la même altitude !

Exemple 9.
L'image qui a servi lors de l'introduction est le graphe et les lignes de niveau de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$(x, y) \mapsto z = \frac{\sin\left(x^2 + 3y^2\right)}{\frac{1}{10} + r^2} + \left(x^2 + 5y^2\right) \cdot \frac{\exp\left(1 - r^2\right)}{2} \qquad \text{avec} \qquad r = \sqrt{x^2 + y^2}.$$





Surfaces de niveau.
Pour les fonctions de 3 variables, le graphe étant dans $\mathbb{R}^4$, on ne peut le dessiner. La notion analogue à la ligne de niveau est celle de ***surface de niveau***, donnée par l'équation $f(x, y, z) = c$.

Exemple 10.
$f(x,y,z) = x^2 + y^2 + z^2$. Les surfaces de niveau sont données par l'équation $x^2 + y^2 + z^2 = c$. Pour $c \geqslant 0$, ces surfaces sont des sphères, centrées à l'origine et de rayon $\sqrt{c}$. Voici ces surfaces pour $c = 1, 3, 5$. Elles ont été découpées pour laisser entrevoir les surfaces des différents niveaux.

2.3. Exemples de surfaces quadratiques

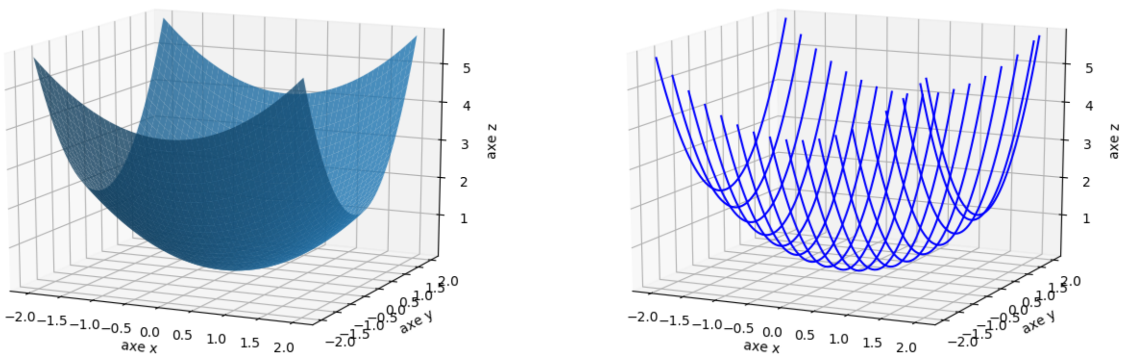
Ce sont des exemples à connaître, car ils seront fondamentaux pour la suite du cours.

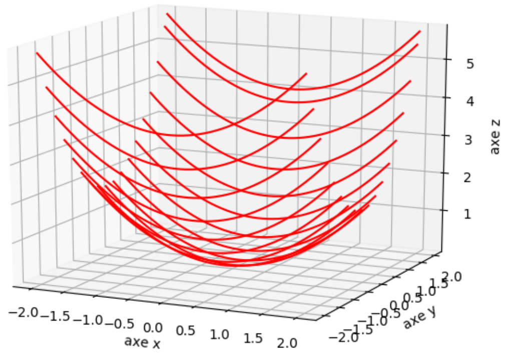
Exemple 11.

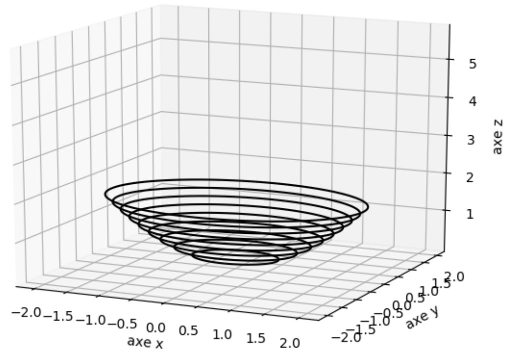
$$f(x,y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1$$

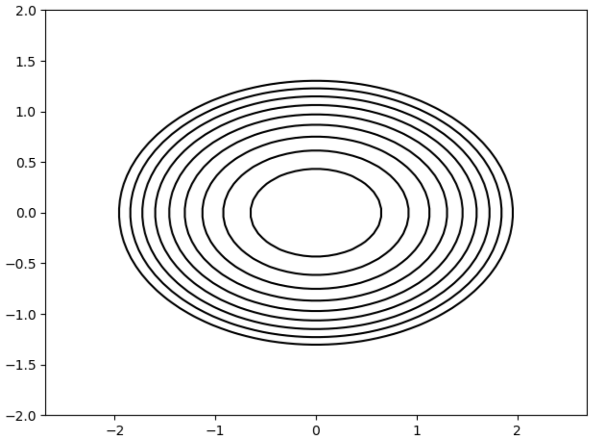
- Les tranches sont des paraboles.
- Les lignes de niveau sont des ellipses.
- Le graphe est donc un ***paraboloïde elliptique***.

Ci-dessous : (a) la surface, (b) des tranches avec x constant, (c) des tranches avec y constant, (d) des courbes de niveau, (e) des lignes de niveau dans le plan.







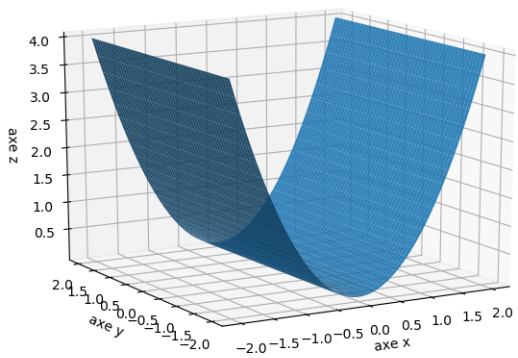


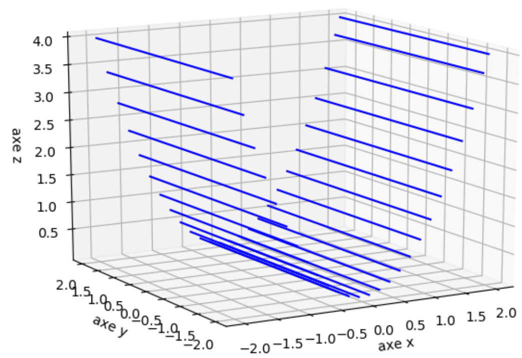
Exemple 12.

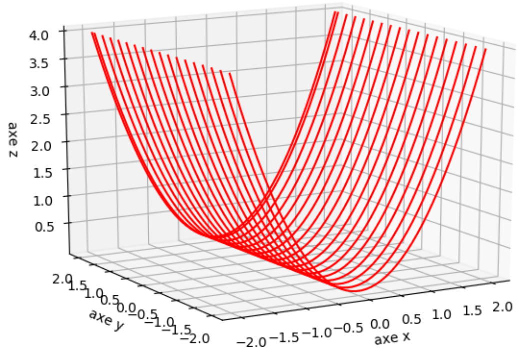
$$f(x, y) = x^2$$

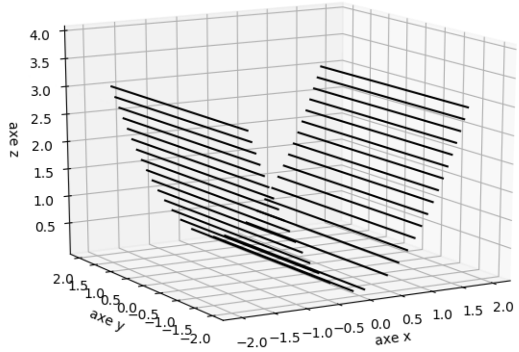
- Les tranches (à y constant) sont des paraboles.
- Les lignes de niveau sont des paires de droites.
- Le graphe est donc un ***cylindre parabolique***.

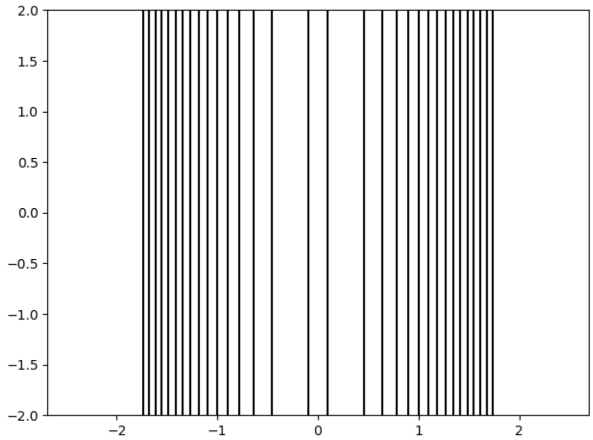
Ci-dessous : (a) la surface, (b) des tranches avec x constant, (c) des tranches avec y constant, (d) des courbes de niveau, (e) des lignes de niveau dans le plan.









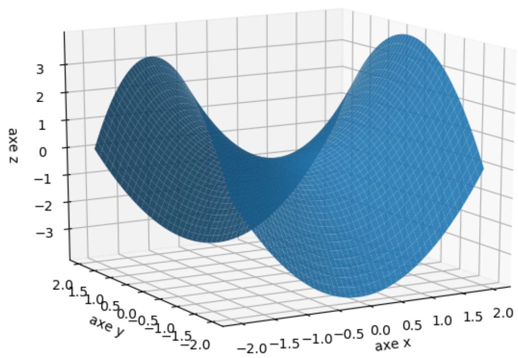


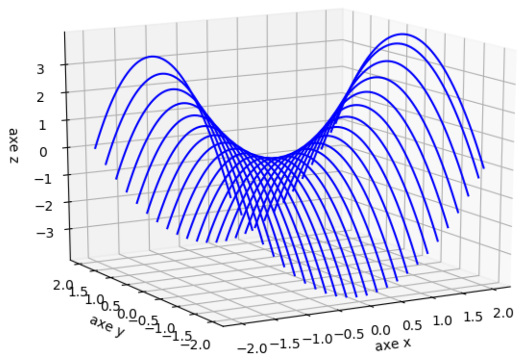
Exemple 13.

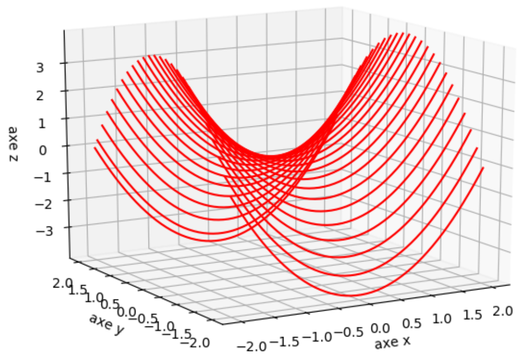
$$f(x,y) = x^2 - y^2$$

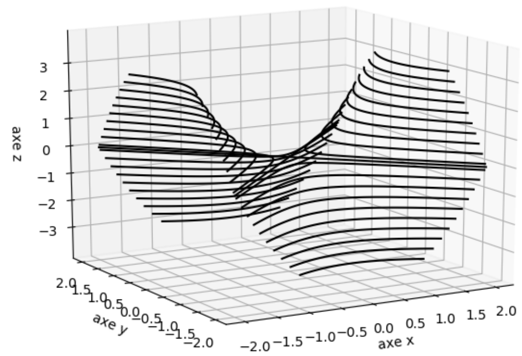
- Les tranches sont des paraboles.
- Les lignes de niveau sont des hyperboles.
- Le graphe est donc un ***paraboloïde hyperbolique***, que l'on appelle aussi la ***selle de cheval***.
- Un autre nom pour cette surface est un ***col*** (du nom d'un col de montagne). En effet, le point $(0,0,0)$ est le point de passage le moins haut pour passer d'un versant à l'autre de la montagne.

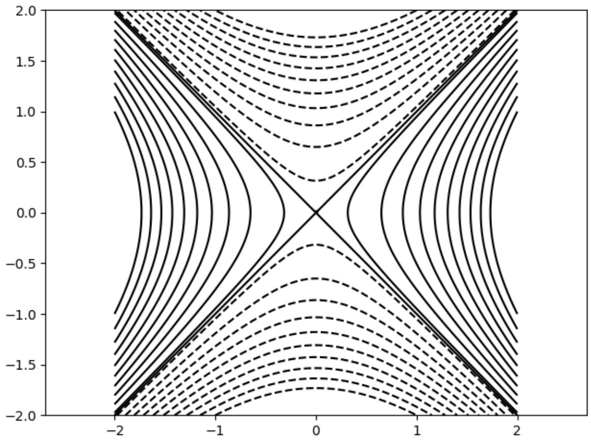
Ci-dessous : (a) la surface, (b) des tranches avec x constant, (c) des tranches avec y constant, (d) des courbes de niveau, (e) des lignes de niveau dans le plan (en pointillé les lignes de niveau négatif).











Mini-exercices.

1. Déterminer et dessiner le domaine de définition de la fonction définie par $f(x,y) = \ln(xy)$. Même question avec $g(x,y) = \sqrt{2x - y^2 + 1}$ et $h(x,y,z) = \frac{1}{x} + \frac{1}{y} + \frac{1}{z}$.
2. Déterminer l'image des fonctions de la question précédente.
3. Soit $f(x,y) = xy$. Dessiner le graphe de f, les tranches et les lignes de niveau. Quelle surface reconnaissez-vous ? Vous pouvez vous aider d'un ordinateur. Mêmes questions avec $g(x,y) = -x^2 - y^2$.

3. Limites

Les notions de limite et de continuité des fonctions d'une seule variable se généralisent en plusieurs variables sans complexité supplémentaire : il suffit de remplacer la valeur absolue par la norme euclidienne.

3.1. Définition

Soit f une fonction $f : E \subset \mathbb{R}^n \to \mathbb{R}$ définie au voisinage de $x_0 \in \mathbb{R}^n$, sauf peut-être en x_0.

Définition 7.
La fonction f admet pour ***limite*** le nombre réel ℓ lorsque x tend vers x_0 si :

$$\forall \epsilon > 0 \quad \exists \delta > 0 \qquad \forall x \in E \qquad 0 < \|x - x_0\| < \delta \implies |f(x) - \ell| < \epsilon$$

On écrit alors

$$\lim_{x_0} f = \ell \qquad \text{ou} \qquad \lim_{x \to x_0} f(x) = \ell \qquad \text{ou} \qquad f(x) \xrightarrow[x \to x_0]{} \ell.$$

On définirait de même $\lim_{x \to x_0} f(x) = +\infty$ par :

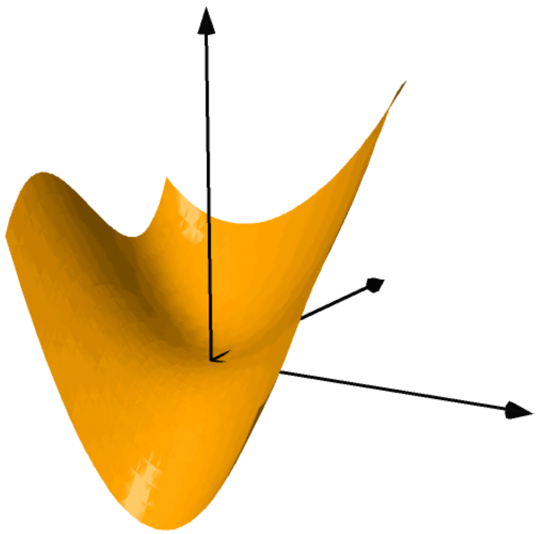
$$\forall A > 0 \quad \exists \delta > 0 \qquad \forall x \in E \qquad 0 < \|x - x_0\| < \delta \implies |f(x)| > A$$

Remarque.

- La notion de limite ne dépend pas ici des normes utilisées.
- Si elle existe, la limite est unique.

Exemple 14.
Soit f la fonction définie par $f(x, y) = x^2 + y\sin(x + y^2)$.

1. Montrer que $f(x, y)$ tend vers 0 lorsque $(x, y) \to (0, 0)$.
2. Trouver un ouvert U contenant l'origine tel que, pour tout $(x, y) \in U$, on ait $|f(x, y)| < \frac{1}{100}$.



Solution.

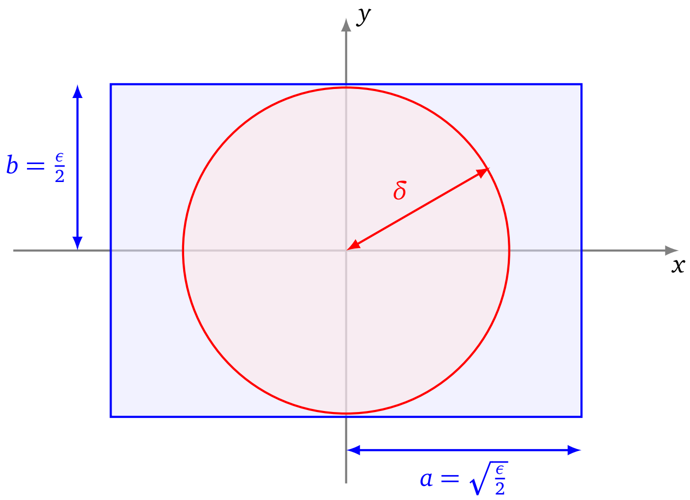
1. On majore $|f(x, y)|$ en utilisant l'inégalité triangulaire et $|\sin(t)| \leqslant 1$:

$$\left|f(x, y)\right| = \left|x^2 + y\sin(x + y^2)\right| \leqslant x^2 + |y|\left|\sin(x + y^2)\right| \leqslant x^2 + |y|$$

Fixons $0 < \epsilon < 1$. Fixons $a = \sqrt{\frac{\epsilon}{2}}$ et $b = \frac{\epsilon}{2}$. Alors, pour $x \in]-a, a[$, on a $x^2 < \frac{\epsilon}{2}$ et, pour $y \in]-b, b[$, on a $|y| < \frac{\epsilon}{2}$. Pour $(x, y) \in]-a, a[\times]-b, b[$, on a donc

$$\left|f(x, y)\right| \leqslant x^2 + |y| < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon.$$

Une valeur δ qui convient est donc $\delta = \frac{\epsilon}{2}$. En effet, si $\|(x, y)\| < \delta$ alors $|x| < \delta = \frac{\epsilon}{2} \leqslant \sqrt{\frac{\epsilon}{2}}$ et $|y| < \delta = \frac{\epsilon}{2}$ donc $|f(x, y)| < \epsilon$. Conclusion : f admet pour limite 0 lorsque (x, y) tend vers $(0, 0)$.



2. Pour $\epsilon = \frac{1}{100}$, on a $a = \frac{1}{\sqrt{200}}$ et $b = \frac{1}{200}$. Pour chaque (x, y) de l'ouvert $]-a, a[\times]-b, b[$, on a $|f(x, y)| < \frac{1}{100}$.

3.2. Opérations sur les limites

Pour calculer les limites, on ne recourt que rarement à cette définition. On utilise plutôt les théorèmes généraux : opérations sur les limites et encadrement. Ce sont les mêmes énoncés que pour les fonctions d'une variable : il n'y a aucune difficulté ni nouveauté.

Proposition 1 (Opérations sur les limites).
Soient $f, g : \mathbb{R}^n \to \mathbb{R}$ définies au voisinage de $x_0 \in \mathbb{R}^n$ et telles que f et g admettent des limites en x_0. Alors :

$$\lim_{x_0}(f+g) = \lim_{x_0} f + \lim_{x_0} g \qquad \lim_{x_0}(f \cdot g) = \lim_{x_0} f \times \lim_{x_0} g$$

Et si g ne s'annule pas dans un voisinage de x_0 :

$$\lim_{x_0} \frac{1}{g} = \frac{1}{\lim_{x_0} g} \qquad \lim_{x_0} \frac{f}{g} = \frac{\lim_{x_0} f}{\lim_{x_0} g}$$

Remarque.

- Les résultats ci-dessus sont aussi valables pour des limites infinies avec les conventions usuelles :

$$\ell + \infty = +\infty, \quad \ell - \infty = -\infty, \quad \frac{1}{0^+} = +\infty, \quad \frac{1}{0^-} = -\infty, \quad \frac{1}{\pm\infty} = 0,$$

$$\ell \times \infty = \infty \ (\ell \neq 0), \ \infty \times \infty = \infty \text{ (avec règle de multiplication des signes).}$$

- Les formes indéterminées sont : $+\infty - \infty$, $\frac{0}{0}$, $\frac{\infty}{\infty}$, $0 \times \infty$, , ∞^0, 1^∞ et 0^0.

La composition est aussi souvent utile :

- soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de plusieurs variables, telle que $\lim_{x \to x_0} f(x) = \ell$,
- soit $g : \mathbb{R} \to \mathbb{R}$ une fonction d'une seule variable, telle que $\lim_{t \to \ell} g(t) = \ell'$,
- alors la fonction de plusieurs variables $g \circ f : \mathbb{R}^n \to \mathbb{R}$ définie par $(g \circ f)(x) = g\big(f(x)\big)$ vérifie $\lim_{x \to x_0} (g \circ f)(x) = \ell'$.

Application : grâce à l'exemple 14, et comme $e^t \to 1$ lorsque $t \to 0$, on en déduit :

$$e^{x^2 + y\sin(x+y^2)} \xrightarrow[(x,y)\to(0,0)]{} 1$$

Il existe aussi un théorème « des gendarmes ».

Théorème 1 (Théorème d'encadrement).
Soient $f, g, h : \mathbb{R}^n \to \mathbb{R}$ trois fonctions définies dans un voisinage U de $x_0 \in \mathbb{R}^n$.

- *Si, pour tout $x \in U$, on a $f(x) \leqslant h(x) \leqslant g(x)$,*

- *et si* $\lim_{x_0} f = \lim_{x_0} g = \ell$,

alors h admet une limite au point x_0 *et* $\lim_{x_0} h = \ell$.

Exemple 15.
Soit h définie par $h(x,y) = \cos(x+y^2)\big(x^2 + y\sin(x+y^2)\big)$. On majore la valeur absolue du cosinus par 1 :

$$\big|h(x,y)\big| \leqslant \big|x^2 + y\sin(x+y^2)\big|.$$

On a vu lors de l'exemple 14 que la fonction définie par $f(x,y) = x^2 + y\sin(x+y^2)$ tend vers 0 en $(0,0)$. Donc, par le théorème des gendarmes, $h(x,y)$ tend aussi vers 0 lorsque (x,y) tend vers $(0,0)$.

3.3. Limite le long d'un chemin

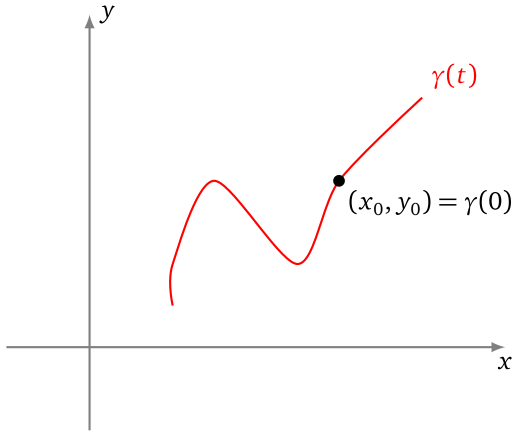
L'unicité de la limite implique que, quelle que soit la façon dont on arrive au point x_0, la valeur limite est toujours la même.

Proposition 2.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ *une fonction définie au voisinage de* $x_0 \in \mathbb{R}^n$, *sauf peut-être en* x_0.

1. *Si* f *admet une limite* ℓ *au point* x_0, *alors la restriction de* f *à toute courbe passant par* x_0 *admet une limite en* x_0 *et cette limite est* ℓ.
2. *Par contraposée, si les restrictions de* f *à deux courbes passant par* x_0 *ont des limites différentes au point* x_0, *alors* f *n'admet pas de limite au point* x_0.

Détaillons dans les cas des fonctions de deux variables :

- Une courbe passant par le point $(x_0, y_0) \in \mathbb{R}^2$ est une fonction continue $\gamma : \mathbb{R} \to \mathbb{R}^2$, $t \mapsto (x(t), y(t))$, telle que $\gamma(0) = (x_0, y_0)$.
- La restriction de f le long de γ est la fonction d'une variable $f \circ \gamma : t \mapsto f\big(x(t), y(t)\big)$.
- Si f a pour limite ℓ en (x_0, y_0) alors la première partie de la proposition affirme que $f\big(x(t), y(t)\big) \xrightarrow[t \to 0]{} \ell$.



Exemple 16.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

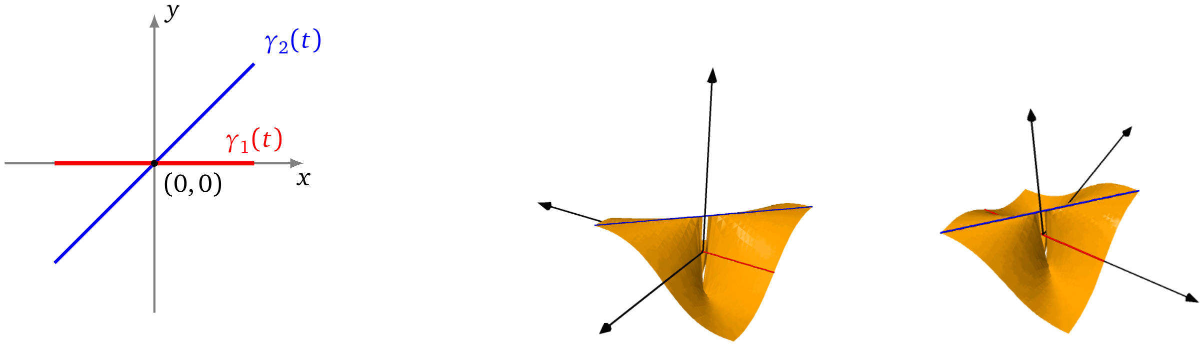
$$f(x,y) = \frac{xy}{x^2+y^2} \text{ si } (x,y) \neq (0,0) \quad \text{et} \quad f(0,0) = 0.$$

La fonction f admet-elle une limite en $(0,0)$?

Solution.

- Si on prend le chemin $\gamma_1(t) = (t, 0)$, alors $(f \circ \gamma_1)(t) = f(t,0) = 0$. Donc, lorsque $t \to 0$, $(f \circ \gamma_1)(t) \to 0$.

- Si on prend le chemin $\gamma_2(t) = (t,t)$, alors $(f \circ \gamma_2)(t) = f(t,t) = \frac{t^2}{2t^2} = \frac{1}{2}$. Donc, lorsque $t \to 0$, $(f \circ \gamma_2)(t) \to \frac{1}{2}$.
  Ci-dessous, sur la figure de gauche, les deux chemins du plan ; sur les deux figures de droite, deux vues différentes des valeurs prises par f le long de ces chemins.



- Si f admettait une limite ℓ alors, quel que soit le chemin $\gamma(t)$ tel que $\gamma(t) \to (0,0)$ lorsque $t \to 0$, on aurait $(f \circ \gamma)(t) \to \ell$. On aurait donc $\ell = 0$ et $\ell = \frac{1}{2}$, ce qui contredirait l'unicité de la limite. Ainsi, f n'a pas de limite en $(0,0)$.

Une autre formulation possible :
Si $f : \mathbb{R}^n \to \mathbb{R}$ a pour limite ℓ en $x_0 \in \mathbb{R}^n$ alors, pour toute suite (u_n) d'éléments de $\mathbb{R}^n$ telle que $u_n \to x_0$, on a $f(u_n) \to \ell$. Pour les fonctions de deux variables, cela s'écrit ainsi : si f a pour limite ℓ en (a,b) alors, pour toute suite $(a_n, b_n) \to (a,b)$, on a $f(a_n, b_n) \to \ell$.

3.4. Fonctions continues

Définition 8.
1. $f : E \subset \mathbb{R}^n \to \mathbb{R}$ est ***continue en*** $x_0 \in E$ si $\lim_{x \to x_0} f(x) = f(x_0)$.
2. f est ***continue sur*** E si elle est continue en tout point de E.

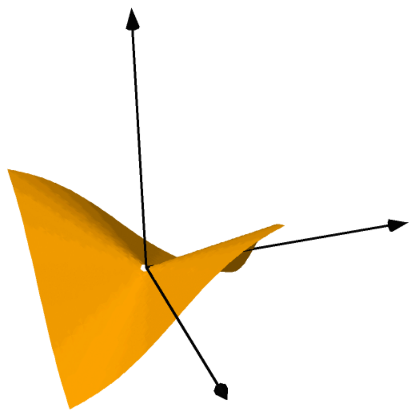
Par les propriétés des limites, si f et g sont deux fonctions continues en x_0, alors :
- la fonction $f + g$ est continue en x_0,
- de même fg et f/g (avec $g(x) \neq 0$ sur un voisinage de x_0) sont continues en x_0,
- si $h : \mathbb{R} \to \mathbb{R}$ est continue, alors $h \circ f$ est continue en x_0.

Exemple 17.
- Les applications définies par $(x,y) \mapsto x + y$, $(x,y) \mapsto xy$, puis toutes les fonctions polynômes en deux variables x et y sont continues sur $\mathbb{R}^2$ (par exemple $(x,y) \mapsto x^2 + 3xy$). De la même façon, toutes les fractions rationnelles en deux variables sont continues là où elles sont définies.
- Comme l'exponentielle est une fonction continue, alors $(x,y) \mapsto e^{xy}$ est continue sur $\mathbb{R}^2$.
- La fonction définie par $f(x,y) = \frac{1}{\sqrt{x^2+y^2}}$ est continue sur $\mathbb{R}^2 \setminus \{(0,0)\}$.

Définition 9 (Prolongement par continuité).
Soit $f : E \subset \mathbb{R}^n \to \mathbb{R}$. Soit x_0 un point adhérent à E n'appartenant pas à E. Si $f(x)$ a une limite ℓ lorsque $x \to x_0$, on peut étendre le domaine de définition de f à $E \cup \{x_0\}$ en posant $f(x_0) = \ell$. La fonction étendue est continue en x_0. On dit que l'on a obtenu un ***prolongement de f par continuité*** au point x_0.

Exemple 18.
Soit $f : \mathbb{R}^2 \setminus \{(0,0)\}$ définie par
$$f(x,y) = \frac{xy}{\sqrt{x^2+y^2}}.$$

Est-il possible de prolonger f par continuité en $(0,0)$?
Sur la figure ci-dessous, la question devient simplement : est-il possible de boucher le trou au milieu de la surface en rajoutant juste un point ?



Solution.

- **Limite à l'origine.**
On utilise que $|x| \leqslant \sqrt{x^2+y^2}$ et $|y| \leqslant \sqrt{x^2+y^2}$. Donc
$$|f(x,y)| = \frac{|x|\cdot|y|}{\sqrt{x^2+y^2}} \leqslant \sqrt{x^2+y^2} \xrightarrow[(x,y)\to(0,0)]{} 0.$$
- **Prolongement.**
Pour prolonger f en $(0,0)$, on choisit comme valeur la limite obtenue. On pose donc $f(0,0)=0$. (On note encore $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction prolongée.)
- **Continuité.**
Par notre choix de $f(0,0)$, f est continue en $(0,0)$. En dehors de l'origine, f est continue comme somme, produit, composition, inverse de fonctions continues. Conclusion : la fonction prolongée est continue sur $\mathbb{R}^2$ tout entier.

Mini-exercices.

1. Soit $f(x,y) = \frac{1+x}{1+y}$. Trouver un ouvert U contenant l'origine tel que $0.999 < f(x,y) < 1.001$ pour tout $(x,y) \in U$.
2. Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue au point $(x_1,\ldots,x_n)$. Montrer que la fonction partielle $f_i : \mathbb{R} \longrightarrow \mathbb{R}$ définie par $f_i(t) = f(x_1,\ldots,x_{i-1},t,x_{i+1},\ldots,x_n)$ est continue en x_i.
3. Sachant que la limite de $f(x,y) = \frac{1+x}{1+y}$ en $(0,0)$ est 1, calculer la limite des fonctions suivantes en $(0,0)$: $\frac{1+x}{1+y} + x^2 + y^2$; $\frac{1+y}{1+x}$; $\sin(xy)\frac{1+x}{1+y}$; $\ln\left(\frac{1+x}{1+y}\right)$.
4. Sachant que $\ln(t) \leqslant t-1$ pour tout $t > 0$, calculer la limite de $\frac{\ln(1+xy)}{1+x^2+y^4}$ en $(0,0)$.
5. Soit $f(x,y) = \frac{\sin(x^2-y^2)}{x^2+y^2}$. Soit $\gamma(t) = (at, bt)$ où $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$ est fixé. Calculer la limite de $(f\circ\gamma)(t)$ lorsque $t \to 0$ en fonction de (a,b). f admet-elle une limite en $(0,0)$? f est-elle prolongeable par continuité en $(0,0)$?
6. Soit f définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$ par $f(x,y) = \frac{xy^3}{x^2+2y^2}$. f admet-elle une limite en $(0,0)$? f est-elle prolongeable par continuité en $(0,0)$? Mêmes questions avec $f(x,y) = \frac{xy^3}{x^4+2y^4}$.

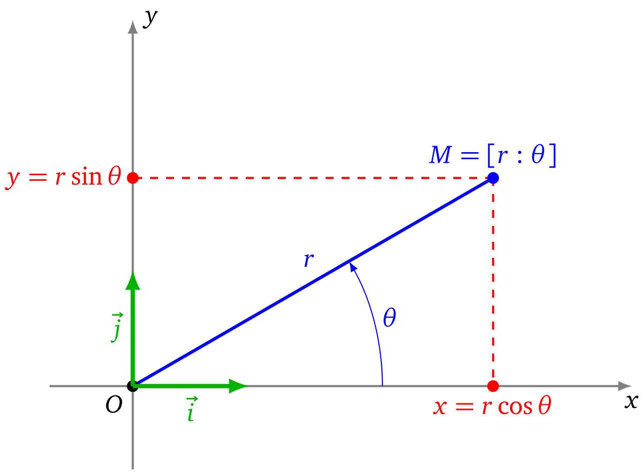
4. Coordonnées polaires

Plutôt que de repérer un point du plan $\mathbb{R}^2$ par ses coordonnées cartésiennes (x, y), on peut le faire au moyen de sa distance à l'origine et de l'angle formé avec l'horizontale : ce sont les coordonnées polaires.

4.1. Définition

Soit M un point du plan $\mathbb{R}^2$. Soit $O = (0, 0)$ l'origine. Soit $(O, \vec{i}, \vec{j})$ un repère orthonormé direct.

- On note $r = \|\overrightarrow{OM}\|$, la distance de M à l'origine.
- On note θ l'angle entre $\vec{i}$ et $\overrightarrow{OM}$.



On note $[r : \theta]$ les *coordonnées polaires* du point M. Dans ce cours, r sera toujours positif. L'angle n'est pas déterminé de manière unique, plusieurs choix sont possibles. Pour avoir unicité, on peut limiter θ à l'intervalle $[0, 2\pi[$, ou bien $]-\pi, +\pi]$. On n'attribue généralement pas de coordonnées polaires au point origine (l'angle n'aurait pas de sens).

Coordonnées polaires vers coordonnées cartésiennes.
On retrouve les coordonnées cartésiennes (x, y) à partir des coordonnées polaires $[r : \theta]$ par les formules

$$x = r\cos\theta \qquad \text{et} \qquad y = r\sin\theta.$$

Autrement dit, on a défini une application :

$$]0, +\infty[\times[0, 2\pi[\to \mathbb{R}^2 \qquad (r, \theta) \mapsto (r\cos\theta, r\sin\theta).$$

Coordonnées cartésiennes vers coordonnées polaires.
On retrouve r et θ à partir de (x, y) par les formules suivantes :

$$r = \sqrt{x^2 + y^2}$$

et, dans le cas $x > 0$ et $y \geqslant 0$,

$$\theta = \arctan\left(\frac{y}{x}\right).$$

Pour les points dans les autres quadrants, on se ramène au quadrant principal où $x > 0$ et $y \geqslant 0$.

4.2. Limite et continuité

Lorsque l'on considère des applications $f : E \subset \mathbb{R}^2 \to \mathbb{R}$, il est quelquefois plus facile de prouver des résultats de limite, continuité, etc., en passant par les coordonnées polaires.

Proposition 3.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction définie au voisinage de $(0,0) \in \mathbb{R}^2$, sauf peut-être en $(0,0)$. Si

$$\lim_{r\to 0} f(r\cos\theta, r\sin\theta) = \ell \in \mathbb{R}$$

existe indépendamment de θ, c'est-à-dire qu'il existe une fonction $\epsilon(r) \xrightarrow[r\to 0]{} 0$ telle que, pour tout $r \geqslant 0$ et tout θ, on a :

$$\left| f(r\cos\theta, r\sin\theta) - \ell \right| \leqslant \epsilon(r),$$

alors $\displaystyle\lim_{(x,y)\to(0,0)} f(x,y) = \ell$.

Pour clarifier cette proposition et expliquer les différents cas pratiques de la limite, voici comment faire. On exprime $f(x,y)$ en coordonnées polaires en calculant $f(r\cos\theta, r\sin\theta)$.

1. Si $\lim_{r\to 0} f(r\cos\theta, r\sin\theta)$ existe et si elle ne dépend pas de la variable θ, alors cette limite est la limite de f au point $(0,0)$.
2. Si $\lim_{r\to 0} f(r\cos\theta, r\sin\theta)$ n'existe pas, alors f n'a pas de limite au point $(0,0)$.
3. Si $\lim_{r\to 0} f(r\cos\theta, r\sin\theta) = \ell(\theta)$ dépend de θ, alors f n'a pas de limite au point $(0,0)$. Pour le justifier, on donne deux valeurs θ_1 et θ_2 telles que $\ell(\theta_1) \neq \ell(\theta_2)$.

Voyons un exemple de chaque situation.

Exemple 19.

1. $f(x,y) = \dfrac{x^3}{x^2+y^2}$

$$f(r\cos\theta, r\sin\theta) = \frac{r^3\cos^3\theta}{r^2(\cos^2\theta + \sin^2\theta)} = \frac{r^3\cos^3\theta}{r^2} = r\cos^3\theta$$

   Comme $\left|\cos^3\theta\right| \leqslant 1$ alors $r|\cos^3\theta| \leqslant r$ (pour tout r mais aussi pour tout θ) avec $\epsilon(r) := r \xrightarrow[r\to 0]{} 0$. Ce qui implique que $f(r\cos\theta, r\sin\theta) \xrightarrow[r\to 0]{} 0$. La limite existe (indépendamment des valeurs prises par θ), donc la fonction f admet bien une limite en $(0,0)$: $f(x,y) \xrightarrow[(x,y)\to(0,0)]{} 0$.

   Pour ceux qui voudraient tout faire à la main avec plus de détails, on peut aussi écrire $\left|f(r\cos\theta, r\sin\theta)\right| \leqslant r$, autrement dit $\left|f(x,y)\right| \leqslant \sqrt{x^2+y^2}$. Donc $f(x,y) \xrightarrow[(x,y)\to(0,0)]{} 0$.

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

$$f(r\cos\theta, r\sin\theta) = \frac{r\sin\theta}{r^2(\cos^2\theta + r\sin^3\theta)} = \frac{1}{r}\,\frac{\sin\theta}{\cos^2\theta + r\sin^3\theta}$$

   Fixons θ tel que $\sin\theta \neq 0$ (c'est-à-dire $\theta \not\equiv 0 \pmod{\pi}$). Alors, lorsque $r \to 0$, $f(r\cos\theta, r\sin\theta)$ n'a pas de limite. En particulier, la fonction $(x,y) \mapsto f(x,y)$ n'a pas de limite en $(0,0)$.

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

$$f(r\cos\theta, r\sin\theta) = \frac{r^2\cos\theta\sin\theta}{r^2} = \cos\theta\sin\theta = \frac{1}{2}\sin(2\theta)$$

   Pour θ fixé, la fonction $r \mapsto f(r\cos\theta, r\sin\theta)$ admet bien une limite $\ell(\theta) = \frac{1}{2}\sin(2\theta)$, lorsque $r \to 0$. Mais cette limite dépend de l'angle θ : si $\theta = 0$, $\ell(\theta) = 0$; par contre, si $\theta = \frac{\pi}{4}$, $\ell(\theta) = \frac{1}{2}$. Comme la limite dépend de l'angle, alors la fonction de deux variables $(x,y) \mapsto f(x,y)$ n'a pas de limite en $(0,0)$.

4.3. Un exemple

Cet exemple est assez subtil et peut être passé en première lecture.

Remarque.
Soit $\ell \in \mathbb{R}$. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction telle que, pour chaque θ fixé, $\lim_{r\to 0} f(r\cos\theta, r\sin\theta) = \ell$. Peut-on en conclure que f admet ℓ pour limite au point $(0,0)$? La réponse est non !
Autrement dit, regarder la limite de f le long des rayons ne permet pas de trouver la limite de f à l'origine.

Attention : la différence entre cette remarque et la proposition 3 est subtile. Dans la proposition 3, on a une hypothèse en terme de limites du type :

$$\forall \epsilon \quad \exists r_0 \quad \forall r < r_0 \quad \forall \theta \qquad \cdots$$

alors que dans la remarque, on note que l'hypothèse (plus faible) suivante est insuffisante :

$$\forall \theta \quad \forall \epsilon \quad \exists r_0 \quad \forall r < r_0 \quad \ldots$$

Exemple 20.
Soit la fonction f définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$ par

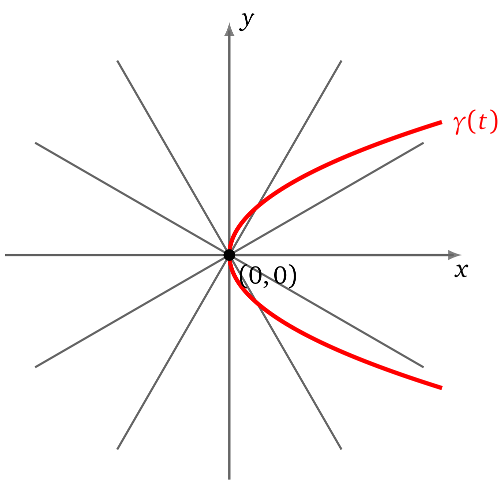
$$f(x,y) = \frac{xy^2}{x^2+y^4}.$$

1. Le long de tous les rayons f tend vers 0, c'est-à-dire, pour θ fixé,
$$f(r\cos\theta, r\sin\theta) \xrightarrow[r\to 0]{} 0.$$
2. Cependant, f n'a pas de limite en $(0,0)$.

Solution.

1. Calculons d'abord :
$$f(r\cos\theta, r\sin\theta) = \frac{r\cos\theta\sin^2\theta}{\cos^2\theta + r^2\sin^4\theta}$$
Fixons θ et discutons selon sa valeur :
   - Si $\cos\theta \neq 0$, alors le numérateur tend vers 0, tandis que le dénominateur tend vers $\cos^2\theta \neq 0$. Donc $f(r\cos\theta, r\sin\theta) \to 0$ lorsque $r \to 0$.
   - Si $\cos\theta = 0$, alors on se trouve sur des points (x,y) où $x = 0$ et donc $f(r\cos\theta, r\sin\theta) = f(0,y) = 0$.

   Dans tous les cas, f tend vers 0 sur tous les rayons définis par un angle θ fixé.
2. Considérons le chemin $\gamma(t) = (t^2, t)$. Alors
$$(f\circ\gamma)(t) = \frac{t^4}{2t^4} = \frac{1}{2}.$$
Mais on a vu que le long des rayons f tend vers 0. Cela contredit l'existence d'une limite pour $f(x,y)$ en $(0,0)$.



Mini-exercices.

1. Calculer l'angle θ des coordonnées polaires $[r : \theta]$ d'un point (x, y) dans le cas $x > 0$, $y < 0$. Puis faire les cas où $x < 0$.
2. La fonction f définie par $f(x, y) = \frac{(2x+3y)^3}{x^2+y^2}$ admet-elle une limite au point $(0, 0)$? Même question avec $f(x, y) = \frac{(2x+3y)^2}{x^2+y^2}$, puis $f(x, y) = \frac{2x+3y}{x^2+y^2}$.

Auteurs du chapitre

Arnaud Bodin. D'après des cours de Abdellah Hanani (Lille), Goulwen Fichou et Stéphane Leborgne (Rennes), Laurent Pujo-Menjouet (Lyon). Relu par Anne Bauval, Vianney Combet et Barbara Tumpach.

Chapitre 2

Calcul différentiel

Pour une fonction de plusieurs variables, il y a une dérivée pour chacune des variables, qu'on appelle dérivée partielle. L'ensemble des dérivées partielles permet de reconstituer une approximation linéaire de la fonction : c'est la différentielle.

1. Dérivées partielles

Rappelons la notion de dérivée. Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction d'une seule variable. La ***dérivée*** de f en $x_0 \in \mathbb{R}$, si elle existe, est :

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

Exemple 1.
La fonction $f : \mathbb{R} \to \mathbb{R}$ définie par $f(x) = x^2$ est dérivable, de dérivée $f'(x_0) = 2x_0$. En effet, lorsque h tend vers 0, on a :

$$\frac{(x_0+h)^2 - x_0^2}{h} = 2x_0 + h \quad \underset{h \to 0}{\longrightarrow} \quad 2x_0.$$

1.1. Définition

Définition 1.
Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$, où U est un ouvert de $\mathbb{R}^n$. On dit que f admet une ***dérivée partielle*** par rapport à la variable x_i au point $x_0 = (a_1, \dots, a_n) \in \mathbb{R}^n$ si la fonction d'une variable

$$x_i \mapsto f(a_1, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_n)$$

est dérivable au point a_i. Dit autrement, on définit la dérivée partielle de f par rapport à x_i au point $x_0 = (a_1, \dots, a_n)$ par

$$\lim_{h \to 0} \frac{f(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_n) - f(a_1, \dots, a_n)}{h}$$

si cette limite existe.

Notation. Cette limite se note

$$\frac{\partial f}{\partial x_i}(x_0).$$

C'est la dérivée partielle de f par rapport à x_i au point x_0. Le symbole « ∂ » se lit « d rond ». Une autre notation est $\partial_{x_i} f(x_0)$ ou bien $f'_{x_i}(x_0)$.
Il y a donc n dérivées partielles au point x_0 :

$$\frac{\partial f}{\partial x_1}(x_0) \qquad \frac{\partial f}{\partial x_2}(x_0) \qquad \dots \qquad \frac{\partial f}{\partial x_n}(x_0)$$

Dans le cas d'une fonction de deux variables $(x,y) \mapsto f(x,y)$, on a :

$$\boxed{\frac{\partial f}{\partial x}(x_0,y_0) = \lim_{h\to 0} \frac{f(x_0+h,y_0)-f(x_0,y_0)}{h} \qquad \frac{\partial f}{\partial y}(x_0,y_0) = \lim_{k\to 0} \frac{f(x_0,y_0+k)-f(x_0,y_0)}{k}}$$

Remarque.
Pour une fonction d'une variable $f : \mathbb{R} \to \mathbb{R}$, on distingue le nombre dérivé $f'(x_0)$ et la fonction dérivée f' définie par $x \mapsto f'(x)$. Il en est de même avec les dérivées partielles. Pour $f : \mathbb{R}^2 \to \mathbb{R}$:

- $\frac{\partial f}{\partial x}(x_0,y_0)$ et $\frac{\partial f}{\partial y}(x_0,y_0)$ sont des nombres réels.
- $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont des fonctions de deux variables, par exemple :

$$\frac{\partial f}{\partial x} : \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \frac{\partial f}{\partial x}(x,y) \end{array}$$

1.2. Exemples

Méthode. Pour calculer une dérivée partielle par rapport à une variable, on n'utilise que rarement la définition avec les limites, car il suffit de dériver par rapport à cette variable en considérant les autres variables comme des constantes.

Exemple 2.
Calculer les dérivées partielles premières de la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

$$f(x,y) = x^2 e^{3y}.$$

Solution.
Pour calculer $\frac{\partial f}{\partial x}$, qui est la dérivée partielle de f par rapport à x, on considère que y est une constante et on dérive x^2e^{3y} comme si c'était une fonction de x :

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

Pour l'autre dérivée partielle $\frac{\partial f}{\partial y}$, on considère que x est une constante et on dérive x^2e^{3y} comme si c'était une fonction de y :

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

Exemple 3.
Pour $f : \mathbb{R}^3 \to \mathbb{R}$ définie par $f(x,y,z) = \cos(x+y^2)e^{-z}$, on a :

$$\frac{\partial f}{\partial x}(x,y,z) = -\sin(x+y^2)e^{-z} \qquad \frac{\partial f}{\partial y}(x,y,z) = -2y\sin(x+y^2)e^{-z} \qquad \frac{\partial f}{\partial z}(x,y,z) = -\cos(x+y^2)e^{-z}$$

Exemple 4.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ définie par $f(x_1,\ldots,x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$. Alors, pour $i = 1,\ldots,n$, on a

$$\frac{\partial f}{\partial x_i}(x_1,\ldots,x_n) = 2x_i.$$

Une fonction peut avoir des dérivées partielles sans être continue ! Nous allons le voir sur l'exemple suivant.

Exemple 5.
La fonction $f : \mathbb{R}^2 \to \mathbb{R}$ suivante admet des dérivées partielles en tout point mais n'est pas continue en $(0,0)$:

$$f(x,y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{en } (0,0) \end{cases}$$

1. **Non continuité à l'origine.**
   Le long du chemin $\gamma(t) = (t,t)$, pour $t \neq 0$, on a $f\big(\gamma(t)\big) = \frac{t^2}{2t^2} = \frac{1}{2}$ qui ne tend pas vers $f(0,0) = 0$. Donc f n'est pas continue en $(0,0)$.
2. **Dérivées partielles en dehors de l'origine.**
   On se place en un point $(x_0, y_0) \neq (0,0)$. Dans un voisinage de ce point, f est définie par $f(x,y) = \frac{xy}{x^2+y^2}$. La fonction $x \mapsto f(x, y_0)$ est donc continue et dérivable au voisinage de x_0. La dérivée partielle s'obtient en dérivant la fonction d'une variable $x \mapsto f(x, y_0)$. Ainsi, on a
   $$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{y_0^3 - x_0^2 y_0}{(x_0^2 + y_0^2)^2}.$$
   De même, en dérivant la fonction $y \mapsto f(x_0, y)$, on trouve
   $$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{x_0^3 - x_0 y_0^2}{(x_0^2 + y_0^2)^2}.$$
3. **Dérivées partielles à l'origine.**
   Comme la fonction f est définie en $(0,0)$ par une formule spéciale, il faut revenir à la définition de ce que sont les dérivées partielles à l'aide des limites.
   Pour calculer $\frac{\partial f}{\partial x}(0,0)$, on calcule en $(x_0, y_0) = (0,0)$:
   $$\frac{f(0+h,0) - f(0,0)}{h} = \frac{0}{h} = 0 \underset{h \to 0}{\longrightarrow} 0$$
   Donc $\frac{\partial f}{\partial x}(0,0) = 0$.
   De même :
   $$\frac{f(0,0+k) - f(0,0)}{k} = \frac{0}{k} = 0 \underset{k \to 0}{\longrightarrow} 0$$
   Donc $\frac{\partial f}{\partial y}(0,0) = 0$.

Conclusion : quel que soit le point $(x_0, y_0) \in \mathbb{R}^2$, les dérivées partielles $\frac{\partial f}{\partial x}(x_0, y_0)$ et $\frac{\partial f}{\partial y}(x_0, y_0)$ existent.

1.3. Dérivée directionnelle

Il est possible de généraliser la notion de dérivée partielle.

> **Définition 2.**
> Soit $f : \mathbb{R}^n \to \mathbb{R}$. Soit $v \in \mathbb{R}^n$ un vecteur non nul. La ***dérivée directionnelle*** de f en $x_0 \in \mathbb{R}^n$ suivant le vecteur v est définie, si elle existe, par
> $$D_v f(x_0) = \lim_{t \to 0} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

Pour une fonction $f : \mathbb{R}^2 \to \mathbb{R}$, la dérivée directionnelle au point (x_0, y_0) suivant le vecteur $v = (h,k)$ est donc donnée par

$$D_v f(x_0, y_0) = \lim_{t \to 0} \frac{f(x_0 + th, y_0 + tk) - f(x_0, y_0)}{t}.$$

Exemple 6.
Soit f la fonction définie sur $\mathbb{R}^2$ par

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

Étudier l'existence de la dérivée directionnelle de f suivant un vecteur non nul au point $(0,0)$.

Solution.
Pour tout vecteur $v=(h,k)$ non nul, on a :

$$\lim_{t\to0}\frac{f(0+th,0+tk)-f(0,0)}{t}=\lim_{t\to0}\frac{\frac{(th)^3+(tk)^3}{(th)^2+(tk)^2}-0}{t}=\frac{h^3+k^3}{h^2+k^2}.$$

Donc f admet une dérivée directionnelle suivant tout vecteur non nul au point $(0,0)$ et, lorsque $v=(h,k)$, $D_vf(0,0)=\frac{h^3+k^3}{h^2+k^2}$.

De façon générale, si le vecteur v est un vecteur de la base canonique, on retrouve une dérivée partielle. Soit $f:\mathbb{R}^2\to\mathbb{R}$.

1. Si $v=(1,0)$, on retrouve $D_vf(x,y)=\frac{\partial f}{\partial x}(x,y)$.
2. Si $v=(0,1)$, on retrouve $D_vf(x,y)=\frac{\partial f}{\partial y}(x,y)$.

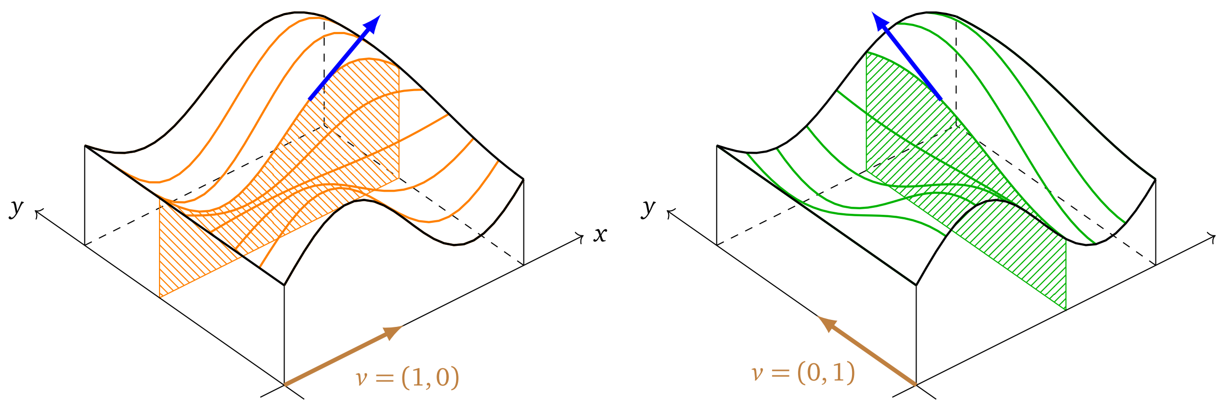
Lorsque f est différentiable (voir plus loin dans ce chapitre), nous aurons une formule simple et directe pour calculer $D_vf(x,y)$ à partir des dérivées partielles. Si f est différentiable et $v=(h,k)$ alors

$$D_vf(x,y)=h\frac{\partial f}{\partial x}(x,y)+k\frac{\partial f}{\partial y}(x,y).$$

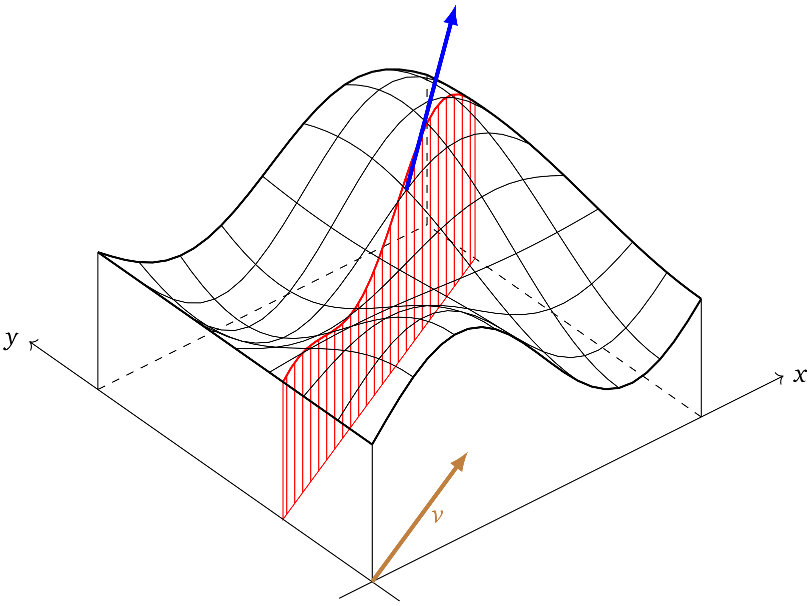
Interprétation géométrique.
Pour une fonction d'une variable, la dérivée en un point est la pente de la tangente au graphe de la fonction en ce point (le graphe est ici une courbe). Pour une fonction de deux variables $(x,y)\mapsto f(x,y)$, les dérivées partielles indiquent les pentes au graphe de f selon certaines directions (le graphe est ici une surface). Plus précisément :

- $\frac{\partial f}{\partial x}(x_0,y_0)$ est la pente au graphe de f en (x_0,y_0) suivant la direction de l'axe (Ox). En effet, cette pente est celle de la tangente à la courbe $z=f(x,y_0)$ et est donnée par la dérivée de $x\mapsto f(x,y_0)$ en x_0. C'est donc bien $\frac{\partial f}{\partial x}(x_0,y_0)$.
- $\frac{\partial f}{\partial y}(x_0,y_0)$ est la pente au graphe de f en (x_0,y_0) suivant la direction de l'axe (Oy).
- Plus généralement, si v est un vecteur unitaire (i.e. de norme 1) alors $D_vf(x_0,y_0)$ est la pente de la tangente suivant la direction v.

Sur la figure de gauche, la dérivée partielle $\frac{\partial f}{\partial x}$ indique la pente en un point d'une tranche parallèle à l'axe (Ox) (en orange). Sur la figure de droite, la dérivée partielle $\frac{\partial f}{\partial y}$ indique la pente en un point d'une tranche parallèle à l'axe (Oy) (en vert).



Ci-dessous, la dérivée directionnelle $D_v f$ indique la pente en un point d'une tranche (en rouge) dans la direction d'un vecteur v.



Mini-exercices.

1. En utilisant seulement la définition avec les limites, calculer les dérivées partielles de la fonction f définie par $f(x,y) = x^2 y$.
2. Calculer les dérivées partielles de la fonction f définie par $f(x,y) = e^{xy^2}$. Même question avec $f(x,y) = x^2 + 3y^2 - 2\sin(xy)$; $f(x,y) = \sqrt{1-x^2-y^2}$; $f(x,y,z) = xy^2 + ze^{y/z}$; $f(x_1,\ldots,x_n) = x_1 \ln(x_1 + \cdots + x_n)$.
3. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = 0$ si $0 < y < x^2$ et $f(x,y) = 1$ sinon. Montrer que f a des dérivées partielles en $(0,0)$, mais n'est pas continue en $(0,0)$.
4. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = xy^2 + x + y$. Calculer la dérivée directionnelle de f en $(0,0)$ suivant tout vecteur non nul $v = (h,k)$. Pour quel vecteur v unitaire cette dérivée est-elle maximale ?

2. Différentielle

La différentielle est une façon de regrouper toutes les dérivées partielles dans une seule fonction.

2.1. Différentiabilité

Pour une fonction $f : \mathbb{R} \to \mathbb{R}$ d'une seule variable, une autre façon d'écrire qu'elle est dérivable en x_0 est de vérifier qu'il existe $\ell \in \mathbb{R}$ tel que

$$\lim_{h \to 0} \frac{f(x_0+h) - f(x_0) - \ell \cdot h}{h} = 0.$$

Et on note ce ℓ par $f'(x_0)$, de sorte que l'on a $f(x_0+h) \simeq f(x_0) + f'(x_0) \cdot h$ (pour h réel, assez petit). Autrement dit, on approche l'application $h \mapsto f(x_0+h) - f(x_0)$ par une fonction linéaire $h \mapsto f'(x_0) \cdot h$.

Nous allons faire ce même travail en dimension supérieure.

Définition 3.
Soit $f : \mathbb{R}^n \to \mathbb{R}$. La fonction f est *différentiable* en $x_0 \in \mathbb{R}^n$ s'il existe une application linéaire $\ell : \mathbb{R}^n \to \mathbb{R}$

> telle que :
> $$\lim_{\|h\|\to 0} \frac{f(x_0+h)-f(x_0)-\ell(h)}{\|h\|} = 0$$
> L'application ℓ est la ***différentielle*** de f en x_0 et se note $\mathrm{d}f(x_0)$.

Dans le cas des fonctions d'une variable, on a $\mathrm{d}f(x_0) = f'(x_0)$ (et $\mathrm{d}f(x_0)(h) = f'(x_0)\cdot h$). Dans le cas des fonctions de plusieurs variables, on verra juste après comment écrire la différentielle à l'aide des dérivées partielles. Noter que $\mathrm{d}f(x_0)$ est une application de $\mathbb{R}^n$ vers $\mathbb{R}$ (comme f), et donc $\mathrm{d}f(x_0)(h)$ est un réel (pour chaque $h \in \mathbb{R}^n$).

De même qu'en une variable, si une fonction est dérivable, alors elle est continue.

> **Proposition 1.**
> *Si f est différentiable en $x_0 \in \mathbb{R}^n$, alors f est continue en x_0.*

Démonstration. Notons g la fonction définie par $g(h) = \frac{f(x_0+h)-f(x_0)-\mathrm{d}f(x_0)(h)}{\|h\|}$. Alors
$$f(x_0+h) = f(x_0) + \mathrm{d}f(x_0)(h) + \|h\|g(h)$$
et il est clair que $\mathrm{d}f(x_0)(h)$ et $\|h\|g(h)$ tendent vers 0 lorsque h tend vers le vecteur nul. Donc la limite de f en x_0 existe et vaut $f(x_0)$, et ainsi f est continue en x_0. □

Exemple 7.
Si $\ell : \mathbb{R}^n \to \mathbb{R}$ est linéaire, alors ℓ est différentiable et sa différentielle en tout point est l'application ℓ elle-même : pour tous $x_0 \in \mathbb{R}^n$ et $h \in \mathbb{R}^n$,
$$\mathrm{d}\ell(x_0)(h) = \ell(h).$$

2.2. Différentielle

> **Proposition 2.**
> *Si $f : \mathbb{R}^n \to \mathbb{R}$ est différentiable en $x_0 \in \mathbb{R}^n$, alors ses dérivées partielles existent et on a :*
> $$\mathrm{d}f(x_0)(h) = h_1\frac{\partial f}{\partial x_1}(x_0) + \cdots + h_n\frac{\partial f}{\partial x_n}(x_0)$$
> *où $h = (h_1,\ldots,h_n)$.*

En particulier, lorsqu'elle existe, la différentielle est unique.

Pour $f : \mathbb{R}^2 \to \mathbb{R}$ différentiable en (x_0, y_0), la formule est :
$$\boxed{\mathrm{d}f(x_0,y_0)(h,k) = h\tfrac{\partial f}{\partial x}(x_0,y_0) + k\tfrac{\partial f}{\partial y}(x_0,y_0)}$$

Démonstration. Prouvons la formule pour deux variables. Soit $f : \mathbb{R}^2 \to \mathbb{R}$ différentiable en $(x_0, y_0) \in \mathbb{R}^2$. Soit $\ell(h,k) = ah + bk$ sa différentielle. Alors, par définition, lorsque $\|(h,k)\| \to 0$, on a :
$$\frac{f(x_0+h,y_0+k)-f(x_0,y_0)-\ell(h,k)}{\|(h,k)\|} \longrightarrow 0$$
Pour $(h,k) = (t,0)$ avec $t > 0$ et $t \to 0$, on a donc :
$$\frac{f(x_0+t,y_0)-f(x_0,y_0)-t\ell(1,0)}{t} = \frac{f(x_0+t,y_0)-f(x_0,y_0)}{t} - \ell(1,0) \longrightarrow 0$$
C'est exactement dire que
$$\frac{\partial f}{\partial x}(x_0,y_0) = \ell(1,0) = a.$$

Avec $(h,k) = (0,t)$, on prouve de même que

$$\frac{\partial f}{\partial y}(x_0, y_0) = \ell(0,1) = b.$$

Ainsi,

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

□

Pour montrer qu'une fonction est différentiable, on peut utiliser que la somme, le produit, l'inverse (d'une fonction ne s'annulant pas) et la composition de fonctions différentiables est différentiable. Sinon, il faut revenir à la définition. Par exemple, pour $f : \mathbb{R}^2 \to \mathbb{R}$:

1. tout d'abord, on calcule les dérivées partielles $\frac{\partial f}{\partial x}(x_0, y_0)$ et $\frac{\partial f}{\partial y}(x_0, y_0)$,
2. on écrit le candidat à être la différentielle $\ell(h,k) = h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0)$,
3. il faut enfin prouver la limite, lorsque $\|(h,k)\| \to 0$:
$$\frac{f(x_0+h, y_0+k) - f(x_0, y_0) - \ell(h,k)}{\|(h,k)\|} \longrightarrow 0.$$

Exemple 8.
Étudier la différentiabilité en tout point de la fonction f définie par

$$f(x,y) = x - 3y + \frac{x^4}{x^2+y^2} \quad \text{si } (x,y) \neq (0,0) \qquad \text{et} \quad f(0,0) = 0.$$

Solution.

- En dehors de $(0,0)$, la fonction f est différentiable, car f est une somme, produit, inverse de fonctions différentiables (car $x^2 + y^2$ ne s'annule qu'à l'origine).
- En $(0,0)$, il faut étudier la différentiabilité à la main.
  — Dérivée partielle par rapport à x :
  $$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} = \lim_{h\to 0} \frac{h+h^2}{h} = 1$$
  — Dérivée partielle par rapport à y :
  $$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0} \frac{f(0,k) - f(0,0)}{k} = \lim_{k\to 0} \frac{-3k}{k} = -3$$
  — Le candidat à être la différentielle est donc :
  $$\ell(h,k) = h\frac{\partial f}{\partial x}(0,0) + k\frac{\partial f}{\partial y}(0,0) = h - 3k$$
  — On calcule :
  $$0 \leqslant \frac{f(0+h, 0+k) - f(0,0) - \ell(h,k)}{\sqrt{h^2+k^2}} = \frac{h^4}{(h^2+k^2)^{\frac{3}{2}}} \leqslant \frac{h^4}{|h|^3} = |h| \xrightarrow[(h,k)\to(0,0)]{} 0.$$
  Donc f est différentiable au point $(0,0)$ et $\mathrm{d}f(0,0)(h,k) = h - 3k$.

2.3. Lien avec les dérivées partielles

Dérivées partielles.
On a vu dans la proposition 2 que si $f : \mathbb{R}^2 \to \mathbb{R}$ est différentiable en (x_0, y_0), alors

$$\mathrm{d}f(x_0,y_0)(1,0) = \frac{\partial f}{\partial x}(x_0, y_0) \qquad \text{et} \qquad \mathrm{d}f(x_0,y_0)(0,1) = \frac{\partial f}{\partial y}(x_0, y_0).$$

En toute dimension, pour $f : \mathbb{R}^n \to \mathbb{R}$ différentiable en $x_0 \in \mathbb{R}^n$, et e_i le i-ème vecteur de la base canonique :

$$\mathrm{d}f(x_0)(e_i) = \frac{\partial f}{\partial x_i}(x_0).$$

Dérivée directionnelle.

Plus généralement, si $f : \mathbb{R}^n \to \mathbb{R}$ est différentiable en $x_0 \in \mathbb{R}^n$, alors $\mathrm{d}f(x_0)(v) = D_v f(x_0)$. Pour $f : \mathbb{R}^2 \to \mathbb{R}$, cela signifie que si $v = (h,k)$, alors :

$$\boxed{D_{(h,k)} f(x_0,y_0) = h\tfrac{\partial f}{\partial x}(x_0,y_0) + k\tfrac{\partial f}{\partial y}(x_0,y_0)}$$

Si f n'est pas différentiable, cette formule peut être fausse.

Gradient.

Le gradient est une autre façon de coder la différentielle. Le ***gradient*** de f en x_0 est le vecteur

$$\operatorname{grad} f(x_0) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x_0) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x_0) \end{pmatrix}.$$

Si f est différentiable en x_0, alors

$$\mathrm{d}f(x_0)(v) = \langle \operatorname{grad} f(x_0) \mid v \rangle,$$

où $\langle u \mid v \rangle$ est le produit scalaire de u et v.

Nous reviendrons en détail sur le gradient et ses applications dans le chapitre « Gradient – Théorème des accroissements finis ».

Résumé.

Lorsque f est différentiable alors la différentielle, la dérivée directionnelle, et le gradient encodent la même information et sont reliés par les formules :

$$\boxed{D_v f(x_0) = \mathrm{d}f(x_0)(v) = \langle \operatorname{grad} f(x_0) \mid v \rangle}$$

Exemple 9.

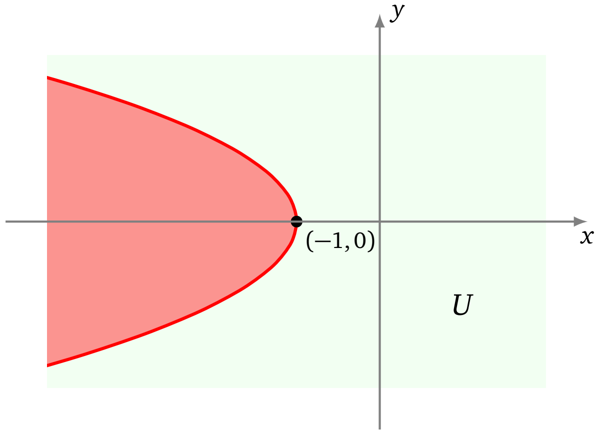
Soit f la fonction définie par

$$f(x,y) = \ln(1 + x + y^2).$$

1. Déterminer le domaine de définition U de f.
2. Calculer les dérivées partielles de f.
3. Montrer que f est différentiable sur U.
4. Calculer le gradient de f en $(0,1)$ et exprimer la différentielle en ce point.
5. Calculer la dérivée directionnelle de f en $(0,1)$ suivant le vecteur $(2,1)$.

Solution.

1. On a $U = \left\{(x,y) \in \mathbb{R}^2 \mid 1 + x + y^2 > 0\right\}$.



2. Les dérivées partielles sont :

$$\frac{\partial f}{\partial x}(x,y)=\frac{1}{1+x+y^2} \qquad \frac{\partial f}{\partial y}(x,y)=\frac{2y}{1+x+y^2}$$

3. f est différentiable sur U comme somme, produit et composée de fonctions différentiables.
4. Le gradient s'obtient directement à partir des dérivées partielles :

$$\operatorname{grad} f(0,1)=\begin{pmatrix}\frac{\partial f}{\partial x}(0,1)\\ \frac{\partial f}{\partial y}(0,1)\end{pmatrix}=\begin{pmatrix}\frac{1}{2}\\ 1\end{pmatrix}$$

La différentielle en ce point $\mathrm{d}f(0,1):\mathbb{R}^2\to\mathbb{R}$ est l'application linéaire définie par

$$\mathrm{d}f(0,1)(h,k)=\langle \operatorname{grad} f(0,1)\mid(h,k)\rangle=\frac{1}{2}h+k.$$

5. Comme f est différentiable, la dérivée directionnelle est simplement la combinaison linéaire des dérivées partielles :

$$D_{(2,1)}f(0,1)=2\times\frac{\partial f}{\partial x}(0,1)+1\times\frac{\partial f}{\partial y}(0,1)=2$$

On aurait aussi pu faire le calcul via la formule $D_{(2,1)}f(0,1)=\mathrm{d}f(0,1)(2,1)$.

Mini-exercices.

1. Soit $g:\mathbb{R}\to\mathbb{R}$ dérivable. Soit $f:\mathbb{R}^2\to\mathbb{R}$ définie par $f(x,y)=g(x+y)$. Montrer que f est différentiable et que $\mathrm{d}f(x_0,y_0)(h,k)=g'(x_0+y_0)h+g'(x_0+y_0)k$.
2. Soit $f(x,y)=2xy-7x+8y$. En utilisant la définition, montrer que f est différentiable et calculer sa différentielle.
3. Soit f définie par $f(x,y)=ye^{x/y}$. Trouver le domaine de définition U de f. Montrer que f est différentiable sur U. Calculer ses dérivées partielles. Calculer la dérivée directionnelle de f en $(4,2)$ suivant le vecteur $v=(-1,1)$.

3. Fonctions de classe $\mathscr{C}^1$

Dans la pratique, les fonctions seront souvent de classe $\mathscr{C}^1$, ce qui implique la différentiabilité, et est plus facile à vérifier.

3.1. Définition

Définition 4.
Soit $f : \mathbb{R}^n \to \mathbb{R}$. On dit que f est de ***classe $\mathscr{C}^1$*** si les dérivées partielles $\frac{\partial f}{\partial x_i}$ existent et sont continues (pour $i = 1, \dots, n$).

On peut bien sûr limiter la définition à un ouvert. Par exemple, si U un ouvert de $\mathbb{R}^2$, $f : U \to \mathbb{R}$ sera de classe $\mathscr{C}^1$ sur U si $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ existent et sont continues sur U.

Théorème 1.
Si f est de classe $\mathscr{C}^1$, alors f est différentiable.

Une autre façon de dire que f est différentiable est de dire que f admet un ***développement limité à l'ordre 1***.
Pour $f : \mathbb{R}^2 \to \mathbb{R}$, si f est différentiable au point (x_0, y_0), alors

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

On rappelle la notation « petit o ».
Notation. Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction définie au voisinage de $(0,0)$. On dit que g est ***négligeable*** par rapport à $\|(x, y)\|^n$ au voisinage de $(0,0)$ et on note $g = o\left(\|(x, y)\|^n\right)$ si

$$\lim_{(x,y)\to(0,0)} \frac{g(x, y)}{\|(x, y)\|^n} = 0.$$

Exemple 10.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x, y) = \sin x \cdot e^{2y}$.

- On a : $\frac{\partial f}{\partial x}(x, y) = \cos x \cdot e^{2y}$ et $\frac{\partial f}{\partial y}(x, y) = 2\sin x \cdot e^{2y}$. Les deux dérivées partielles existent et sont continues donc f est de classe $\mathscr{C}^1$ sur tout $\mathbb{R}^2$.
- En particulier, f est différentiable en tout point $(x_0, y_0) \in \mathbb{R}^2$ et
$$\mathrm{d}f(x_0, y_0)(h, k) = h\cos x_0 e^{2y_0} + 2k\sin x_0 e^{2y_0}.$$
- Par exemple, pour $(x_0, y_0) = (\frac{\pi}{6}, 1)$, on a le développement limité :
$$f(\tfrac{\pi}{6} + h, 1 + k) = f(\tfrac{\pi}{6}, 1) + h\frac{\partial f}{\partial x}(\tfrac{\pi}{6}, 1) + k\frac{\partial f}{\partial y}(\tfrac{\pi}{6}, 1) + o\left(\|(h,k)\|\right)$$
Ainsi :
$$f(\tfrac{\pi}{6} + h, 1 + k) = \frac{1}{2}e^2 + \frac{\sqrt{3}}{2}e^2 h + e^2 k + \epsilon(h, k)\sqrt{h^2 + k^2}$$
où $\epsilon(h, k) \to 0$ lorsque $(h, k) \to (0, 0)$.

Démonstration. Soit f une fonction de classe $\mathscr{C}^1$ au voisinage du point (x_0, y_0) : elle admet donc des dérivées partielles continues au voisinage de (x_0, y_0).
Pour $(h, k) \in \mathbb{R}^2$, on a

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \big[f(x_0 + h, y_0) - f(x_0, y_0)\big] + \big[f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)\big].$$

La fonction $x \mapsto f(x, y_0)$ est dérivable en x_0, donc

$$f(x_0 + h, y_0) - f(x_0, y_0) = h\frac{\partial f}{\partial x}(x_0, y_0) + h\epsilon_1(h)$$

avec $\lim_{h\to 0} \epsilon_1(h) = 0$.

Fixons h : la fonction $y \mapsto f(x_0+h, y)$ est dérivable autour de y_0. Appliquons le théorème des accroissements finis à cette fonction d'une variable sur l'intervalle $[y_0, y_0+k]$; il existe donc $\ell \in]0,k[$ tel que

$$f(x_0+h, y_0+k) - f(x_0+h, y_0) = k\frac{\partial f}{\partial y}(x_0+h, y_0+\ell).$$

Le réel ℓ dépend de k et de h et tend vers 0 lorsque k tend vers 0 (uniformément en h). Or, $\frac{\partial f}{\partial y}$ est continue au point (x_0, y_0), donc $\frac{\partial f}{\partial y}(x_0+h, y_0+\ell) = \frac{\partial f}{\partial y}(x_0, y_0) + \epsilon_2(h,k)$ avec $\lim_{(h,k)\to(0,0)} \epsilon_2(h,k) = 0$, d'où

$$f(x_0+h, y_0+k) - f(x_0+h, y_0) = k\frac{\partial f}{\partial y}(x_0, y_0) + k\epsilon_2(h,k).$$

Ainsi,

$$f(x_0+h, y_0+k) - f(x_0, y_0) = h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0) + h\epsilon_1(h) + k\epsilon_2(h,k).$$

Or

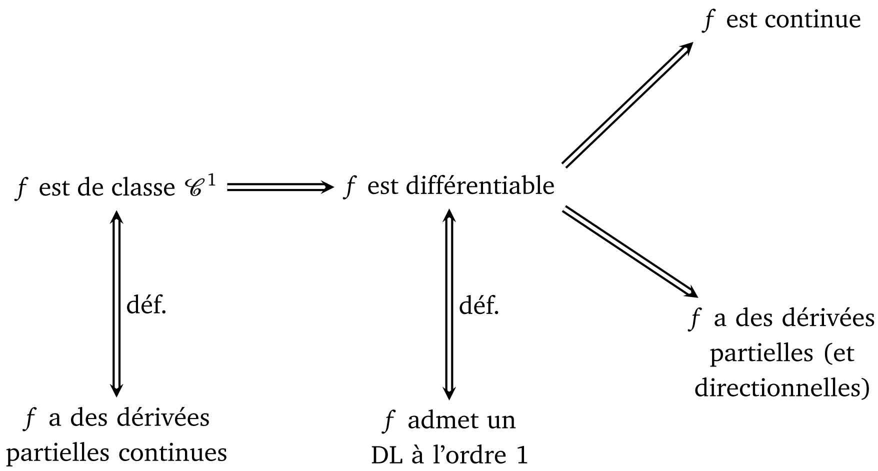
$$\frac{|h\epsilon_1(h) + k\epsilon_2(h,k)|}{\|(h,k)\|} \leqslant |\epsilon_1(h)| + |\epsilon_2(h,k)| \xrightarrow[(h,k)\to(0,0)]{} 0,$$

donc

$$f(x_0+h, y_0+k) - f(x_0, y_0) = \mathrm{d}f(x_0, y_0)(h,k) + o\left(\|(h,k)\|\right).$$

Ainsi f admet un développement limité d'ordre 1 au point (x_0, y_0), autrement dit f est différentiable en ce point. □

3.2. Résumé



- Les équivalences sont issues des définitions.
- Les implications viennent du théorème 1, de la proposition 1 et de la proposition 2.
- Les implications inverses sont fausses : voir les exemples ci-dessous.

3.3. Contre-exemples

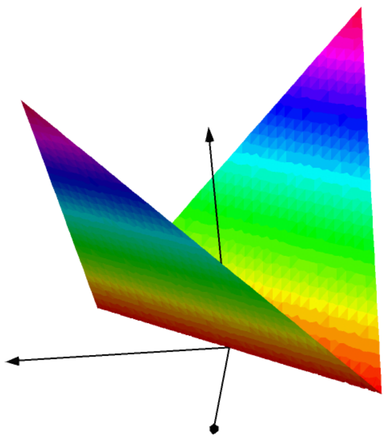
Dans cette partie, on justifie que les énoncés précédents ne peuvent pas être améliorés. Cette section peut être passée en première lecture.

Si f est différentiable, alors f est continue. La réciproque est fausse, comme le prouve l'exemple suivant.

Exemple 11.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = |x+y|$. Alors f est continue (comme somme et composée de fonctions continues), mais n'est pas différentiable. Par exemple, $\frac{\partial f}{\partial x}$ n'est pas définie en $(0,0)$, car

$$\frac{f(0+h,0) - f(0,0)}{h} = \frac{|h|}{h}$$

n'a pas de limite lorsque $h \to 0$. (Plus précisément, c'est $+1$ pour les $h > 0$ et -1 pour les $h < 0$.) Donc f n'est pas différentiable en $(0,0)$.



Sur la figure du graphe de f, on devine que, en tout point de la droite $(y = -x)$, f n'est pas différentiable.

Si f est différentiable, alors f admet des dérivées partielles et des dérivées directionnelles dans toutes les directions. La réciproque est fausse, comme le prouve l'exemple suivant.

Exemple 12.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$f(x,y) = \frac{y^3}{\sqrt{x^2+y^4}} \quad \text{si } (x,y) \neq (0,0) \qquad \text{et} \quad f(0,0) = 0.$$

Montrer que f admet une dérivée directionnelle suivant tout vecteur non nul au point $(0,0)$, mais qu'elle n'y est pas différentiable.

Solution.

1. Soit $v = (h,k) \neq (0,0)$.
   - Si $h = 0$, on a $\dfrac{f(t\cdot v) - f(0,0)}{t} = \dfrac{f(0,tk)}{t} = k$.
   - Si $h \neq 0$, on a $\left|\dfrac{f(t\cdot v) - f(0,0)}{t}\right| = \left|\dfrac{k^3 t^2}{\sqrt{h^2t^2 + h^4t^4}}\right| \leqslant \left|\dfrac{k^3}{h}\right| |t| \xrightarrow[t\to 0]{} 0$.

   Ainsi, $D_v f(0,0) = \begin{cases} k & \text{si} \quad h = 0, \\ 0 & \text{si} \quad h \neq 0. \end{cases}$
2. Avec $v = (1,0)$, on aura $\frac{\partial f}{\partial x}(0,0) = 0$, et avec $v = (0,1)$, on aura $\frac{\partial f}{\partial y}(0,0) = 1$. Le candidat à être la différentielle en $(0,0)$ est donc $\ell(h,k) = k$. Mais l'expression
$$\epsilon(h,k) = \frac{f(h,k) - f(0,0) - \ell(h,k)}{\sqrt{h^2+k^2}} = \frac{k^3 - k\sqrt{h^2+k^4}}{\sqrt{h^2+k^2}\sqrt{h^2+k^4}}$$
ne tend pas vers 0 lorsque $(h,k) \to 0$, car $\lim_{t\to 0^+} \epsilon(t,t) = -\frac{1}{\sqrt{2}}$. Donc f n'est pas différentiable au point $(0,0)$.

Si f est de classe $\mathscr{C}^1$, alors elle différentiable. La réciproque est fausse, comme le prouve l'exemple suivant.

Exemple 13.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ telle que

$$f(x,y) = y^2 \sin\left(\frac{1}{x^2+y^2}\right) \quad \text{si } (x,y) \neq (0,0) \qquad \text{et} \quad f(0,0) = 0.$$

Montrer que f est différentiable en tout point de $\mathbb{R}^2$ sans être de classe $\mathscr{C}^1$ à l'origine.

Solution.

- **En dehors de l'origine.**
  Les dérivées partielles
  $$\frac{\partial f}{\partial x}(x,y) = -\frac{2xy^2}{(x^2+y^2)^2}\cos\left(\frac{1}{x^2+y^2}\right)$$
  et
  $$\frac{\partial f}{\partial y}(x,y) = 2y\sin\left(\frac{1}{x^2+y^2}\right) - \frac{2y^3}{(x^2+y^2)^2}\cos\left(\frac{1}{x^2+y^2}\right)$$
  existent et sont continues sur $\mathbb{R}^2\setminus\{(0,0)\}$. Ainsi, f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2\setminus\{(0,0)\}$ et, par le théorème 1, est donc différentiable sur $\mathbb{R}^2\setminus\{(0,0)\}$.
- **Différentiabilité au point $(0,0)$.**
  — Calculons les dérivées partielles de f au point $(0,0)$. Comme $f(x,0)=0$ alors
  $$\frac{\partial f}{\partial x}(0,0) = \lim_{h\to 0}\frac{f(h,0)-f(0,0)}{h} = 0.$$
  Et comme $f(0,y) = y^2\sin\left(\frac{1}{y^2}\right)$, alors
  $$\frac{\partial f}{\partial y}(0,0) = \lim_{k\to 0}\frac{f(0,k)-f(0,0)}{k} = \lim_{k\to 0} k\sin\left(\frac{1}{k^2}\right) = 0.$$
  — Le candidat à être la différentielle en $(0,0)$ est donc $\ell(h,k)=0$.
  — De plus,
  $$\lim_{(h,k)\to(0,0)}\frac{f(h,k)-f(0,0)-\ell(h,k)}{\sqrt{h^2+k^2}} = \lim_{(h,k)\to(0,0)}\frac{k^2}{\sqrt{h^2+k^2}}\sin\left(\frac{1}{h^2+k^2}\right).$$
  Or $\left|\frac{k^2}{\sqrt{h^2+k^2}}\sin\left(\frac{1}{h^2+k^2}\right)\right| \leqslant |k| \xrightarrow[(h,k)\to(0,0)]{} 0$, donc f est différentiable au point $(0,0)$.
- **Conclusion.**
  La fonction f est différentiable sur $\mathbb{R}^2$. Par ailleurs,
  $$\frac{\partial f}{\partial x}(t,t) = -\frac{1}{2t}\cos\left(\frac{1}{2t^2}\right)$$
  n'a pas de limite lorsque $t\to 0$. La dérivée partielle $\frac{\partial f}{\partial x}$ n'est donc pas continue en $(0,0)$. Ainsi, f n'est pas de classe $\mathscr{C}^1$ à l'origine.

Mini-exercices.

1. Justifier que la fonction définie par $f(x,y) = \ln(1+x+2y)\cos(y)$ est différentiable sur son ensemble de définition. Écrire le développement limité à l'ordre 1 de f en $(0,0)$. Même travail avec $f(x,y) = \sqrt{1+x-2y}$.
2. Montrer que la fonction $f(x,y) = \frac{x^3}{x^2+y^2}$ si $(x,y)\neq(0,0)$ et $f(0,0)=0$ admet des dérivées partielles en tout point, mais n'est pas différentiable en $(0,0)$.
3. Montrer que la fonction de deux variables $f(x,y) = x^2\sin\left(\frac{1}{x}\right)$ si $x\neq 0$ et $f(x,y)=0$ sinon est différentiable, mais que $\frac{\partial f}{\partial x}$ n'est pas continue en $(0,0)$.

Auteurs du chapitre

Arnaud Bodin. D'après des cours de Abdellah Hanani (Lille), Goulwen Fichou et Stéphane Leborgne (Rennes), Laurent Pujo-Menjouet (Lyon). Relu par Anne Bauval, Vianney Combet et Barbara Tumpach.

Chapitre 3

Matrice jacobienne

Pour une fonction de plusieurs variables, il n'y a pas une dérivée mais plusieurs : une pour chaque variable. Si en plus la fonction est à valeurs vectorielles alors, pour chaque composante et pour chaque variable, il y a une dérivée. Toute ces dérivées sont regroupées dans la matrice jacobienne.

1. Matrice jacobienne

1.1. Fonctions vectorielles

Une fonction est dite *fonction vectorielle* lorsque l'espace d'arrivée n'est pas $\mathbb{R}$ mais $\mathbb{R}^p$, avec $p \geqslant 2$:

$$F : \begin{array}{ccc} \mathbb{R}^n & \longrightarrow & \mathbb{R}^p \\ x = (x_1, \ldots, x_n) & \longmapsto & \big(f_1(x), \ldots, f_p(x)\big) \end{array}$$

Chaque composante f_j, pour $j = 1, \ldots, p$, est une fonction de plusieurs variables à valeurs réelles : $f_j : \mathbb{R}^n \to \mathbb{R}$.
On note $x \mapsto F(x)$ ou bien encore $(x_1, \ldots, x_n) \mapsto F(x_1, \ldots, x_n)$.
On a déjà rencontré des fonctions à valeurs vectorielles. Quelques exemples :

- De $\mathbb{R}$ dans $\mathbb{R}^2$: $F(t) = (t^2, t)$.
- De $\mathbb{R}^2$ dans $\mathbb{R}^2$: $F(x, y) = (e^x \cos y, e^x \sin y)$.
- De $\mathbb{R}^2$ dans $\mathbb{R}^3$: $F(x, y) = (x^2, y^3, x^2 + y^2)$.
- De $\mathbb{R}^n$ dans $\mathbb{R}^n$: $F(x) = \dfrac{x}{\|x\|}$ où $x = (x_1, \ldots, x_n) \neq 0$.

Un exemple important est le cas d'une application linéaire.

- Par exemple, $L(x, y, z) = (2x + 3y - z, 5y - 7z)$ est une application linéaire $L : \mathbb{R}^3 \to \mathbb{R}^2$. Elle s'exprime aussi : $L(x, y, z) = A \times \left(\begin{smallmatrix} x \\ y \\ z \end{smallmatrix}\right)$ avec $A = \begin{pmatrix} 2 & 3 & -1 \\ 0 & 5 & -7 \end{pmatrix}$.
- Plus généralement, pour une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$, il existe une matrice A avec p lignes et n colonnes telle que

$$L(x_1, \ldots, x_n) = A \times \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

1.2. Matrice jacobienne

Soit $F : \mathbb{R}^n \to \mathbb{R}^p$ une fonction, dont les composantes sont $F = (f_1, \ldots, f_p)$. Soit $x \in \mathbb{R}^n$. On suppose que les dérivées partielles $\frac{\partial f_j}{\partial x_i}$ existent en x (pour tous $i = 1, \ldots, n$ et $j = 1, \ldots, p$).

Définition 1.

La ***matrice jacobienne*** de F en $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ est
$$J_F(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(x) & \cdots & \frac{\partial f_p}{\partial x_n}(x) \end{pmatrix}.$$

C'est une matrice à p lignes et n colonnes. La première ligne correspond aux dérivées partielles de f_1, la seconde ligne aux dérivées partielles de f_2, etc.

Voici ce que cela donne pour $F : \mathbb{R}^2 \to \mathbb{R}^2$ avec $F = (f_1, f_2)$, en $(x, y) \in \mathbb{R}^2$:
$$J_F(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y) \\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y) \end{pmatrix}$$

Exemple 1.
Soit $F : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F(x,y) = (x^2 + y^2, e^{x-y})$. Au point (x, y), on a :
$$J_F(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y) \\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 2x & 2y \\ e^{x-y} & -e^{x-y} \end{pmatrix}.$$
Par exemple, au point $(x_0, y_0) = (2, 1)$, la matrice jacobienne est
$$J_F(2,1) = \begin{pmatrix} 4 & 2 \\ e & -e \end{pmatrix}.$$

Exemple 2.
Les coordonnées polaires d'un point du plan définissent l'application $F : \mathbb{R}_+ \times [0, 2\pi[\to \mathbb{R}^2$, $F(r, \theta) = (r\cos\theta, r\sin\theta)$. Alors
$$J_F(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

Voyons une autre situation, où $F : \mathbb{R}^3 \to \mathbb{R}^2$ avec $F = (f_1, f_2)$, en $(x, y, z) \in \mathbb{R}^3$:
$$J_F(x,y,z) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y,z) & \frac{\partial f_1}{\partial y}(x,y,z) & \frac{\partial f_1}{\partial z}(x,y,z) \\ \frac{\partial f_2}{\partial x}(x,y,z) & \frac{\partial f_2}{\partial y}(x,y,z) & \frac{\partial f_2}{\partial z}(x,y,z) \end{pmatrix}$$

Exemple 3.
Pour $F(x,y,z) = (e^{xy}, z\sin x)$, on a
$$J_F(x,y,z) = \begin{pmatrix} ye^{xy} & xe^{xy} & 0 \\ z\cos x & 0 & \sin x \end{pmatrix}.$$

Exemple 4.
Soit $F : \mathbb{R} \to \mathbb{R}^p$ une fonction d'une seule variable, mais à valeurs vectorielles, définie par $F(x) = \big(f_1(x), \dots, f_p(x)\big)$. Alors
$$J_F(x) = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_p'(x) \end{pmatrix}.$$

1.3. Opérateurs différentiels classiques

Gradient

Pour une fonction à valeurs scalaires $f : \mathbb{R}^n \to \mathbb{R}$ dont les dérivées partielles existent, le vecteur ***gradient*** est :

$$\operatorname{grad} f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

C'est un vecteur colonne qui est la transposée de la matrice jacobienne (qui elle est ici un vecteur ligne) :

$$\operatorname{grad} f(x) = J_f(x)^T.$$

On reviendra en détail sur le gradient dans le chapitre « Gradient – Théorème des accroissements finis ».

Les physiciens notent le gradient $\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$ où ∇ (qui se lit « nabla ») correspond à l'opérateur

$$\nabla = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}.$$

Divergence

Pour une fonction $F : \mathbb{R}^n \to \mathbb{R}^n$ $(n = p)$ de composantes $f_1, \ldots, f_n$ dont toutes les dérivées partielles existent, on définit sa ***divergence*** par

$$\operatorname{div} F(x) = \operatorname{tr} J_F(x) = \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(x)$$

où $\operatorname{tr} J_F(x)$ est la trace de la matrice jacobienne.

Attention ! Ne pas confondre les notions de gradient et de divergence : $\operatorname{grad} F(x)$ est un vecteur alors que $\operatorname{div} F(x)$ est un nombre réel !

Les physiciens notent la divergence $\operatorname{div} F(x) = \nabla \cdot F(x)$, où $u \cdot v$ est le produit scalaire canonique des vecteurs u et v sur $\mathbb{R}^n$, ce qui fait que

$$\operatorname{div} F(x) = \nabla \cdot F(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix} \cdot \begin{pmatrix} f_1(x) \\ \vdots \\ f_n(x) \end{pmatrix}.$$

Exemple 5.

Soit $F : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $F(x, y, z) = \big(x^2 y, \sin(yz), e^{xyz}\big)$. Alors

$$\operatorname{div} F(x, y, z) = \frac{\partial f_1}{\partial x}(x, y, z) + \frac{\partial f_2}{\partial y}(x, y, z) + \frac{\partial f_3}{\partial z}(x, y, z) = 2xy + z\cos(yz) + xye^{xyz}.$$

Rotationnel en dimension 2

Pour une fonction $F : \mathbb{R}^2 \to \mathbb{R}^2$ de composantes f_1, f_2 dont toutes les dérivées partielles existent, on définit le ***rotationnel*** de F par

$$\operatorname{rot} F(x, y) = \frac{\partial f_2}{\partial x}(x, y) - \frac{\partial f_1}{\partial y}(x, y).$$

Le rotationnel est ici un nombre réel.

Exemple 6.
Soit $F(x,y) = \left(\frac{y}{x^3}, y\ln x\right)$ définie sur $]0,+\infty[\times\mathbb{R}$. Alors

$$\operatorname{rot} F(x,y) = \frac{\partial(y\ln x)}{\partial x} - \frac{\partial(\frac{y}{x^3})}{\partial y} = \frac{y}{x} - \frac{1}{x^3}.$$

Rotationnel en dimension 3
Pour une fonction $F : \mathbb{R}^3 \to \mathbb{R}^3$ de composantes f_1, f_2, f_3 dont toutes les dérivées partielles existent, on définit le ***rotationnel*** de F par

$$\operatorname{rot} F(x,y,z) = \begin{pmatrix} \frac{\partial f_3}{\partial y}(x,y,z) - \frac{\partial f_2}{\partial z}(x,y,z) \\ \frac{\partial f_1}{\partial z}(x,y,z) - \frac{\partial f_3}{\partial x}(x,y,z) \\ \frac{\partial f_2}{\partial x}(x,y,z) - \frac{\partial f_1}{\partial y}(x,y,z) \end{pmatrix}.$$

Le rotationnel est donc ici un vecteur. Pour se souvenir de la formule, les physiciens écrivent $\operatorname{rot} F(x,y,z) = \nabla \wedge F(x,y,z)$, où $u \wedge v$ désigne le produit vectoriel entre les vecteurs u et v :

$$\operatorname{rot} F(x,y,z) = \nabla \wedge F(x,y,z) = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \wedge \begin{pmatrix} f_1(x,y,z) \\ f_2(x,y,z) \\ f_3(x,y,z) \end{pmatrix}$$

Exemple 7.
Soit $F : \mathbb{R}^3 \to \mathbb{R}^3$ définie par $F(x,y,z) = (x^3, yz^2, xyz)$. Alors

$$\operatorname{rot} F(x,y,z) = \begin{pmatrix} xz - 2yz \\ -yz \\ 0 \end{pmatrix}.$$

1.4. Différentielle

Le pendant théorique de la matrice jacobienne est la différentielle associée à $F : \mathbb{R}^n \to \mathbb{R}^p$ en un point x. Cette section est plus théorique : pour une première lecture, on peut juste retenir que la différentielle $\mathrm{d}F(x)$ est une application linéaire dont la matrice (dans la base canonique) est la matrice jacobienne $J_F(x)$. Autrement dit :

$$\boxed{\mathrm{d}F(x)(h) = J_F(x) \times h}$$

où $x \in \mathbb{R}^n$ et $h \in \mathbb{R}^n$, alors que le résultat $\mathrm{d}F(x)(h)$ est un élément de $\mathbb{R}^p$.

Voici les explications de ces notions en détail. Les notions de limite et de continuité pour $F : \mathbb{R}^n \to \mathbb{R}^p$ sont similaires à celles des fonctions $f : \mathbb{R}^n \to \mathbb{R}$: on remplace dans l'espace d'arrivée la valeur absolue de $\mathbb{R}$ par une norme sur $\mathbb{R}^p$.
Nous allons voir ce qu'il en est pour la différentielle d'une fonction à valeurs vectorielles. Soit $F : \mathbb{R}^n \to \mathbb{R}^p$ dont les composantes sont $F = (f_1, \ldots, f_p)$ avec chaque $f_j : \mathbb{R}^n \to \mathbb{R}$.

Définition 2.
- $F : \mathbb{R}^n \to \mathbb{R}^p$ est ***différentiable*** en $x \in \mathbb{R}^n$ si chacune des composantes $f_j : \mathbb{R}^n \to \mathbb{R}$ ($j = 1, \ldots, p$) est différentiable en x. On note $\mathrm{d}f_j(x) : \mathbb{R}^n \to \mathbb{R}$ la différentielle de f_j en x.
- La ***différentielle*** d'une application vectorielle différentiable $F : \mathbb{R}^n \to \mathbb{R}^p$ en $x \in \mathbb{R}^n$ est l'application linéaire $\mathrm{d}F(x) : \mathbb{R}^n \to \mathbb{R}^p$ définie par

$$\mathrm{d}F(x) = \big(\mathrm{d}f_1(x), \ldots, \mathrm{d}f_p(x)\big).$$

Attention ! La différentielle $\mathrm{d}F(x)$ de F en $x \in \mathbb{R}^n$ est une application linéaire, donc c'est bien une fonction (et pas un vecteur). L'évaluation de cette fonction donne une expression avec des vecteurs :

$$\forall h \in \mathbb{R}^n \qquad \mathrm{d}F(x)(h) = \big(\mathrm{d}f_1(x)(h), \dots, \mathrm{d}f_p(x)(h)\big).$$

Proposition 1.
Soit $F : \mathbb{R}^n \to \mathbb{R}^p$ différentiable en $x \in \mathbb{R}^n$. Alors

$$\boxed{\mathrm{d}F(x)(h) = J_F(x) \times h}$$

où $J_F(x)$ est la matrice jacobienne de F en x, quel que soit $h \in \mathbb{R}^n$.

Autrement dit, trouver la différentielle en x revient à calculer la matrice jacobienne en x. Cette proposition découle de l'expression de chaque différentielle $\mathrm{d}f_j(x)$ à l'aide des dérivées partielles $\frac{\partial f_j}{\partial x_i}$ $(i = 1, \dots, n)$.

Exemple 8.
Soit $F : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F(x,y) = (ye^{x^2}, x^2 - y)$.
Calculons $\mathrm{d}F(x,y)(h,k)$ quels que soient (x,y), $(h,k) \in \mathbb{R}^2$.

- La matrice jacobienne de F est :
$$J_F(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y) \\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y) \end{pmatrix} = \begin{pmatrix} 2xye^{x^2} & e^{x^2} \\ 2x & -1 \end{pmatrix}$$
- En (x,y) et pour $(h,k) \in \mathbb{R}^2$, on a donc :
$$\mathrm{d}F(x,y)(h,k) = J_F(x,y) \times \begin{pmatrix} h \\ k \end{pmatrix} = \begin{pmatrix} (2xyh + k)e^{x^2} \\ 2xh - k \end{pmatrix}$$
- Par exemple, au point $(x_0, y_0) = (1,1)$, on a $\mathrm{d}F(1,1)(h,k) = \big((2h+k)e, 2h-k\big)$.

Remarque.

- Si F a des composantes de classe $\mathscr{C}^1$ (c'est-à-dire toutes les dérivées partielles existent et sont continues), alors elles sont différentiables et F est également différentiable.
- Si F est différentiable en x, alors F est continue en x.
- Si $L : \mathbb{R}^n \to \mathbb{R}^p$ est une application linéaire alors, en tout point, sa différentielle est l'application elle-même : autrement dit, $\mathrm{d}L(x) = L$, pour tout $x \in \mathbb{R}^n$.

Remarque.
Il existe une autre définition équivalente des deux notions rencontrées.

- $F : \mathbb{R}^n \to \mathbb{R}^p$ est ***différentiable*** en $x \in \mathbb{R}^n$ s'il existe une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^p$ telle que :
$$\lim_{\|h\| \to 0} \frac{F(x+h) - F(x) - L(h)}{\|h\|} = 0.$$
- Dans ce cas, L est la ***différentielle*** de F en x et on la note $\mathrm{d}F(x)$.

Mini-exercices.

1. Soient $F, G : \mathbb{R}^n \to \mathbb{R}^p$. Soient $x, y \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$. Justifier que les égalités suivantes sont vraies : $J_{F+G}(x) = J_F(x) + J_G(x)$; $J_{\lambda F}(x) = \lambda J_F(x)$. Trouver un exemple où $J_F(x+y)$ n'est pas égal à $J_F(x) + J_F(y)$.
2. Calculer en tout point la matrice jacobienne de l'application F définie par $F(x,y) = (x^2+y^2, e^{xy}, x+y)$. Même question avec $F(x,y,z) = (x^{y+z}, z\arctan(y))$.
3. Calculer la divergence et le rotationnel de F définie par $F(x,y) = (y\,\mathrm{sh}(x), \mathrm{ch}(x/y))$. On rappelle que $\mathrm{ch}\,x = \frac{e^x + e^{-x}}{2}$ et $\mathrm{sh}\,x = \frac{e^x - e^{-x}}{2}$. Même question avec $F(x,y,z) = (x+yz, \sin(y)\sin(z), \sqrt{x+z})$.
4. À quelle condition sur la matrice jacobienne $J_F(x)$ la différentielle $\mathrm{d}F(x)$ est-elle bijective ?
5. Exprimer la différentielle de $F(x,y) = \left(\frac{1}{x}\ln(y-1), \frac{e^y - x}{x^2}\right)$ en tout point $(x,y) \in \mathbb{R}^* \times]1, +\infty[$.

2. Matrice jacobienne d'une composée

Les dérivées partielles d'une composée de fonctions sont compliquées à obtenir. C'est l'objet de cette section.

2.1. Formule

Rappelons tout d'abord la formule de dérivée d'une composée pour les fonctions de $\mathbb{R}$ dans $\mathbb{R}$.

> **Proposition 2.**
> *Soient $f, g : \mathbb{R} \to \mathbb{R}$ des fonctions dérivables. Alors $g \circ f$ est dérivable et*
>
> $$\boxed{(g \circ f)'(x) = g'\big(f(x)\big) \times f'(x)}$$

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "f"] \arrow[rr, bend left, "g \circ f"] & \mathbb{R} \arrow[r, "g"] & \mathbb{R} \\
x \arrow[r, mapsto] & f(x) & \\
& y \arrow[r, mapsto] & g(y)
\end{tikzcd}$$

Remarque.
Il peut être intéressant de nommer x la variable de la fonction f et y la variable de la fonction g. La formule peut alors aussi s'écrire :

$$\frac{\mathrm{d}(g \circ f)}{\mathrm{d}x}(x) = \frac{\mathrm{d}g}{\mathrm{d}y}(f(x)) \times \frac{\mathrm{d}f}{\mathrm{d}x}(x).$$

En notant $y = f(x)$, alors on peut considérer g comme une fonction de la variable y, mais aussi (par composée) de la variable x. On peut alors écrire comme les physiciens :

$$\frac{\mathrm{d}g}{\mathrm{d}x} = \frac{\mathrm{d}g}{\mathrm{d}y} \times \frac{\mathrm{d}y}{\mathrm{d}x}.$$

C'est une formule que l'on mémorise facilement en disant que l'on simplifie la fraction en éliminant les $\mathrm{d}y$ au numérateur et au dénominateur.

Passons maintenant au cas de $F : \mathbb{R}^n \to \mathbb{R}^p$ et $G : \mathbb{R}^p \to \mathbb{R}^q$. La composée est alors $G \circ F : \mathbb{R}^n \to \mathbb{R}^q$, et est bien sûr définie par $(G \circ F)(x) = G\big(F(x)\big)$.

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "F"] \arrow[rr, bend left, "G \circ F"] & \mathbb{R}^p \arrow[r, "G"] & \mathbb{R}^q \\
x \arrow[r, mapsto] & F(x) & \\
& y \arrow[r, mapsto] & G(y)
\end{tikzcd}$$

> **Théorème 1.**
> *Si F et G sont différentiables alors $G \circ F$ est différentiable et les matrices jacobiennes sont reliées par la formule suivante :*
>
> $$\boxed{J_{G \circ F}(x) = J_G\big(F(x)\big) \times J_F(x)}$$

Ici, « $\times$ » est le produit des deux matrices jacobiennes.

On rappelle en particulier que si les composantes de F et G sont de classe $\mathscr{C}^1$ (i.e. les dérivées partielles existent et sont continues) alors les fonctions sont différentiables et la formule est valable. Et en plus $G \circ F$ est de classe $\mathscr{C}^1$.

Attention ! Noter que $J_F(x)$ et $J_{G \circ F}(x)$ sont des matrices jacobiennes calculées en x mais que, dans la formule, ***$J_G\big(F(x)\big)$ est la matrice jacobienne de G en $F(x)$*** (et pas en x, ce qui pourrait même ne pas avoir de sens). C'est une source fréquente d'erreurs !

Exemple 9.
Soient $F : \mathbb{R}^2 \to \mathbb{R}^2$, $F(x,y) = (x+y, e^{2x-y})$ et $G : \mathbb{R}^2 \to \mathbb{R}^3$, $G(x,y) = (xy, y\sin x, x^2)$. Les matrices jacobiennes de F et de G sont :

$$J_F(x,y) = \begin{pmatrix} 1 & 1 \\ 2e^{2x-y} & -e^{2x-y} \end{pmatrix} \qquad J_G(x,y) = \begin{pmatrix} y & x \\ y\cos x & \sin x \\ 2x & 0 \end{pmatrix}$$

Attention, nous avons besoin de $J_G(F(x,y))$. Donc, dans $J_G(x,y)$, on remplace x par la première composante de F (c'est $x+y$) et y par la seconde composante de F (c'est e^{2x-y}). Ainsi,

$$J_G(F(x,y)) = \begin{pmatrix} e^{2x-y} & x+y \\ e^{2x-y}\cos(x+y) & \sin(x+y) \\ 2(x+y) & 0 \end{pmatrix}.$$

Pour obtenir la matrice jacobienne de la composée $G \circ F : \mathbb{R}^2 \to \mathbb{R}^3$, on applique la formule donnée par le produit de matrices :

$$J_{G\circ F}(x,y) = J_G\big(F(x,y)\big) \times J_F(x,y)$$

On trouve

$$J_{G\circ F}(x,y) = \begin{pmatrix} (1+2x+2y)e^{2x-y} & (1-x-y)e^{2x-y} \\ \big(\cos(x+y)+2\sin(x+y)\big)e^{2x-y} & \big(\cos(x+y)-\sin(x+y)\big)e^{2x-y} \\ 2x+2y & 2x+2y \end{pmatrix}.$$

Voici la version du théorème en termes de différentielles.

> **Théorème 2.**
> *Si $F : \mathbb{R}^n \to \mathbb{R}^p$ est différentiable en x, et si $G : \mathbb{R}^p \to \mathbb{R}^q$ est différentiable en $F(x)$, alors $G \circ F : \mathbb{R}^n \to \mathbb{R}^q$ est différentiable en x et on a :*
> $$\mathrm{d}(G \circ F)(x) = \mathrm{d}G(F(x)) \circ \mathrm{d}F(x).$$

Autrement dit, l'application linéaire $\mathrm{d}(G \circ F)(x)$ est la composée de l'application linéaire $\mathrm{d}G(F(x))$ avec l'application linéaire $\mathrm{d}F(x)$.

2.2. Applications

Nous allons appliquer la formule de la matrice jacobienne d'une composée pour calculer des dérivées partielles. Le plus compliqué est d'identifier quelles sont les fonctions à composer et de s'adapter aux noms des variables qui peuvent changer selon les situations.
Les deux seules choses à retenir, c'est d'abord la formule $J_{G\circ F}(x) = J_G\big(F(x)\big) \times J_F(x)$, et ensuite comment l'appliquer. Il est donc inutile d'apprendre les formules qui suivent.

Cas $F : \mathbb{R} \to \mathbb{R}^2$, $G : \mathbb{R}^2 \to \mathbb{R}$

> **Proposition 3.**
> *Soit $F : \mathbb{R} \to \mathbb{R}^2$, $t \mapsto F(t) = (x(t), y(t))$ une fonction, avec $t \mapsto x(t)$ et $t \mapsto y(t)$ dérivables, et soit $G : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto G(x,y)$ une fonction différentiable. Alors $h = G \circ F : \mathbb{R} \to \mathbb{R}$, $t \mapsto h(t) = G(x(t), y(t))$*

est dérivable et

$$h'(t) = \frac{\partial G}{\partial x}\big(x(t), y(t)\big) \cdot x'(t) + \frac{\partial G}{\partial y}\big(x(t), y(t)\big) \cdot y'(t).$$

C'est une application directe de la formule $J_h(t) = J_G(F(t)) \times J_F(t)$, avec :

$$J_h(t) = \frac{\mathrm{d}h}{\mathrm{d}t}(t) = h'(t) \qquad J_G(x,y) = \begin{pmatrix} \frac{\partial G}{\partial x}(x,y) & \frac{\partial G}{\partial y}(x,y) \end{pmatrix} \qquad J_F(t) = \begin{pmatrix} \frac{\mathrm{d}x}{\mathrm{d}t}(t) \\ \frac{\mathrm{d}y}{\mathrm{d}t}(t) \end{pmatrix} = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}$$

Exemple 10.
Soit $G(x,y) = \cos(y)e^x$. Calculer la dérivée de la fonction $h : t \mapsto G(t^2, \sin t)$.

Solution.
Une première méthode serait d'écrire $h(t) = \cos(\sin t)e^{t^2}$ puis de dériver h...
Mais utilisons ici la formule $J_h(t) = J_G(F(t)) \times J_F(t)$, où l'on définit $F(t) = (t^2, \sin t)$, de sorte que $h = G \circ F$.
Sachant que :

$$J_h(t) = h'(t) \qquad J_G(x,y) = \begin{pmatrix} \cos(y)e^x & -\sin(y)e^x \end{pmatrix} \qquad J_F(t) = \begin{pmatrix} 2t \\ \cos t \end{pmatrix},$$

on calcule $J_G(F(t))$ et on obtient

$$h'(t) = 2t\big(\cos(\sin t)e^{t^2}\big) + \cos(t)\big(-\sin(\sin t)e^{t^2}\big) = \big(2t\cos(\sin t) - \cos(t)\sin(\sin t)\big)e^{t^2}.$$

Exemple 11.
Soit $G : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable. Soit $h : \mathbb{R} \to \mathbb{R}$ telle que $h(t) = G(2t, 1+t^2)$. Exprimer la dérivée de h en fonction des dérivées partielles de G.

Solution.
On pose $F : \mathbb{R} \to \mathbb{R}^2$ définie par $F(t) = (2t, 1+t^2)$, de sorte que $h = G \circ F$. Nous avons donc

$$J_h(t) = h'(t) \qquad J_G(x,y) = \begin{pmatrix} \frac{\partial G}{\partial x}(x,y) & \frac{\partial G}{\partial y}(x,y) \end{pmatrix} \qquad J_F(t) = \begin{pmatrix} 2 \\ 2t \end{pmatrix}.$$

Ainsi :

$$h'(t) = J_h(t) = J_G(F(t)) \times J_F(t) = 2\frac{\partial G}{\partial x}(2t, 1+t^2) + 2t\frac{\partial G}{\partial y}(2t, 1+t^2).$$

Cas $F : \mathbb{R} \to \mathbb{R}^n$, $G : \mathbb{R}^n \to \mathbb{R}$.

Plus généralement, on a le résultat suivant.

Proposition 4.
Soit $F : \mathbb{R} \to \mathbb{R}^n$ une fonction dont chacune des composantes est dérivable, et soit $G : \mathbb{R}^n \to \mathbb{R}$ différentiable. Alors $h : \mathbb{R} \to \mathbb{R}$ définie par $h(t) = G\big(F(t)\big)$ est dérivable et :

$$h'(t) = \big\langle \operatorname{grad} G\big(F(t)\big) \mid F'(t) \big\rangle.$$

Cas $F : \mathbb{R}^2 \to \mathbb{R}^2$, $G : \mathbb{R}^2 \to \mathbb{R}$

Proposition 5.
Soient $F : \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto \big(f_1(x,y), f_2(x,y)\big)$, $G : \mathbb{R}^2 \to \mathbb{R}$, $(u,v) \mapsto G(u,v)$ des fonctions différentiables. La fonction $H = G \circ F : \mathbb{R}^2 \to \mathbb{R}$, $(x,y) \mapsto G\big(F(x,y)\big)$ est différentiable et :

$$\left\{ \begin{array}{rcl} \frac{\partial H}{\partial x}(x,y) & = & \frac{\partial G}{\partial u}(F(x,y))\frac{\partial f_1}{\partial x}(x,y) + \frac{\partial G}{\partial v}(F(x,y))\frac{\partial f_2}{\partial x}(x,y) \\ \frac{\partial H}{\partial y}(x,y) & = & \frac{\partial G}{\partial u}(F(x,y))\frac{\partial f_1}{\partial y}(x,y) + \frac{\partial G}{\partial v}(F(x,y))\frac{\partial f_2}{\partial y}(x,y) \end{array} \right. .$$

C'est encore une fois la formule $J_H(x,y) = J_G(F(x,y)) \times J_F(x,y)$, avec :

$$J_H(x,y) = \begin{pmatrix} \frac{\partial H}{\partial x}(x,y) & \frac{\partial H}{\partial y}(x,y) \end{pmatrix} \qquad J_G(u,v) = \begin{pmatrix} \frac{\partial G}{\partial u}(u,v) & \frac{\partial G}{\partial v}(u,v) \end{pmatrix}$$

et

$$J_F(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(x,y) & \frac{\partial f_1}{\partial y}(x,y) \\ \frac{\partial f_2}{\partial x}(x,y) & \frac{\partial f_2}{\partial y}(x,y) \end{pmatrix}.$$

Exemple 12.
Calculer les dérivées partielles de la fonction $(x,y) \mapsto G(x-y, x+y)$ où $G : \mathbb{R}^2 \to \mathbb{R}$ est une fonction différentiable.

Solution.
On pose $F(x,y) = (x-y, x+y)$, on note (u,v) les variables de la fonction G et $H(x,y) = (G \circ F)(x,y) = G(x-y, x+y)$.
On a donc :

$$J_H(x,y) = \begin{pmatrix} \frac{\partial H}{\partial x}(x,y) & \frac{\partial H}{\partial y}(x,y) \end{pmatrix} \qquad J_G(u,v) = \begin{pmatrix} \frac{\partial G}{\partial u}(u,v) & \frac{\partial G}{\partial v}(u,v) \end{pmatrix} \qquad J_F(x,y) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Donc :

$$\begin{cases} \frac{\partial H}{\partial x}(x,y) & = & \frac{\partial G}{\partial u}(x-y, x+y) + \frac{\partial G}{\partial v}(x-y, x+y) \\ \frac{\partial H}{\partial y}(x,y) & = & -\frac{\partial G}{\partial u}(x-y, x+y) + \frac{\partial G}{\partial v}(x-y, x+y) \end{cases}$$

Un autre exemple.

Exemple 13.
Prenons $F : \mathbb{R}^2 \to \mathbb{R}^3$ et $G : \mathbb{R}^3 \to \mathbb{R}$ deux fonctions définies par

$$F(x,y) = (x + y^4, y - 3x^2, 2x^2 - 3y) \quad \text{et} \quad G(x,y,z) = 2xy - 3(x+z).$$

Calculer les dérivées partielles de la fonction $H = G \circ F$.

Solution.

- Tout d'abord, on note que H est une fonction de deux variables à valeurs réelles, c'est-à-dire $H : \mathbb{R}^2 \to \mathbb{R}$. Pour calculer $\frac{\partial H}{\partial x}$ et $\frac{\partial H}{\partial y}$, il suffit de calculer la matrice jacobienne de H.
- La formule de la matrice jacobienne d'une composée s'écrit :
$$J_H(x,y) = J_G(F(x,y)) \times J_F(x,y).$$
- On a
$$J_H(x,y) = \begin{pmatrix} \frac{\partial H}{\partial x}(x,y) & \frac{\partial H}{\partial y}(x,y) \end{pmatrix} \qquad J_G(x,y,z) = \begin{pmatrix} 2y-3 & 2x & -3 \end{pmatrix} \qquad J_F(x,y) = \begin{pmatrix} 1 & 4y^3 \\ -6x & 1 \\ 4x & -3 \end{pmatrix}.$$
- On en déduit que
$$J_G(F(x,y)) = \begin{pmatrix} 2(y-3x^2)-3 & 2(x+y^4) & -3 \end{pmatrix}.$$
- On obtient $\frac{\partial H}{\partial x}(x,y)$ comme la première composante de $J_H(x,y)$:
$$\frac{\partial H}{\partial x}(x,y) = 1 \cdot \big(2(y-3x^2)-3\big) - 6x \cdot \big(2(x+y^4)\big) + 4x \cdot \big(-3\big) = -12xy^4 - 18x^2 + 2y - 12x - 3$$
- À vous de faire le calcul de $\frac{\partial H}{\partial y}$!

Mini-exercices.

1. Calculer de deux façons différentes la dérivée de la fonction $t \mapsto G(\sin t, e^t)$, où $G(x,y) = \frac{x}{y}$. Même question avec $t \mapsto G(t+1, t^2, \frac{1}{t})$ et $G(x,y,z) = x^2 + \sqrt{yz}$.
2. Exprimer les dérivées partielles de $(x,y) \mapsto G(x^2 - y^3, \ln(x) - y)$ en fonction des dérivées partielles

de $G : \mathbb{R}^2 \to \mathbb{R}$. Même question avec $(x, y, z) \mapsto G(x + y^2, 2y - z, xz)$ et $G : \mathbb{R}^3 \to \mathbb{R}$.

3. Soit $G : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $G(x, y) = \left(\frac{x}{y}, \ln(x + y)\right)$. Calculer la matrice jacobienne de la fonction définie par $(x, y) \mapsto G(ax + by, cx + dy)$, où $a, b, c, d \in \mathbb{R}$ sont des constantes.

Auteurs du chapitre

Arnaud Bodin. D'après des cours de Abdellah Hanani (Lille), Goulwen Fichou et Stéphane Leborgne (Rennes), Laurent Pujo-Menjouet (Lyon). Relu par Anne Bauval, Vianney Combet et Barbara Tumpach.

Chapitre 4

Gradient – Théorème des accroissements finis

Le calcul différentiel s'applique au calcul des équations des tangentes aux courbes et des plans tangents aux surfaces. Il permet aussi d'approcher les fonctions de plusieurs variables par des formules linéaires.

1. Gradient

Le gradient est un vecteur dont les coordonnées sont les dérivées partielles. Il est très important en physique et a des nombreuses applications géométriques, car il indique la direction perpendiculaire aux courbes et surfaces.

1.1. Rappel de la définition

Définition 1.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction admettant des dérivées partielles. Le *gradient* en $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, noté $\operatorname{grad} f(x)$, est le vecteur

$$\operatorname{grad} f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}.$$

Les physiciens notent souvent $\nabla f(x)$ pour $\operatorname{grad} f(x)$. Le symbole ∇ se lit « nabla ».

Pour une fonction $f(x, y)$ de deux variables, au point (x_0, y_0), on a donc

$$\operatorname{grad} f(x_0, y_0) = \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) \\ \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix}.$$

Exemple 1.

- Si $f(x, y) = x^2 y^3$ alors $\operatorname{grad} f(x, y) = \begin{pmatrix} 2xy^3 \\ 3x^2y^2 \end{pmatrix}$. Au point $(x_0, y_0) = (2, 1)$, on a $\operatorname{grad} f(2, 1) = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$.
- Si $f(x, y, z) = x^2 \sin(yz)$ alors $\operatorname{grad} f(x, y, z) = \begin{pmatrix} 2x \sin(yz) \\ x^2 z \cos(yz) \\ x^2 y \cos(yz) \end{pmatrix}$.
- Si $f(x_1, \ldots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$ alors $\operatorname{grad} f(x_1, \ldots, x_n) = \begin{pmatrix} 2x_1 \\ \vdots \\ 2x_n \end{pmatrix}$.

Remarque.
Le gradient est un élément de $\mathbb{R}^n$ écrit comme un vecteur colonne. C'est la transposée de la matrice jacobienne qui est ici un vecteur ligne. Parfois, pour alléger l'écriture, on peut aussi écrire le gradient sous la forme d'un vecteur ligne.

1.2. Rappel du lien avec la différentielle

Lien avec la différentielle.
Le gradient est une autre écriture possible de la différentielle. Si f est différentiable en $x \in \mathbb{R}^n$ et si $h \in \mathbb{R}^n$, alors :

$$\boxed{\mathrm{d}f(x)(h) = \langle \operatorname{grad} f(x) \mid h \rangle}$$

où $\langle \cdot \mid \cdot \rangle$ désigne le produit scalaire. Pour $f : \mathbb{R}^2 \to \mathbb{R}$ et $(x_0, y_0), (h, k) \in \mathbb{R}^2$, cela s'écrit :

$$\mathrm{d}f(x_0, y_0)(h, k) = \langle \operatorname{grad} f(x_0, y_0) \mid \left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right) \rangle = h \frac{\partial f}{\partial x}(x_0, y_0) + k \frac{\partial f}{\partial y}(x_0, y_0).$$

La différentielle $\mathrm{d}f(x)$ est une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}$ et $\operatorname{grad} f(x)$ est la transposée de sa matrice dans la base canonique. On pourrait donc aussi écrire le produit de matrices $\mathrm{d}f(x)(h) = \operatorname{grad} f(x)^T \cdot h$.

Lien avec la dérivée directionnelle.
Si f est différentiable en $x \in \mathbb{R}^n$ et si $v \in \mathbb{R}^n$, alors :

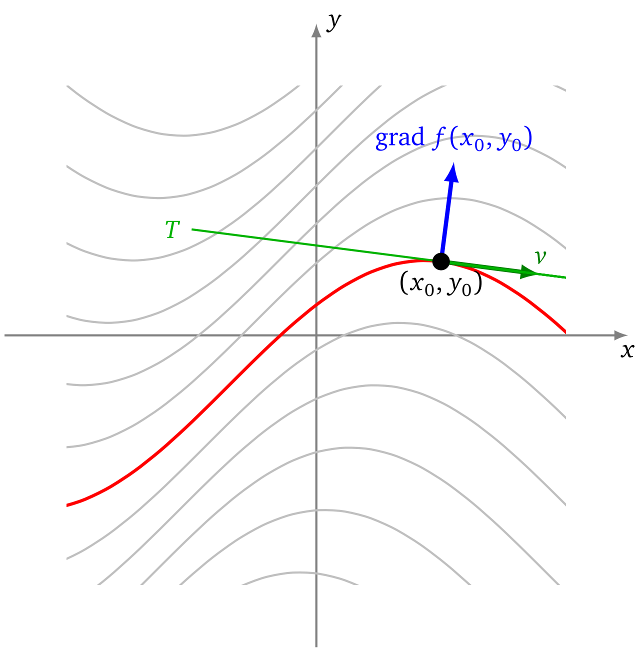
$$D_v f(x) = \mathrm{d}f(x)(v) = \langle \operatorname{grad} f(x) \mid v \rangle.$$

1.3. Tangentes aux lignes de niveau

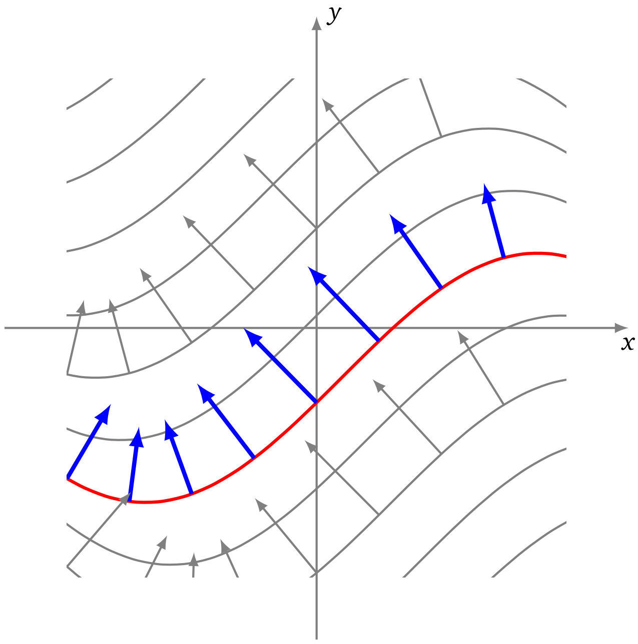
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction différentiable et soit $k \in \mathbb{R}$. On considère les lignes de niveau $f(x, y) = k$, c'est-à-dire l'ensemble des $(x, y) \in \mathbb{R}^2$ qui vérifient l'équation $f(x, y) = k$.

Proposition 1.
Le vecteur gradient $\operatorname{grad} f(x_0, y_0)$ *est orthogonal à la ligne de niveau de* f *passant au point* (x_0, y_0).

Sur ce premier dessin, vous avez (en rouge) la ligne de niveau passant par le point (x_0, y_0). En ce point est dessiné (en vert) un vecteur tangent v et la tangente à la ligne de niveau. Le vecteur gradient (en bleu) est orthogonal à la ligne de niveau en ce point.



En chaque point du plan part un vecteur gradient. Ce vecteur gradient est orthogonal à la ligne de niveau passant par ce point.



Précisons la notion de tangente :

- On se place en un point (x_0, y_0) où les deux dérivées partielles ne s'annulent pas en même temps, c'est-à-dire $\operatorname{grad} f(x_0, y_0)$ n'est pas le vecteur nul. Considérons $\mathscr{C}$, la ligne de niveau de f qui passe par ce point (x_0, y_0). Le théorème des fonctions implicites (qui sera vu plus tard) montre qu'il est possible de trouver une paramétrisation de $\mathscr{C}$ au voisinage de (x_0, y_0). Notons $\gamma : [-1,1] \to \mathbb{R}^2$, $t \mapsto \gamma(t) = (\gamma_1(t), \gamma_2(t))$ cette paramétrisation, en supposant que $\gamma(0) = (x_0, y_0)$.
- La ***tangente*** à la courbe $\mathscr{C}$ en (x_0, y_0) est la droite passant par le point (x_0, y_0) et de vecteur directeur le vecteur dérivé $\gamma'(0) = (\gamma_1'(0), \gamma_2'(0))$.
- Un vecteur v est ***orthogonal*** (ou ***normal*** si v n'est pas nul) à la courbe $\mathscr{C}$ en (x_0, y_0) s'il est orthogonal à la tangente en ce point, c'est-à-dire si $\langle v \mid \gamma'(0) \rangle = 0$.

On peut maintenant prouver la proposition.

Démonstration. Notons $k = f(x_0, y_0)$. Alors $\mathscr{C}$ est la ligne de niveau $f(x, y) = k$. Dire que $\gamma(t)$ est une paramétrisation de $\mathscr{C}$ (autour de (x_0, y_0)), c'est exactement dire que :

$$\forall t \in [-1,1] \qquad f\big(\gamma(t)\big) = k$$

Comme $f \circ \gamma$ est une fonction constante, alors sa dérivée est nulle. La formule de la différentielle d'une composition s'écrit

$$J_f(\gamma(t)) \times J_\gamma(t) = 0$$

et donc ici

$$\begin{pmatrix} \frac{\partial f}{\partial x}(\gamma(t)) & \frac{\partial f}{\partial y}(\gamma(t)) \end{pmatrix} \times \begin{pmatrix} \gamma_1'(t) \\ \gamma_2'(t) \end{pmatrix} = 0.$$

En $t = 0$, on trouve exactement

$$\langle \operatorname{grad} f(x_0, y_0) \mid \gamma'(0) \rangle = 0,$$

ce qui signifie que le gradient est orthogonal au vecteur tangent. □

Dans la pratique, c'est l'équation de la tangente qui nous intéresse :

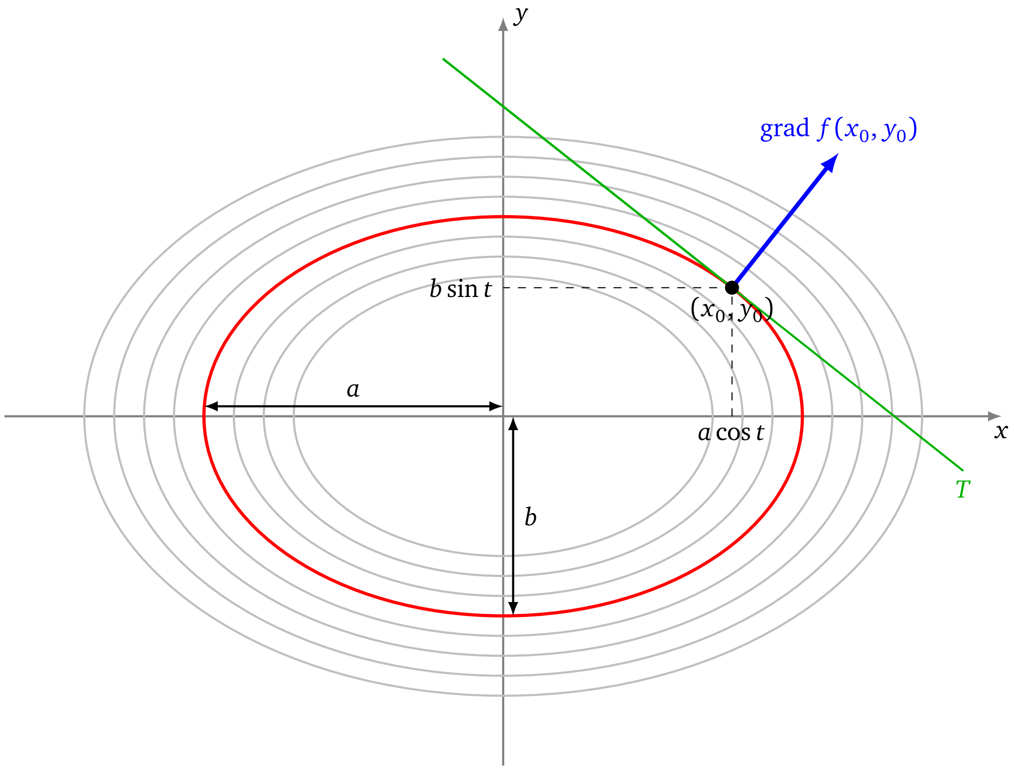
Proposition 2.
L'équation de la tangente à la ligne de niveau de f en (x_0, y_0) est

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

pourvu que le gradient de f en ce point ne soit pas le vecteur nul.

Démonstration. C'est l'équation de la droite dont un vecteur normal est $\operatorname{grad} f(x_0, y_0)$ et qui passe par (x_0, y_0). □

Exemple 2 (Tangente à une ellipse)**.**
Trouvons les tangentes à l'ellipse $\mathscr{E}$ d'équation $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ (avec $a, b > 0$).



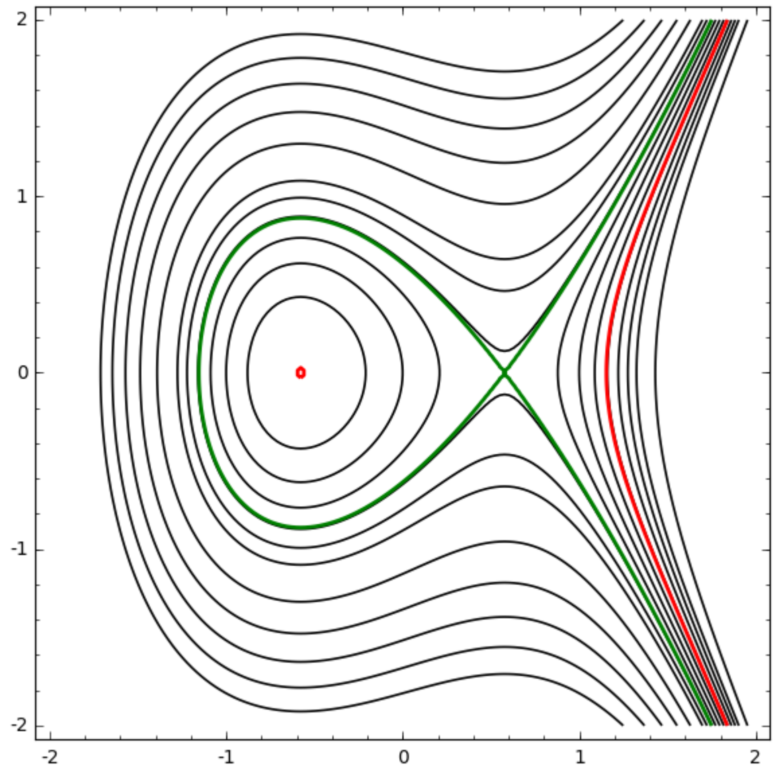
- **Par les lignes de niveau.**
Cette ellipse $\mathscr{E}$ est la ligne de niveau $f(x, y) = 1$ de la fonction $f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2}$. Les dérivées partielles de f sont :
$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{2x_0}{a^2} \qquad \frac{\partial f}{\partial y}(x_0, y_0) = \frac{2y_0}{b^2}$$
Donc l'équation de la tangente à l'ellipse $\mathscr{E}$ en un de ces points (x_0, y_0) est
$$\frac{2x_0}{a^2}(x - x_0) + \frac{2y_0}{b^2}(y - y_0) = 0.$$
Mais comme $\frac{x_0^2}{a^2} + \frac{y_0^2}{b^2} = 1$ alors l'équation de la tangente se simplifie en $\dfrac{x_0}{a^2}x + \dfrac{y_0}{b^2}y = 1$.
- **Par une paramétrisation.**
Une autre approche est de paramétrer l'ellipse $\mathscr{E}$ par $\gamma(t) = (a \cos t, b \sin t)$, $t \in [0, 2\pi[$. Le vecteur dérivé étant $\gamma'(t) = (-a \sin t, b \cos t)$, la tangente en $\gamma(t)$ est dirigée par le vecteur $(-b \cos t, -a \sin t)$. L'équation de la tangente en $\gamma(t)$ est donc
$$-b \cos t(x - a \cos t) - a \sin t(y - b \sin t) = 0.$$
En posant $x_0 = a \cos t$ et $y_0 = b \sin t$, on retrouve l'équation ci-dessus.

Exemple 3.
Soit $f(x,y) = x^3 - y^2 - x$. Nous allons calculer l'équation des tangentes aux courbes de niveau de f.

- **Calcul du gradient.** $\operatorname{grad} f(x,y) = \begin{pmatrix} 3x^2-1 \\ -2y \end{pmatrix}$.
- **Points où le gradient s'annule.** $P_1 = \left(-\frac{\sqrt{3}}{3}, 0\right)$, $P_2 = \left(+\frac{\sqrt{3}}{3}, 0\right)$.
  On calcule $f(P_1) = +\frac{2\sqrt{3}}{9}$, $f(P_2) = -\frac{2\sqrt{3}}{9}$. Ainsi P_1 est sur la ligne de niveau $f(x,y) = +\frac{2\sqrt{3}}{9}$ et P_2 sur celle $f(x,y) = -\frac{2\sqrt{3}}{9}$.
- **Équation de la tangente.** En dehors de ces deux points, les courbes de niveau ont une tangente. Au point (x_0, y_0), l'équation est

$$(3x_0^2 - 1)(x - x_0) - 2y_0(y - x_0) = 0.$$

Voici quelques lignes de niveau de f. Le point P_1 est le point isolé du niveau rouge et il n'y a pas de tangente en ce point. Le point P_2 est le point double du niveau vert et il n'y a pas de tangente en ce point (en fait on pourrait dire qu'il y a deux tangentes).

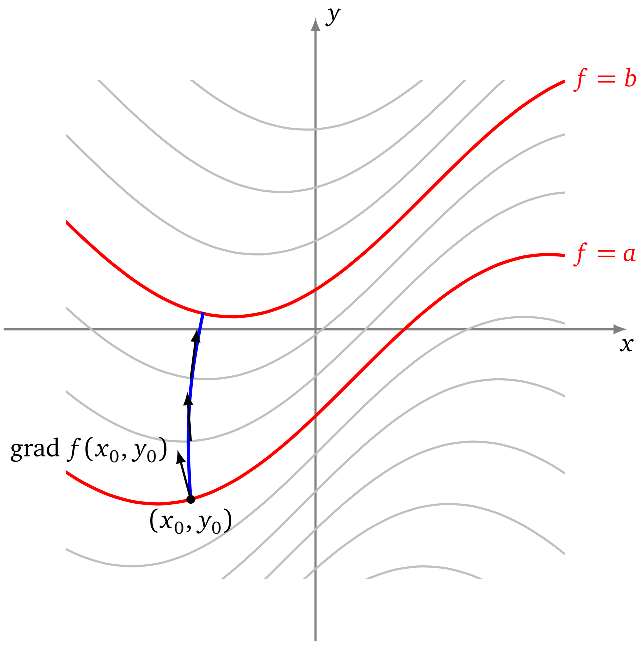


1.4. Lignes de plus forte pente

Considérons les lignes de niveau $f(x,y) = k$ d'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$. On se place en un point (x_0, y_0). On cherche dans quelle direction se déplacer pour augmenter le plus vite la valeur de f.

> **Proposition 3.**
> *Le vecteur gradient* $\operatorname{grad} f(x_0, y_0)$ *indique la direction de plus forte pente à partir du point* (x_0, y_0).

Autrement dit, si l'on veut passer le plus vite possible du niveau a à un niveau $b > a$, à partir d'un point donné (x_0, y_0) de niveau $f(x_0, y_0) = a$, alors il faut démarrer en suivant la direction du gradient $\operatorname{grad} f(x_0, y_0)$.



Comme illustration, un skieur voulant aller vite choisit la plus forte pente descendante en un point de la montagne : c'est la direction opposée au gradient.

Démonstration. La dérivée suivant le vecteur non nul v au point (x_0, y_0) décrit la variation de f autour de ce point lorsqu'on se déplace dans la direction v. La direction selon laquelle la croissance est la plus forte est celle du gradient de f. En effet, on a

$$D_v f(x_0, y_0) = \langle \operatorname{grad} f(x_0, y_0) \mid v \rangle = \| \operatorname{grad} f(x_0, y_0) \| \cdot \|v\| \cdot \cos \theta$$

où θ est l'angle entre le vecteur $\operatorname{grad} f(x_0, y_0)$ et le vecteur v. Le maximum est atteint lorsque l'angle θ est nul, c'est-à-dire lorsque v pointe dans la même direction que $\operatorname{grad} f(x_0, y_0)$. □

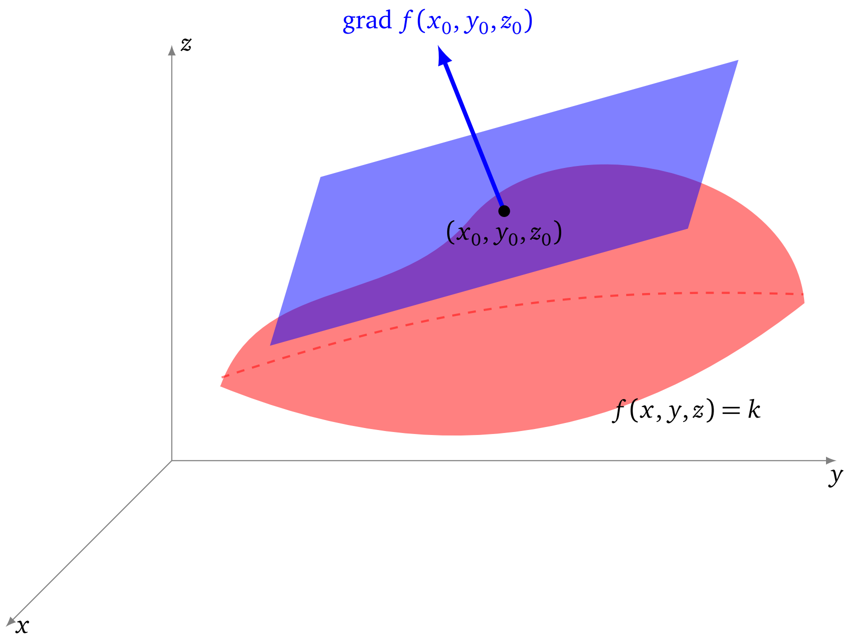
1.5. Surfaces de niveau

On a des résultats similaires pour les surfaces de niveau $f(x, y, z) = k$ d'une fonction f différentiable.
Rappelons que le plan de $\mathbb{R}^3$ passant par (x_0, y_0, z_0) et de vecteur normal $n = (a, b, c)$ a pour équation cartésienne : $a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$.
De même qu'il existe une droite tangente à une ligne de niveau, il existe un ***plan tangent*** à une surface de niveau.

Proposition 4.
Le vecteur gradient $\operatorname{grad} f(x_0, y_0, z_0)$ *est orthogonal à la surface de niveau de* f *passant au point* (x_0, y_0, z_0).
Autrement dit, l'équation du plan tangent à la surface de niveau de f *en* (x_0, y_0, z_0) *est*

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

pourvu que le gradient de f *en ce point ne soit pas le vecteur nul.*



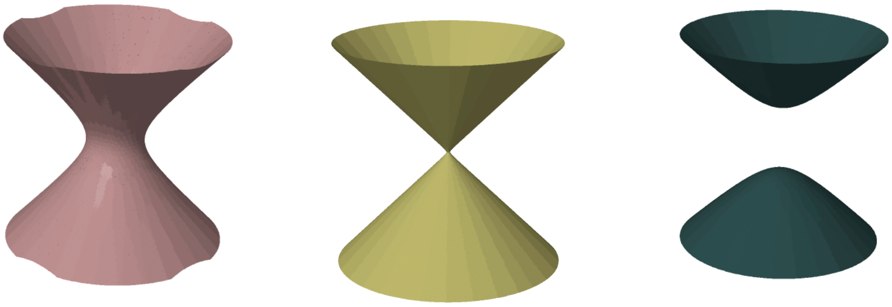
Plus généralement, pour $f : \mathbb{R}^n \to \mathbb{R}$, grad $f(x_0)$ est orthogonal à l'espace tangent à l'hypersurface de niveau $f(x) = k$ passant par le point $x_0 \in \mathbb{R}^n$.

Exemple 4.
Pour quelles valeurs de k la surface de niveau $x^2 + y^2 - z^2 = k$ admet-elle un plan tangent horizontal (c'est-à-dire parallèle au plan $(z = 0)$) ?

Solution. On pose $f(x, y, z) = x^2 + y^2 - z^2$.

- **Calcul du gradient.** $\operatorname{grad} f(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ -2z \end{pmatrix}$.
- **Gradient nul.** Le gradient est le vecteur nul uniquement au point $(0, 0, 0)$, donc au niveau $k = 0$. En ce point, il n'y a pas de plan tangent.
- **Plan tangent horizontal.** Le plan tangent est horizontal exactement lorsque le gradient est un vecteur colinéaire à $\left(\begin{smallmatrix} 0 \\ 0 \\ 1 \end{smallmatrix}\right)$ (et n'est pas le vecteur nul). Il faut donc $\frac{\partial f}{\partial x} = 0$ et $\frac{\partial f}{\partial y} = 0$, ce qui implique ici $x = 0$ et $y = 0$.
- **Analyse.** En un tel point $(0, 0, z)$, on a $f(x, y, z) = -z^2$, donc le niveau k est strictement négatif.
- **Synthèse.** Réciproquement, étant donné $k < 0$, alors aux points $(0, 0, \pm\sqrt{|k|})$, le vecteur gradient est vertical, donc le plan tangent est horizontal.
- **Conclusion.**
  — Pour $k > 0$, il n'y a pas de plan tangent horizontal. La surface $f(x, y, z) = k$ est un *hyperboloïde à une nappe*.
  — Pour $k = 0$, il n'y a pas de plan tangent horizontal. Le point $(0, 0, 0)$ est singulier. La surface $f(x, y, z) = 0$ est un *cône*.
  — Pour $k < 0$, il y a deux points ayant un plan tangent horizontal. La surface $f(x, y, z) = k$ est un *hyperboloïde à deux nappes*.

De gauche à droite : l'hyperboloïde à une nappe, le cône, l'hyperboloïde à deux nappes.



Les trois surfaces ensemble, comme des surfaces de niveau de f (avec une découpe pour voir l'intérieur).

1.6. Plan tangent au graphe d'une fonction

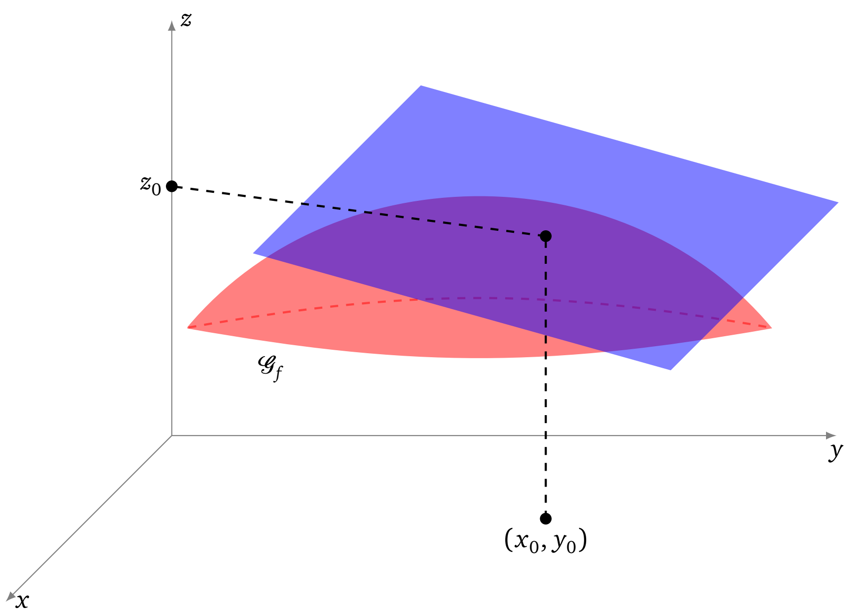
Soit $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ différentiable. On s'intéresse maintenant au graphe de f. Rappelons que c'est la surface

$$\mathcal{G}_f = \left\{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in U \text{ et } z = f(x, y)\right\}.$$

Attention ! Il ne faut pas confondre le graphe d'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ avec les surfaces de niveau de fonctions $f : \mathbb{R}^3 \to \mathbb{R}$.

Proposition 5.
Soit $f : U \subset \mathbb{R}^2 \to \mathbb{R}$ différentiable. Soit $(x_0, y_0) \in U$ et soit $M_0 = (x_0, y_0, f(x_0, y_0))$ un point du graphe $\mathcal{G}_f$ de f. Le plan tangent au graphe $\mathcal{G}_f$ en M_0 a pour équation :

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



Démonstration. On introduit la fonction F définie par $F(x,y,z) = z - f(x,y)$, pour tout $(x,y,z) \in U \times \mathbb{R}$. Le graphe de f est la surface $\mathscr{G}_f = \{(x,y,z) \in U \times \mathbb{R} \mid F(x,y,z) = 0\}$. Ainsi, $\mathscr{G}_f$ est aussi une surface de niveau de F. On calcule :

$$\operatorname{grad} F(x,y,z) = \left(-\frac{\partial f}{\partial x}(x,y), -\frac{\partial f}{\partial y}(x,y), 1\right)$$

Notez que ce vecteur n'est jamais nul et donc une équation du plan tangent en (x_0, y_0, z_0) est :

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

où on a noté $z_0 = f(x_0, y_0)$. □

Exemple 5.
Soit $f(x,y) = 3x^2 - 2y^3$.

1. Trouver l'équation du plan tangent au graphe de f au-dessus de (x_0, y_0).
   On a $\frac{\partial f}{\partial x}(x,y) = 6x$, $\frac{\partial f}{\partial y}(x,y) = -6y^2$. Donc l'équation du plan tangent au point $(x_0, y_0, f(x_0, y_0))$ est :
   $$z = 3x_0^2 - 2y_0^3 + 6x_0(x - x_0) - 6y_0^2(y - y_0)$$
   ou encore
   $$6x_0 x - 6y_0^2 y - z = 3x_0^2 - 4y_0^3.$$
2. *Application.* Trouver l'équation du plan $\mathscr{P}_0$ tangent au-dessus du point $(x_0, y_0) = (1, 2)$.
   On pose $z_0 = f(x_0, y_0) = -13$. L'équation est alors $z = -13 + 6(x - 1) - 24(y - 2)$, autrement dit $6x - 24y - z = -29$.
3. Trouver les points pour lesquels le plan tangent est parallèle à $\mathscr{P}_0$.
   On cherche un point (x_1, y_1) qui vérifie $(6, -24, -1) = (6x_1, -6y_1^2, -1)$ et $(x_1, y_1) \neq (x_0, y_0)$. On trouve comme seul autre point $(x_1, y_1) = (1, -2)$.

Mini-exercices.

1. Calculer le gradient en tout point de la fonction définie par $f(x,y) = xe^y$. Même question pour $f(x,y,z) = x^2 y^3 z^4$, puis $f(x_1, \ldots, x_n) = \sqrt{x_1^2 + \cdots + x_n^2}$.
2. Calculer le gradient de $f(x,y) = \ln(x + y^2)$ en tout point (x_0, y_0). Exprimer la différentielle $\mathrm{d}f(x_0, y_0)(h,k)$. Calculer la dérivée directionnelle de f en $(x_0, y_0) = (1,2)$ le long du vecteur $(2,3)$.
3. Soient $f, g : \mathbb{R}^n \to \mathbb{R}$. Montrer que $\operatorname{grad}(f + g)(x) = \operatorname{grad} f(x) + \operatorname{grad} g(x)$. Que vaut $\operatorname{grad}(\lambda \cdot f)(x)$ (où $\lambda \in \mathbb{R}$) ?

4. Trouver $f:\mathbb{R}^n\to\mathbb{R}$ et $x,y\in\mathbb{R}^n$ tels que $\operatorname{grad} f(x+y)$ n'est pas égal à $\operatorname{grad} f(x)+\operatorname{grad} f(y)$. Trouver $f:\mathbb{R}^n\to\mathbb{R}$, $x\in\mathbb{R}^n$ et $\lambda\in\mathbb{R}$ tels que $\operatorname{grad} f(\lambda\cdot x)$ n'est pas égal à $\lambda\cdot\operatorname{grad} f(x)$.
5. Soit l'hyperbole $\mathscr{H}$ d'équation $x^2-y^2=1$. Dessiner $\mathscr{H}$. Calculer l'équation cartésienne de la tangente en un point $(x_0,y_0)\in\mathscr{H}$. Trouver les points de $\mathscr{H}$ où la tangente est colinéaire au vecteur $(2,1)$.
6. Trouver les points de la surface d'équation $x^2-y^2z=0$ où le gradient s'annule. Calculer l'équation du plan tangent en dehors de ces points.
7. Soit $f(x,y)=\ln(x+y^2)$. Trouver l'équation du plan tangent au graphe de f au-dessus du point $(-2,3)$.

2. Calcul d'incertitudes

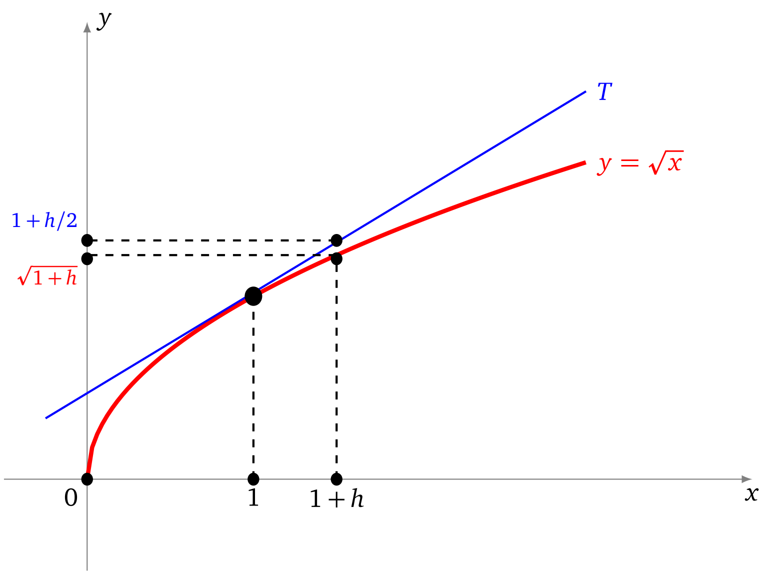
Pour les fonctions d'une variable, la dérivée permet de calculer un développement limité à l'ordre 1 et donc d'approcher une fonction autour d'un point par une formule linéaire. Pour les fonctions de plusieurs variables, nous avons besoin des dérivées partielles pour obtenir une formule linéaire.

2.1. Calcul approché

Motivation : comment estimer la valeur $\sqrt{1,01}$ sans calculatrice ? On pose $f(x)=\sqrt{x}$. Le développement limité en 1 s'écrit $f(1+h)\simeq f(1)+hf'(1)$. Dans ce cas, on obtient :

$$\sqrt{1+h}\simeq 1+\frac{h}{2}.$$

Cela donne l'estimation $\sqrt{1,01}\simeq 1,005$ (au lieu de $\sqrt{1,01}=1,00498\ldots$).
Géométriquement, la tangente au graphe de f en 1 donne une bonne approximation des valeurs de f autour de ce point.



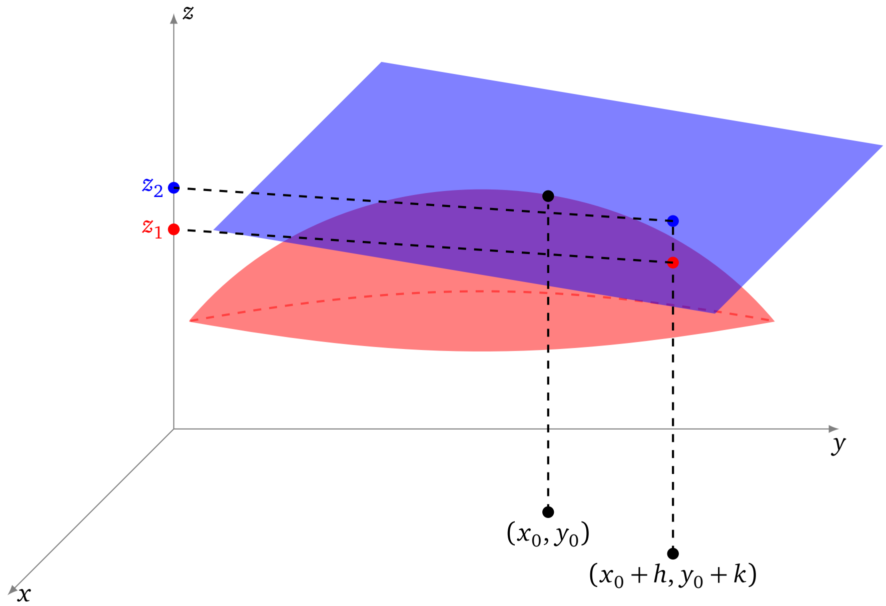
Nous allons voir l'analogue pour les fonctions de deux variables.
Si f est différentiable au point $A=(x_0,y_0)$, alors

$$f(x_0+h,y_0+k)=f(x_0,y_0)+\mathrm{d}f(x_0,y_0)(h,k)+o(\sqrt{h^2+k^2}).$$

On propose alors comme approximation de $f(x_0+h,y_0+k)$ la quantité $f(x_0,y_0)+\mathrm{d}f(x_0,y_0)(h,k)$, appelée ***approximation linéaire***, c'est-à-dire :

$$\boxed{f(x_0+h,y_0+k) \simeq f(x_0,y_0)+h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0)}$$

C'est une approximation qui est valable pour h et k petits.
L'interprétation géométrique est la suivante : on approche le graphe de f en (x_0, y_0) par le plan tangent au graphe en ce point. Sur la figure ci-dessous sont représentés : le graphe de f (en rouge), le plan tangent au-dessus du point (x_0, y_0) (en bleu). La valeur $z_1 = f(x_0+h, y_0+k)$ est la valeur exacte donnée par le point de la surface au-dessus de (x_0+h, y_0+k). On approche cette valeur par la valeur $z_2 = f(x_0,y_0)+h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0)$ donnée par le point du plan tangent au-dessus de (x_0+h, y_0+k).



Exemple 6.
Valeur approchée de $f(1,002; 0,997)$ si $f(x,y) = x^2 y$.

Solution. Ici, on a $(x_0, y_0) = (1,1)$, $h = 2\times 10^{-3}$, $k = -3\times 10^{-3}$, $\frac{\partial f}{\partial x} = 2xy$, $\frac{\partial f}{\partial y} = x^2$, donc $\frac{\partial f}{\partial x}(x_0,y_0) = 2$, $\frac{\partial f}{\partial x}(x_0,y_0) = 1$. Ainsi, on a

$$f(1+h,1+k) \simeq f(1,1)+2h+k$$

et donc

$$f(1,002;0,997) \simeq 1+2\times 2\times 10^{-3}-3\times 10^{-3} = 1,001.$$

Avec une calculatrice, on trouve $f(1,002;0,997) = 1,000991988$: l'approximation est bonne.

Exemple 7.
Deux résistances R_1 et R_2 sont connectées en parallèle. La résistance totale R du circuit est donnée par la formule

$$\frac{1}{R} = \frac{1}{R_1}+\frac{1}{R_2}.$$

La résistance R_1 vaut environ 1 ; R_2 vaut environ 2 (en kilo-ohms). Écrire l'approximation linéaire correspondante, puis donner une valeur approchée de R lorsque $R_1 = 1,01$ et $R_2 = 1,98$.

Solution. Notons $f(x,y) = \frac{1}{\frac{1}{x}+\frac{1}{y}} = \frac{xy}{x+y}$, de sorte que $R = f(R_1, R_2)$. Par exemple, si $R_1 = 1$ et $R_2 = 2$, on trouve $R = 0,6666\ldots$
On calcule

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

Posons $(x_0, y_0) = (1, 2)$. On a $f(x_0, y_0) = \frac{2}{3}$, $\frac{\partial f}{\partial x}(x_0, y_0) = \frac{4}{9}$, $\frac{\partial f}{\partial y}(x_0, y_0) = \frac{1}{9}$.
L'approximation linéaire de f au voisinage de (x_0, y_0) s'écrit

$$f(1+h, 2+k) \simeq \frac{2}{3} + \frac{4}{9}h + \frac{1}{9}k.$$

Avec $h = 0,01$ et $k = -0,02$, on obtient $f(1,01; 1,98) \simeq 0,6689$.

2.2. Calcul d'incertitudes

Soit $f(x, y)$ une grandeur qui dépend de deux mesures x et y. La valeur de x est proche d'une valeur fixe x_0, mais n'est connue qu'à une incertitude près, c'est-à-dire $x \in [x_0 - \Delta x, x_0 + \Delta x]$ où Δx est un nombre réel positif, appelé l'***incertitude*** sur x. De même, $y \in [y_0 - \Delta y, y_0 + \Delta y]$.
Quelle va être l'erreur commise en approchant $f(x, y)$ par $f(x_0, y_0)$?
On note $(x, y) = (x_0 + h, y_0 + k)$. On a alors

$$\left|f(x_0+h, y_0+k) - f(x_0, y_0)\right| \simeq \left|h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0)\right| \leqslant \Delta x\left|\frac{\partial f}{\partial x}(x_0, y_0)\right| + \Delta y\left|\frac{\partial f}{\partial y}(x_0, y_0)\right|.$$

On obtient ainsi une majoration de l'***incertitude estimée*** δf :

$$\delta f \leqslant \Delta x\left|\frac{\partial f}{\partial x}(x_0, y_0)\right| + \Delta y\left|\frac{\partial f}{\partial y}(x_0, y_0)\right|$$

Exemple 8.
Une usine produit des cylindres de rayon $r = 2 \pm 0,1$ et de hauteur $h = 10 \pm 0,2$. Quelle est l'incertitude estimée sur le volume du cylindre ?

Solution. Le volume est donné par la formule $V(r, h) = \pi r^2 h$. On pose $r_0 = 2$, $\Delta r = 0,1$, $h_0 = 10$, $\Delta h = 0,2$. Ainsi l'incertitude estimée vérifie

$$\delta V \leqslant \Delta r\left|\frac{\partial V}{\partial r}(r_0, h_0)\right| + \Delta h\left|\frac{\partial V}{\partial h}(r_0, h_0)\right|.$$

On a $\frac{\partial V}{\partial r} = 2\pi r h$ et $\frac{\partial V}{\partial h} = \pi r^2$, donc ici : $\delta V \leqslant 0,1 \times 40\pi + 0,2 \times 4\pi = 4,8\pi \simeq 15$. Le volume sans erreur est $V_0 = V(r_0, h_0) = 40\pi \simeq 126$. Au vu de notre estimation de l'incertitude, on écrit

$$V(r, h) = 126 \pm 15.$$

Remarque.
Sur cet exemple, r est ici connu avec une incertitude relative $\frac{\Delta r}{r_0} = \frac{0,1}{2} = 5\,\%$, et pour h l'incertitude relative est $\frac{\Delta h}{h_0} = \frac{0,2}{10} = 2\,\%$. L'estimation de l'incertitude relative du volume est $\frac{\delta V}{V_0} \simeq \frac{15}{126} \simeq 12\,\%$. L'erreur relative du volume est donc bien plus élevée que les erreurs relatives sur r et h.

Remarque.
La formule de l'incertitude est seulement une estimation. Pour une majoration exacte de l'erreur, il faut utiliser le théorème ou l'inégalité des accroissements finis.

Mini-exercices.

1. Donner l'approximation linéaire de $y^2 e^x$ en $x_0 = 1$ et $y_0 = 2$. Sans calculatrice, en déduire une valeur approchée de $(1,99)^2 e^{1,03}$. Faire le même travail pour $\frac{\ln(9,99 \times 2,02)}{2,02}$. On admettra que $e \simeq 2,718$ et $\ln(20) \simeq 2,996$.
2. La tension U, la résistance R et l'intensité I sont reliées par la loi d'Ohm $U = RI$. Écrire l'approximation linéaire de la résistance, pour $U_0 = 120$ et $I_0 = 1$. Estimer la résistance lorsque $U = 118$ et $I = 0,9$. Estimer l'erreur commise lorsque l'on approche la résistance R par $R_0 = U_0/I_0$ pour $U \in [118, 122]$ et $I \in [0,9; 1,1]$.
3. Soit $f(x, y) = \sqrt{x^2 + y^2}$. Écrire l'approximation linéaire pour $f(x, y)$ autour de $(x_0, y_0) = (4, 3)$. Pour x connu avec une incertitude relative de $5\,\%$ autour de x_0, et y connu avec une incertitude relative

de 10 % autour de y_0, estimer l'erreur relative commise lorsque l'on approche $f(x,y)$ par $f(x_0,y_0)$.

3. Théorème des accroissements finis

Le théorème des accroissements finis est une façon exacte de mesurer l'écart entre deux valeurs de f. On peut aussi en tirer des inégalités. Cette section est beaucoup plus théorique que le reste du chapitre et peut être passée lors d'une première lecture.

3.1. Théorème des accroissements finis en une variable (rappel)

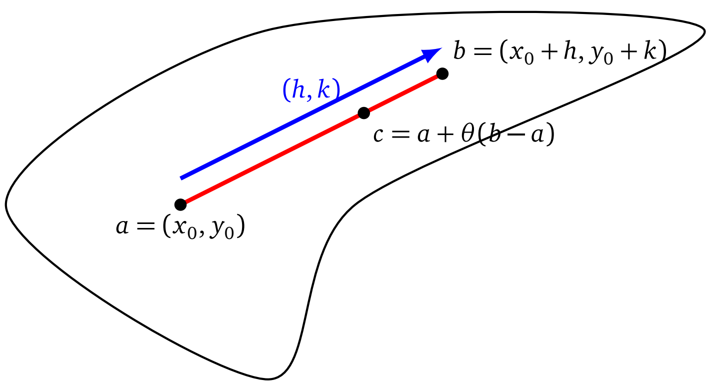
Théorème 1 (Théorème des accroissements finis en une variable)**.**
Soit $f : [a,b] \to \mathbb{R}$ continue sur $[a,b]$, dérivable sur $]a,b[$, où $a<b$. Il existe $c \in]a,b[$ tel que
$$f(b)-f(a)=f'(c)(b-a).$$

On renvoie au cours de première année pour la preuve.

3.2. Théorème des accroissements finis en deux variables

Théorème 2 (Théorème des accroissements finis en deux variables)**.**
Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathscr{C}^1$ sur un ouvert $U \subset \mathbb{R}^2$. Soient a,b deux points de U. Si le segment $[a,b]$ est inclus dans U, alors il existe $c \in]a,b[$ tel que
$$f(b)-f(a)=\langle \operatorname{grad} f(c) \mid b-a \rangle .$$

Énoncé ainsi, le théorème est valable aussi pour $U \subset \mathbb{R}^n$; en plus, en se souvenant que la différentielle s'exprime à l'aide du gradient $\mathrm{d}f(c)(h) = \langle \operatorname{grad} f(c) \mid h \rangle$, on obtient une reformulation similaire au cas d'une variable :
$$f(b)-f(a)=\mathrm{d}f(c)(b-a)$$
Revenons au cas de deux variables et à une autre reformulation. Pour $\theta \in]0,1[$, on note
$$a=(x_0,y_0), \quad b=(x_0+h,y_0+k), \quad c=a+\theta(b-a)=(x_0+\theta h, y_0+\theta k) \in]a,b[$$
et
$$\operatorname{grad} f(c) = \begin{pmatrix} \frac{\partial f}{\partial x}(c) \\ \frac{\partial f}{\partial y}(c) \end{pmatrix} \qquad b-a=\begin{pmatrix} h \\ k \end{pmatrix}.$$



Théorème 3 (Théorème des accroissements finis en deux variables)**.**
Soit $f : U \to \mathbb{R}$ *une fonction de classe* $\mathscr{C}^1$ *sur un ouvert* $U \subset \mathbb{R}^2$ *et soit* $(x_0, y_0) \in U$*. Pour tout* $(h,k) \in \mathbb{R}^2$ *tel que* $(x_0+h, y_0+k) \in U$ *et tel que* $(x_0+th, y_0+tk) \in U$ *pour tout* $t \in [0,1]$*, il existe* $\theta \in]0,1[$ *tel que*

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

Exemple 9.
Soient $f(x,y) = e^{x-y^2}$, $a = (0,0)$, $b = (2,1)$.

- Le théorème des accroissements finis nous dit qu'il existe $c \in]a,b[$ tel que
$$f(b) - f(a) = \langle \operatorname{grad} f(c) \mid b-a \rangle .$$
- On a $\frac{\partial f}{\partial x} = e^{x-y^2}$, $\frac{\partial f}{\partial y} = -2ye^{x-y^2}$. Donc le théorème des accroissements finis affirme qu'il existe $\theta \in]0,1[$ tel que
$$f(2,1) - f(0,0) = \left\langle (e^{2\theta-\theta^2}, -2\theta e^{2\theta-\theta^2}) \mid (2,1) \right\rangle = 2(1-\theta)e^{2\theta-\theta^2}.$$
Comme $f(2,1) = e$ et $f(0,0) = 1$, on en déduit qu'il existe θ tel que $e - 1 = 2(1-\theta)e^{2\theta-\theta^2}$.

Démonstration. Soit $g : [0,1] \to \mathbb{R}$ la fonction définie par

$$g(t) = f(x_0+th, y_0+tk).$$

Alors $g = f \circ \gamma$ où $\gamma : [0,1] \to \mathbb{R}^2$ est définie par $\gamma(t) = (x_0+th, y_0+tk)$. La formule de différentiabilité d'une composée s'écrit

$$J_g(t) = J_f(\gamma(t)) \times J_\gamma(t)$$

donc ici :

$$g'(t) = \begin{pmatrix} \frac{\partial f}{\partial x}(\gamma(t)) & \frac{\partial f}{\partial y}(\gamma(t)) \end{pmatrix} \times \begin{pmatrix} \gamma_1'(t) \\ \gamma_2'(t) \end{pmatrix} = \begin{pmatrix} \frac{\partial f}{\partial x}(\gamma(t)) & \frac{\partial f}{\partial y}(\gamma(t)) \end{pmatrix} \times \begin{pmatrix} h \\ k \end{pmatrix} = h\frac{\partial f}{\partial x}(\gamma(t)) + k\frac{\partial f}{\partial y}(\gamma(t))$$

Par le théorème des accroissements finis en une variable appliqué à la fonction $g : [0,1] \to \mathbb{R}$, il existe $\theta \in]0,1[$ tel que $g(1) - g(0) = g'(\theta)(1-0)$, c'est-à-dire :

$$f(x_0+h,y_0+k) - f(x_0,y_0) = h\frac{\partial f}{\partial x}(x_0+\theta h, y_0+\theta k) + k\frac{\partial f}{\partial y}(x_0+\theta h, y_0+\theta k),$$

ou autrement dit

$$f(b) - f(a) = \langle \operatorname{grad} f(c) \mid b-a \rangle .$$

□

Nous allons voir quelques applications du théorème des accroissements finis.

3.3. Inégalité des accroissements finis en deux variables

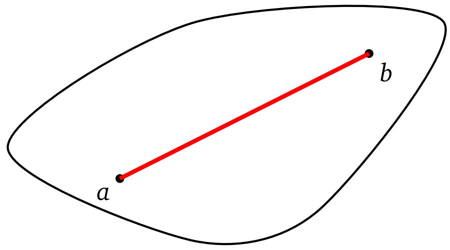
Corollaire 1 (Inégalité des accroissements finis)**.**
Soit $f : U \to \mathbb{R}$ *une fonction de classe* $\mathscr{C}^1$ *sur un ouvert* **convexe** $U \subset \mathbb{R}^2$*. On suppose qu'il existe* $k > 0$ *tel que*

$$\forall c \in U \qquad \|\operatorname{grad} f(c)\| \leqslant k.$$

Alors

$$\forall a,b \in U \qquad |f(b) - f(a)| \leqslant k\|b-a\|.$$

L'hypothèse de convexité permet de s'assurer que tous les points du segment $[a,b]$ appartiennent à U. Le corollaire est une conséquence immédiate du théorème des accroissements finis, à l'aide de l'inégalité de Cauchy-Schwarz $|\langle x \mid y \rangle| \leqslant \|x\| \cdot \|y\|$.



Exemple 10.
Soit $f(x,y) = \sin\left(\frac{x+\pi}{y+1}\right)$. Nous allons utiliser l'inégalité des accroissements finis pour majorer $f(x,y)$ sur $[-\frac{\pi}{2},+\frac{\pi}{2}] \times [-\frac{1}{2},+\frac{1}{2}]$.

- Soit $a = (0,0)$. Alors $f(a) = f(0,0) = 0$.
- Soit $b = (x,y)$ avec $x \in [-\frac{\pi}{2},+\frac{\pi}{2}]$ et $y \in [-\frac{1}{2},+\frac{1}{2}]$.
- Calculons les dérivées partielles :
$$\frac{\partial f}{\partial x}(x,y) = \frac{1}{y+1}\cos\left(\frac{x+\pi}{y+1}\right) \qquad \frac{\partial f}{\partial y}(x,y) = -\frac{x+\pi}{(y+1)^2}\cos\left(\frac{x+\pi}{y+1}\right)$$
- On majore la valeur absolue du cosinus par 1. Pour $x \in [-\frac{\pi}{2},+\frac{\pi}{2}]$ et $y \in [-\frac{1}{2},+\frac{1}{2}]$, on obtient
$$\left|\frac{\partial f}{\partial x}(x,y)\right| \leqslant \frac{1}{-\frac{1}{2}+1} = 2 \qquad \left|\frac{\partial f}{\partial y}(x,y)\right| \leqslant \frac{\frac{\pi}{2}+\pi}{(-\frac{1}{2}+1)^2} = 6\pi$$
- Ainsi, pour ces (x,y), on a $\|\operatorname{grad} f(x,y)\| \leqslant \sqrt{2^2+(6\pi)^2} = 2\sqrt{1+9\pi^2}$. On note $k = 2\sqrt{1+9\pi^2}$ et alors $\|\operatorname{grad} f(x,y)\| \leqslant k$, pour $(x,y) \in [-\frac{\pi}{2},+\frac{\pi}{2}] \times [-\frac{1}{2},+\frac{1}{2}]$.
- L'inégalité des accroissements finis s'écrit
$$|f(b)-f(a)| \leqslant k\|b-a\|.$$
Ici, $b-a = (x,y)$ et $f(a) = 0$. Alors, pour tout $(x,y) \in [-\frac{\pi}{2},+\frac{\pi}{2}] \times [-\frac{1}{2},+\frac{1}{2}]$, on a
$$|f(x,y)| \leqslant k\|(x,y)\|,$$
c'est-à-dire
$$\left|\sin\left(\frac{x+\pi}{y+1}\right)\right| \leqslant 2\sqrt{1+9\pi^2}\sqrt{x^2+y^2}.$$

3.4. Fonctions dont la différentielle est nulle

Il est clair que pour une fonction constante, ses dérivées partielles sont partout nulles. La réciproque est vraie sur un ensemble connexe.

Corollaire 2.
Soit $f : U \to \mathbb{R}$ de classe $\mathscr{C}^1$, où U est un ouvert **connexe** *de $\mathbb{R}^2$. Si $\operatorname{grad} f(x,y) = (0,0)$ pour tout $(x,y) \in U$, alors f est constante sur U.*

Dans la pratique, pour $f : \mathbb{R}^2 \to \mathbb{R}$ qui vérifie $\operatorname{grad} f(x,y) = (0,0)$ pour tout (x,y), on écrit :

- comme $\frac{\partial f}{\partial x}(x,y) = 0$, alors f ne dépend pas de la variable x, c'est-à-dire $f(x,y) = g(y)$ où $g : \mathbb{R} \to \mathbb{R}$;
- mais comme en plus $\frac{\partial f}{\partial y}(x,y) = 0$, alors $g'(y) = 0$, donc $f(x,y) = g(y) = \text{cst}$.

Démonstration. Par le théorème des accroissements finis, si V est une boule ouverte incluse dans U alors, pour tous $a,b \in V$, $f(b) = f(a)$, car $\operatorname{grad} f(c) = 0$ pour tous $c \in [a,b] \subset V \subset U$. La fonction f est donc localement constante, c'est-à-dire constante dans un voisinage de chaque point.
Sur un ensemble connexe, une fonction localement constante est constante. C'est une propriété purement topologique des ensembles connexes. Voici une preuve. Fixons $a \in U$ et $z_0 = f(a)$. Notons $V_1 = f^{-1}(\{z_0\}) = \{b \in U \mid f(b) = z_0\}$. Par le premier paragraphe de cette preuve, V_1 est un ensemble ouvert. Notons $V_2 = f^{-1}(\mathbb{R} \setminus \{z_0\})$. Comme f est une fonction continue, alors l'image réciproque d'un ouvert est un ouvert,

donc V_2 est un ouvert. Il est clair que $U = V_1 \cup V_2$ et que V_1 et V_2 sont disjoints. Comme U est connexe et que V_1 n'est pas vide (car $a \in V_1$) alors V_2 est vide, donc $U = V_1$ et, pour tout $b \in U$, $f(b) = z_0$. □

Mini-exercices.

1. Soient $f(x, y) = x + y^2$, $a = (0, 0)$, $b = (h, k)$. Appliquer le théorème des accroissement finis entre a et b. Trouver explicitement la valeur de $\theta \in]0, 1[$ (ou bien le point $c \in]a, b[$) du théorème.
2. Soit $f(x, y) = \ln(1 + x^2 - y)$. Trouver une constante k telle que $|f(x, y)| \leqslant k\sqrt{x^2 + y^2}$ pour tout $(x, y) \in [-\frac{1}{2}, +\frac{1}{2}] \times [-\frac{1}{2}, +\frac{1}{2}]$.
3. Trouver toutes les fonctions $f : \mathbb{R}^2 \to \mathbb{R}$ telles que $\operatorname{grad} f(x, y) = (1, -1)$ pour tout $(x, y) \in \mathbb{R}^2$.

Auteurs du chapitre

Arnaud Bodin. D'après des cours de Abdellah Hanani (Lille), Goulwen Fichou et Stéphane Leborgne (Rennes), Laurent Pujo-Menjouet (Lyon). Relu par Anne Bauval et Vianney Combet.

Chapitre 5

Difféomorphismes

1. Difféomorphismes

1.1. Définition

Les applications de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qui sont bijectives et de classe $\mathscr{C}^1$ ainsi que leur réciproque, sont utilisées comme changements de variables. On les appelle des difféomorphismes.

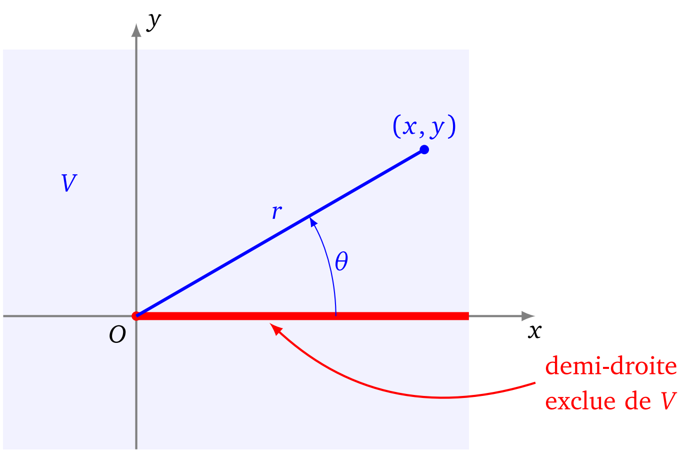
Définition 1.
Soient U et V deux ouverts de $\mathbb{R}^n$ et $\Phi : U \to V$. On dit que Φ est un ***difféomorphisme*** si :

1. Φ est une bijection de U sur V,
2. Φ est de classe $\mathscr{C}^1$ sur U, c'est-à-dire différentiable sur U, de différentielle continue,
3. et Φ^{-1} est de classe $\mathscr{C}^1$ sur V.

Remarque : on peut être plus précis en parlant d'un $\mathscr{C}^1$-difféomorphisme.

Exemple 1.

1. $f :]0,+\infty[\to]0,+\infty[$ avec $f(x) = x^3$ est un difféomorphisme. Son inverse est $f^{-1} : \mathbb{R} \to \mathbb{R}$ définie par $f^{-1}(x) = x^{\frac{1}{3}}$. f et f^{-1} sont bien des applications $\mathscr{C}^1$.
2. Un exemple fondamental qu'on développera dans ce chapitre est le passage en coordonnées polaires : on remplace les coordonnées cartésiennes (x, y) par les coordonnées polaires (r, θ). Il faut exclure certaines parties du plan afin d'obtenir un difféomorphisme. Soit $U = \{(r, \theta) \in \mathbb{R}^2 \mid r > 0 \text{ et } 0 < \theta < 2\pi\}$ et $V = \mathbb{R}^2 \setminus (\mathbb{R}_+ \times \{0\})$ alors $\Phi : U \to V$ défini par :
$$\Phi(r, \theta) = (r \cos\theta, r \sin\theta)$$
est un difféomorphisme. Trouver son inverse revient à exprimer les coordonnées polaires (r, θ) en fonction des coordonnées cartésiennes (x, y).



1.2. Difféomorphisme et jacobienne

Du théorème de composition découle que les matrices jacobiennes de Φ et Φ^{-1} sont inverses l'une de l'autre. Soient U et V deux ouverts de $\mathbb{R}^n$ et $\Phi : U \to V$. Soit $x \in U$ et $y = \Phi(x) \in V$. Notons $J_\Phi(x)$ la matrice jacobienne de Φ en x et $J_{\Phi^{-1}}(y)$ la matrice jacobienne de Φ^{-1} en y. Ce sont des matrices carrées $n \times n$.

Proposition 1.

$$\boxed{J_{\Phi^{-1}}(y) = (J_\Phi(x))^{-1}}$$

Comme on l'a dit : la matrice jacobienne de l'inverse est l'inverse de la matrice jacobienne !

Démonstration. Notons $J = J_\Phi(x)$, $y = \Phi(x)$, $\Psi = \Phi^{-1}$ et $K = J_\Psi(y)$. Comme Ψ est l'inverse de Φ alors :

$$\Psi \circ \Phi(x) = x.$$

On applique la formule de la composition des jacobiennes :

$$J_{\Psi\circ\Phi}(x) = J_\Psi(\Phi(x)) \times J_\Phi(x).$$

L'application $\Psi \circ \Phi$ est l'identité, donc sa matrice jacobienne est la matrice identité : $J_{\Psi\circ\Phi}(x) = I$. Quant au terme de droite, c'est exactement $K \times J$. Ainsi $KJ = I$ et donc $K = J^{-1}$. □

Exemple 2 (Cas d'une variable).
Soit $f : I \to J$ une fonction d'une variable où I et J sont des intervalles ouverts de $\mathbb{R}$. On suppose que f est une fonction de classe $\mathscr{C}^1$, bijective et que la dérivée de f ne s'annule pas sur I. Alors, pour $y \in J$, on retrouve la formule connue :

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)}.$$

Rappelons que pour une fonction d'une variable, la matrice jacobienne $J_f(x)$ est la matrice 1×1 avec pour seul coefficient $f'(x)$. Par exemple, pour $f(x) = \sin(x)$ en $x_0 = \frac{\pi}{3}$, on sait que $y_0 = \sin(\frac{\pi}{3}) = \frac{\sqrt{3}}{2}$ et $f'(x_0) = \cos(\frac{\pi}{3}) = \frac{1}{2}$. La bijection réciproque est $f^{-1}(y) = \arcsin(y)$, et on peut calculer $\arcsin'(\frac{\sqrt{3}}{2})$ directement :

$$\arcsin'\left(\frac{\sqrt{3}}{2}\right) = \arcsin'(y_0) = \frac{1}{\sin'(x_0)} = 2.$$

Exemple 3.
Soit $F : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F(x,y) = (x - y^2, y+1)$.

1. *Matrice jacobienne de F.* Notons $f_1(x,y) = x - y^2$ et $f_2(x,y) = y+1$. Alors la matrice jacobienne de F est :
$$J_F(x,y) = \begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & -2y \\ 0 & 1 \end{pmatrix}$$
2. *Inverse de F.* Notons $F(x,y) = (X,Y)$, alors F est bijective et son inverse est $F^{-1}(X,Y) = G(X,Y) = (X + (Y-1)^2, Y-1)$: à vous de vérifier qu'on a bien $G \circ F(x,y) = (x,y)$.
3. *Jacobienne de F^{-1}.* Par la formule $G(X,Y) = (g_1(X,Y), g_2(X,Y))$ définissant l'inverse on a directement :
$$J_{F^{-1}}(X,Y) = \begin{pmatrix} \frac{\partial g_1}{\partial X} & \frac{\partial g_1}{\partial Y} \\ \frac{\partial g_2}{\partial X} & \frac{\partial g_2}{\partial Y} \end{pmatrix} = \begin{pmatrix} 1 & 2(Y-1) \\ 0 & 1 \end{pmatrix}$$
4. *Jacobienne de F^{-1} (encore).* Une autre façon d'obtenir la matrice jacobienne de F^{-1} est de d'abord calculer l'inverse de $J_F(x,y)$ (qui est inversible car de déterminant 1) puis on applique la proposition 1, en se souvenant que $F(x,y) = (X,Y)$ avec $Y = y+1$ et donc $y = Y - 1$:
$$J_{F^{-1}}(X,Y) = \left(J_F(x,y)\right)^{-1} = \begin{pmatrix} 1 & +2y \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2(Y-1) \\ 0 & 1 \end{pmatrix}$$

On retrouve le résultat précédent.

1.3. Déterminant jacobien

Pour un difféomorphisme, le déterminant de la matrice jacobienne joue un rôle particulier.
On appelle ***déterminant jacobien*** de Φ, en $x \in U$, le déterminant de la matrice jacobienne de Φ en x :

$$\det\left(J_\Phi(x)\right).$$

Il est clair que le déterminant jacobien d'un difféomorphisme ne s'annule pas, puisque la matrice jacobienne est inversible. La réciproque sera donnée par le théorème d'inversion locale.

Mini-exercices.

1. Soit $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ une application linéaire, de matrice A inversible. Monter que Φ est un difféomorphisme (on sait qu'une application linéaire est continue). Quelle est la matrice de Φ ? Et celle de Φ^{-1} ?
2. Le changement de coordonnées cylindriques est le changement de variables $(x, y, z) = (r\cos\theta, r\sin\theta, z)$. On se limite ici au domaine U défini par $r > 0$, $\theta \in [0, \frac{\pi}{2}]$, $z \in \mathbb{R}$. Calculer l'ensemble V des (x, y, z) correspondants. On définit $\Phi : U \to V$ par $(r, \theta, z) \mapsto (x, y, z)$. Montrer que Φ est un difféomorphisme. Calculer sa matrice jacobienne et le déterminant jacobien. Calculer la matrice jacobienne de Φ^{-1}.

2. Théorème d'inversion locale

Le théorème d'inversion locale et celui d'inversion globale permettent de montrer qu'une application est un difféomorphisme.

2.1. Théorème d'inversion locale

Théorème 1 (Théorème d'inversion locale)**.**
Soient U et V deux ouverts de $\mathbb{R}^n$ et $\Phi : U \to V$. Soit $x \in U$. Si :

1. *$\Phi : U \to V$ est de classe $\mathscr{C}^1$,*
2. *et si la matrice jacobienne $J_\Phi(x)$ est inversible,*

alors Φ est localement inversible, c'est-à-dire il existe un ouvert U_x contenant x et un ouvert V_y contenant $y = \Phi(x)$ tels que $\Phi : U_x \to V_y$ soit un difféomorphisme.

Nous admettons ce théorème dont la preuve est assez technique.

On obtient deux autres formulations équivalentes, si on remplace la condition 2. par la condition :
2'. le déterminant jacobien $\det J_\Phi(x)$ est non nul,
ou, si on préfère le langage des différentielles :
2". la différentielle $\mathrm{d}F(x) : \mathbb{R}^n \to \mathbb{R}^n$ est une application linéaire inversible.

Conclusion : on sait qu'une application linéaire $L : \mathbb{R}^n \to \mathbb{R}^n$ est inversible si le déterminant de la matrice A associée est non nul. Le théorème d'inversion locale est donc une extension puissante de ce résultat pour une application non linéaire : si le déterminant jacobien est non nul alors l'application est (localement) inversible.

2.2. Théorème d'inversion globale

Le théorème d'inversion globale est encore plus intéressant car il conduit à une fonction qui est partout un difféomorphisme.

Théorème 2 (Théorème d'inversion globale)**.**
Soient U et V deux ouverts de $\mathbb{R}^n$ et $\Phi : U \to V$. Si

1. *$\Phi : U \to V$ est de classe $\mathscr{C}^1$,*
2. *$\Phi : U \to V$ est bijective,*
3. *et le déterminant jacobien $\det J_\Phi(x)$ est non nul, pour tout $x \in U$,*

alors $\Phi : U \to V$ est un difféomorphisme.

C'est une application du théorème d'inversion locale en tout point de l'ensemble de définition.

2.3. Exemple

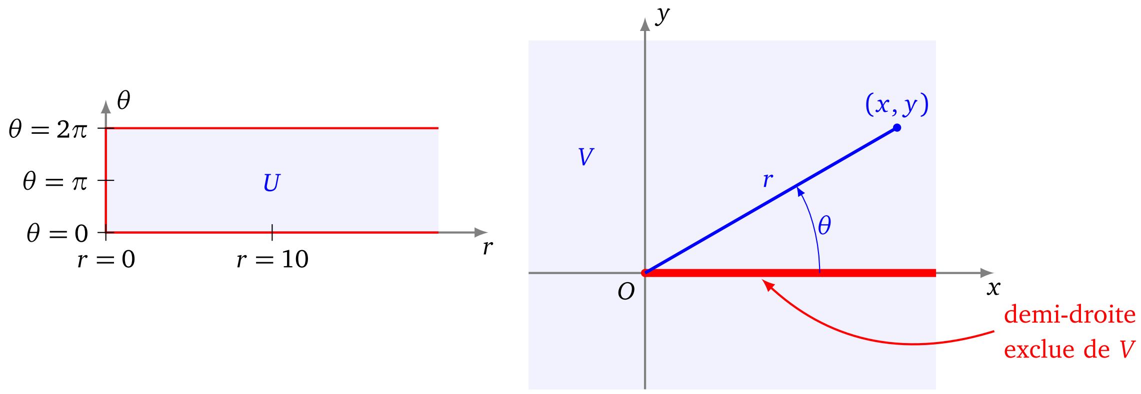
Les passages en coordonnées polaires, cylindriques ou sphériques, sont très souvent utilisés. Détaillons le premier qui consiste à remplacer les coordonnées cartésiennes (x, y) d'un point du plan, par le module r et l'argument θ du point dans le plan complexe.

Coordonnées polaires.

- *Définition de Φ.*

$$\begin{array}{cccc} \Phi & : & U =]0,+\infty[\times]0,2\pi[& \to & V = \mathbb{R}^2 \setminus \left(\mathbb{R}^+ \times \{0\}\right) \\ & & (r,\theta) & \mapsto & (x,y) = (r\cos\theta, r\sin\theta). \end{array}$$

  Clairement Φ est une application $\mathscr{C}^1$.



- *Inverse de Φ.* Il s'agit d'exprimer (r,θ) en fonction de (x,y). On sait $x^2+y^2=r^2$, donc $r=\sqrt{x^2+y^2}$. On note $\theta \in]0,2\pi[$ l'angle entre l'axe des abscisses et le point (x,y). Il est définit de manière unique (et pourrait être déterminé par des formules, par exemple $\theta = \arctan(\frac{y}{x})$ pour $x>0$, $y>0$).
  L'inverse de Φ est alors $\Phi^{-1} : V \to U$ défini par $\Phi^{-1}(x,y) = (r,\theta)$.
- *Matrice jacobienne de Φ.*

$$J_\Phi(r,\theta) = \begin{pmatrix} \frac{\partial(r\cos\theta)}{\partial r} & \frac{\partial(r\cos\theta)}{\partial\theta} \\ \frac{\partial(r\sin\theta)}{\partial r} & \frac{\partial(r\sin\theta)}{\partial\theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}.$$

- *Déterminant jacobien de Φ.*

$$\det J_\Phi(r,\theta) = \det\begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = r\cos^2\theta + r\sin^2\theta = r.$$

- *Théorème d'inversion locale.* Comme $r>0$ sur U, le déterminant jacobien ne s'annule jamais, donc dans un voisinage de chaque point $(r,\theta)\in U$, Φ est localement un difféomorphisme. Ainsi Φ est inversible (on l'avait déjà dit) mais en plus Φ^{-1} est $\mathscr{C}^1$ (on ne le savait pas car on n'avait pas explicité Φ^{-1}).

- *Théorème d'inversion globale.* On peut faire mieux en appliquant le théorème d'inversion globale (vérifier les hypothèses !) : $\Phi : U \to V$ est un difféomorphisme. On retrouve que Φ est inversible d'inverse de classe $\mathscr{C}^1$ sur V.
- *Matrice jacobienne de Φ^{-1}.*

$$J_{\Phi^{-1}}(x,y) = \big(J_\Phi(r,\theta)\big)^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} & \frac{y}{\sqrt{x^2+y^2}} \\ -\frac{y}{x^2+y^2} & \frac{x}{x^2+y^2} \end{pmatrix}$$

où on a utilisé $x = r\cos\theta$, $y = r\sin\theta$ et $r = \sqrt{x^2+y^2}$.

Application.
Considérons maintenant une application $(x,y) \mapsto f(x,y)$ de U (défini ci-dessus) dans $\mathbb{R}$. Pour « passer en coordonnées polaires », on doit remplacer les anciennes coordonnées (x,y) par les nouvelles coordonnées (r,θ). On compose par la fonction Φ pour définir :

$$g(r,\theta) = f \circ \Phi(r,\theta) = f\,(r\cos\theta, r\sin\theta).$$

Supposons que l'on connaisse les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$. Nous avons besoin de calculer les dérivées partielles de la nouvelle fonction g : $\frac{\partial g}{\partial r}$ et $\frac{\partial g}{\partial \theta}$. Elles sont données par les formules :

$$\boxed{\begin{cases} \dfrac{\partial g}{\partial r}(r,\theta) &= \cos\theta\dfrac{\partial f}{\partial x}(x,y) + \sin\theta\dfrac{\partial f}{\partial y}(x,y) \\[2ex] \dfrac{\partial g}{\partial \theta}(r,\theta) &= -r\sin\theta\dfrac{\partial f}{\partial x}(x,y) + r\cos\theta\dfrac{\partial f}{\partial y}(x,y) \end{cases}}$$

Dans ces formules, pour n'avoir que des variables r et θ, il faut substituer $x = r\cos\theta$ et $y = r\sin\theta$.

Preuve. D'une part la matrice jacobienne de g est la matrice ligne :

$$J_g(r,\theta) = \Big(\tfrac{\partial g}{\partial r}(r,\theta) \quad \tfrac{\partial g}{\partial \theta}(r,\theta)\Big)$$

Mais d'autre part, comme $g(r,\theta) = f \circ \Phi(r,\theta)$, alors

$$J_g(r,\theta) = J_f\big(\Phi(r,\theta)\big) \times J_\Phi(r,\theta).$$

On sait que $\Phi(r,\theta) = (x,y)$, on a déjà calculé $J_\Phi(r,\theta)$ et la matrice jacobienne de f est la matrice ligne :

$$J_f(x,y) = \Big(\tfrac{\partial f}{\partial x}(x,y) \quad \tfrac{\partial f}{\partial y}(x,y)\Big)$$

La formule résulte alors du produit de matrice $J_g(r,\theta) = J_f(x,y) \times J_\Phi(r,\theta)$.

Exercice. Exprimer $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ en fonction de $\frac{\partial g}{\partial r}$ et $\frac{\partial g}{\partial \theta}$. Indication : $f(x,y) = g \circ \Phi^{-1}(x,y)$ et nous avons déjà calculé $J_{\Phi^{-1}}(x,y)$.

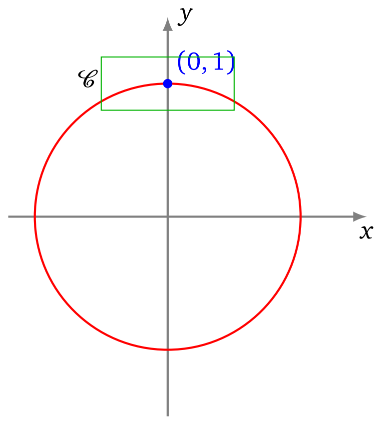
Mini-exercices.

1. Soit $U = \mathbb{R}^2 \setminus \{(0,0)\}$ le plan privé de l'origine et $\Phi(x,y) = (x^2 - y^2, 2xy)$. Montrer que Φ est un difféomorphisme local au voisinage de tout point de U. Montrer que Φ n'est pas un difféomorphisme global (indication : montrer que Φ n'est pas injective).
2. Soit $U = \mathbb{R}\times]0,2\pi[$ et $\Phi : U \to \mathbb{R}^2$ l'application définie par $(x,y) \mapsto (e^x\cos y, e^x\sin y)$. Montrer que Φ est un difféomorphisme local au voisinage de tout point de U. Montrer que Φ est injective. Déterminer $\Phi(U)$ et vérifier que c'est un ouvert. En déduire que $\Phi : U \to \Phi(U)$ est un difféomorphisme (global).

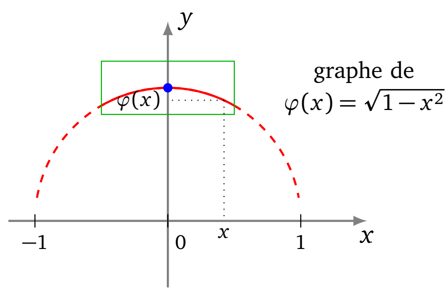
3. Théorème des fonctions implicites

3.1. Motivation

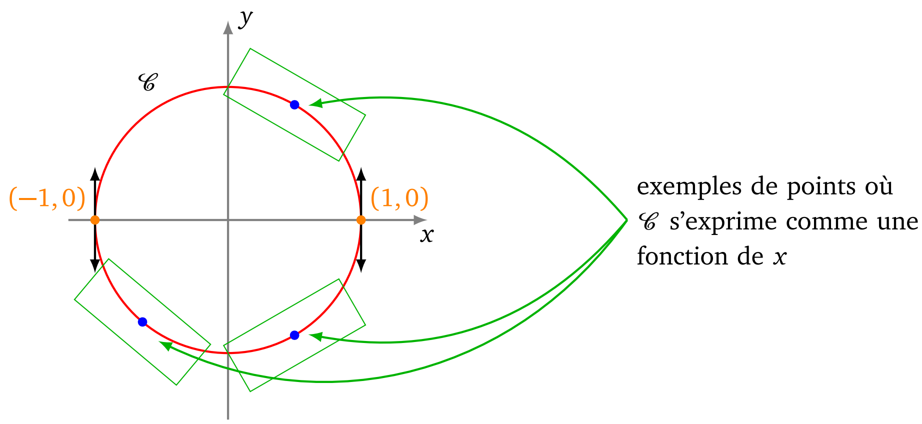
Le théorème des fonctions implicites est assez technique et son énoncé est difficile à comprendre, cependant l'idée sous-jacente est assez simple. Comprenons donc d'abord le théorème sur un exemple. Il s'agit de remplacer l'étude d'une fonction de deux variables, par une fonction d'une seule variable. Partons de la fonction de deux variables $F : \mathbb{R}^2 \to \mathbb{R}$ définie par $F(x, y) = x^2 + y^2 - 1$. On souhaite étudier l'ensemble $(F(x, y) = 0)$, c'est-à-dire l'ensemble des points $(x, y) \in \mathbb{R}^2$ tels que $x^2 + y^2 = 1$. Géométriquement c'est donc le cercle $\mathscr{C}$ de rayon 1, centré à l'origine.



Plaçons-nous autour du point $(0, 1) \in \mathscr{C}$. Alors, dans un voisinage de ce point, la portion de cercle $\mathscr{C}$ est le graphe $y = \varphi(x)$ d'une fonction $x \mapsto \varphi(x)$. On ne cherche pas à expliciter cette fonction mais en fait c'est $\varphi(x) = \sqrt{1 - x^2}$.

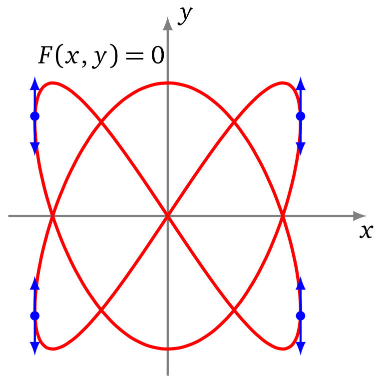


Est-ce qu'on peut faire ce travail pour tous les points du cercle ? Presque ! C'est possible pour tous les points, sauf autour des points $(1, 0)$ et $(-1, 0)$. Autour de ces deux points la portion de cercle n'est pas le graphe d'une fonction de x. L'explication c'est qu'en ces points il y a une tangente verticale, et que le graphe d'une fonction de x dérivable n'a jamais de tangente verticale.



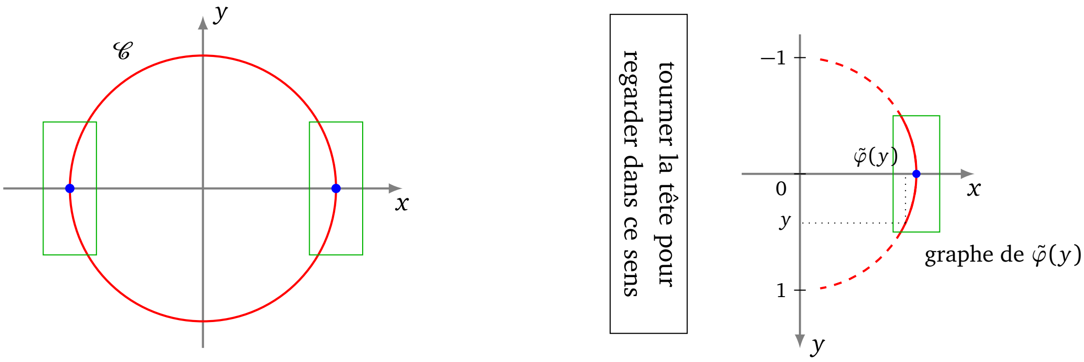
Comment détecter les points de tangence verticale d'une courbe quelconque définie par $(F(x,y)=0)$?

$$\boxed{\text{les points de tangence verticale sur } (F(x,y)=0) \text{ sont les points où } \frac{\partial F}{\partial y}(x,y)=0 \text{ (et } \frac{\partial F}{\partial x}(x,y)\neq 0).}$$



C'est une formule valable pour n'importe quelle fonction $F:\mathbb{R}^2\to\mathbb{R}$. Faîtes le calcul pour notre fonction : si $x^2+y^2=1$ et $\frac{\partial F}{\partial y}(x,y)=0$ alors $(x,y)=(\pm 1,0)$.

Comment faire pour les deux points spéciaux, par exemple en $(1,0)$? On peut tout simplement considérer la portion de cercle $\mathscr{C}$ autour de $(1,0)$ comme le graphe d'une fonction de y : $y\mapsto\tilde{\varphi}(y)$. C'est un peu déroutant car les rôles de x et de y sont inversés ; c'est comme si on avait tourné la feuille de 90°.



En fait le cercle $\mathscr{C}$ est localement le graphe d'une fonction $y\mapsto\tilde{\varphi}(y)$, sauf au point ayant une tangente *horizontale*, qui sont donnés cette fois par l'équation $\frac{\partial F}{\partial x}(x,y)=0$ (et $\frac{\partial F}{\partial y}(x,y)\neq 0$).

De façon plus générale pour $F:\mathbb{R}^n\to\mathbb{R}$ une fonction de n variables, on remplace l'étude de $(F(x_1,\ldots,x_n)=0)$ par l'étude d'une fonction φ de $n-1$ variables à condition que l'une des dérivées partielle ne s'annule pas. Le théorème des fonctions implicites est surtout un résultat théorique, car la fonction φ n'est pas fournie explicitement.

3.2. Cas $F(x,y)=0$

Soit $F:\mathbb{R}^2\to\mathbb{R}$. On considère la courbe de niveau $\mathscr{C}:\big(F(x,y)=0\big)$.
On dit que la fonction $y=\varphi(x)$ est ***définie implicitement*** par $F(x,y)=0$ si $F(x,\varphi(x))=0$, c'est-à-dire si $(x,\varphi(x))\in\mathscr{C}$. Alors on dit que $y=\varphi(x)$ est une ***fonction implicite*** de $F(x,y)=0$.

Théorème 3 (des fonctions implicites)**.**
Soient $F:\mathbb{R}^2\to\mathbb{R}$ une fonction de classe $\mathscr{C}^1$ et (x_0,y_0) un point tel que $F(x_0,y_0)=0$. Si

$$\boxed{\frac{\partial F}{\partial y}(x_0,y_0)\neq 0}$$

alors :

1. *Il existe une fonction $\varphi : I \to J$ de classe $\mathscr{C}^1$ définissant une fonction implicite $y = \varphi(x)$ où I est un intervalle ouvert contenant x_0, J est un intervalle ouvert contenant y_0 et $y_0 = \varphi(x_0)$. Plus précisément, pour tout $(x, y) \in I \times J$ on a :*

$$\boxed{F(x,y) = 0 \iff y = \varphi(x).}$$

En particulier :

$$\boxed{\text{pour tout } x \in I,\ F(x,\varphi(x)) = 0.}$$

2. *De plus, la dérivée de φ est donnée par*

$$\boxed{\varphi'(x) = \frac{-\frac{\partial F}{\partial x}(x,\varphi(x))}{\frac{\partial F}{\partial y}(x,\varphi(x))}}$$

en tout point de $x \in I$ où $\frac{\partial F}{\partial y}(x,\varphi(x)) \neq 0$.

On a un énoncé symétrique en échangeant x et y si c'est l'autre dérivée partielle qui ne s'annule pas. Si $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$, il existe une fonction $\tilde{\varphi} : I \to \mathbb{R}$ de la variable y, définie autour de y_0, qui définit la fonction implicite $\tilde{\varphi}(y)$, de sorte que dans un voisinage ouvert de (x_0, y_0) :

$$F(x,y) = 0 \iff x = \tilde{\varphi}(y).$$

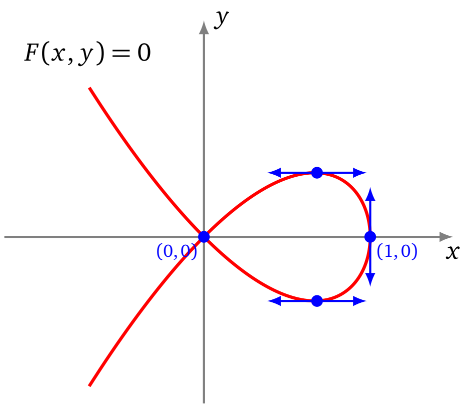
(ce qui implique ici $F(\tilde{\varphi}(y), y) = 0$).
Si les deux dérivées partielles $\frac{\partial F}{\partial x}(x_0, y_0)$ et $\frac{\partial F}{\partial y}(x_0, y_0)$ s'annulent simultanément, le point (x_0, y_0) s'appelle un ***point singulier*** et le théorème des fonctions implicites ne s'applique pas.

3.3. Exemple

Soit

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

Courbe $F(x,y) = 0$**.**
Commençons par tracer la courbe $\mathscr{C}$ d'équation ($F(x,y) = 0$), ce qui nous permettra d'anticiper les calculs. La courbe forme une boucle. On note qu'il y a une tangente verticale au point $(1,0)$. Le point $(0,0)$ est un point double, on verra que les deux dérivées partielles s'annulent en même temps : c'est un point singulier. En dehors de ces deux points, on peut donc exprimer localement la courbe $\mathscr{C}$ comme le graphe d'une fonction d'une variable $y = \varphi(x)$.



Il y a deux points P_1 et P_2 où la tangente est horizontale. En dehors des trois points P_1, P_2 et $(0,0)$ on pourrait aussi exprimer localement la courbe $\mathscr{C}$ comme le graphe d'une fonction d'une variable $x = \tilde{\varphi}(y)$.

Dérivées partielles.
On a $F(x,y) = x^3 + xy^2 - x^2 + y^2$. Donc

$$\frac{\partial F}{\partial x}(x,y) = 3x^2 + y^2 - 2x \qquad \text{et} \qquad \frac{\partial F}{\partial y}(x,y) = 2xy + 2y.$$

Pour chercher les points de tangence verticale, on résout d'abord l'équation $\frac{\partial F}{\partial y}(x,y) = 0$.

$$\frac{\partial F}{\partial y}(x,y) = 0 \iff 2y(x+1) = 0 \iff y = 0 \text{ ou } x = -1.$$

En plus on cherche un point $(x,y) \in \mathscr{C}$.
Cas $y = 0$. Alors, comme on veut en plus $F(x,y) = 0$, on a $x^3 - x^2 = 0$, donc $x = 0$ ou $x = 1$. Ainsi il y a deux solutions $(0,0)$ et $(1,0)$.
Cas $x = -1$. On veut en plus $F(x,y) = 0$, mais l'équation $F(-1,y) = 0$ équivaut à $-2 + y^2 - y^2 = 0$ qui n'a pas de solution.

Tangence verticale.
Conclusion :
- $(1,0) \in \mathscr{C}$ est un point de tangence verticale : la dérivée $\frac{\partial F}{\partial y}$ s'annule mais pas $\frac{\partial F}{\partial x}$ (faites le calcul).
- $(0,0) \in \mathscr{C}$ est un point singulier : les deux dérivées partielles s'annulent.

Théorème des fonctions implicites.
Ainsi pour n'importe quel point (x_0, y_0) de $\mathscr{C}$ autre que $(0,0)$ et $(1,0)$, il existe $\varphi : I \to J$ (où I est un intervalle ouvert contenant x_0 et J un intervalle ouvert contenant y_0) tel que

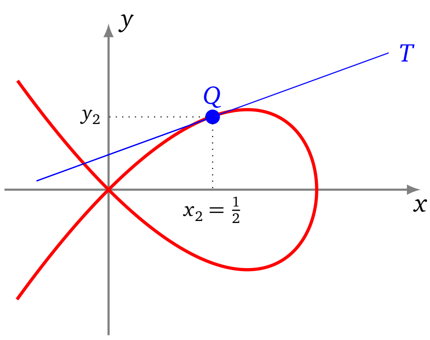
$$F(x,y) = 0 \iff y = \varphi(x)$$

(pour $x \in I$, $y \in J$).

Théorème des fonctions implicites (bis).
On peut aussi appliquer le théorème des fonctions implicites là où la courbe n'a pas de tangentes horizontales (ni point singulier).
La courbe $\mathscr{C}$ admet deux tangentes horizontales en $P_1 = (x_1, y_1)$ et $P_2 = (x_1, -y_1)$ où $x_1 = \frac{\sqrt{5}-1}{2}$ et $y_1 = \sqrt{\frac{5\sqrt{5}-11}{2}}$. Pour tout $(x_0, y_0) \in \mathscr{C}$ autre que P_1, P_2 et $(0,0)$ il existe $\tilde{\varphi} : I \to J$ de la variable y (cette fois I est un intervalle ouvert contenant y_0 et J un intervalle ouvert contenant x_0) tel que

$$F(x,y) = 0 \iff x = \tilde{\varphi}(y).$$

Équation d'une tangente.
Même si on ne connaît pas l'expression de φ, le théorème des fonctions implicites permet de calculer des valeurs de la dérivée de $\varphi(x)$ (à condition de connaître la valeur $\varphi(x)$ correspondante).
Considérons le point $Q = (x_2, y_2) \in \mathscr{C}$ pour lequel $x_2 = \frac{1}{2}$ et $y_2 > 0$. En résolvant l'équation $F(\frac{1}{2}, y) = 0$ on trouve la solution positive $y_2 = \frac{\sqrt{3}}{6}$. Calculons l'équation de la tangente à $\mathscr{C}$ en Q.



Appliquons le théorème des fonctions implicites autour de Q : il existe une fonction $y = \varphi(x)$ dont le graphe (autour de $x = \frac{1}{2}$) est exactement la courbe $\mathscr{C}$ (autour de Q). Ainsi la tangente T à $\mathscr{C}$ en Q est aussi la tangente au graphe de φ en $x_2 = \frac{1}{2}$. On sait comment calculer l'équation d'une tangente au graphe de la fonction d'une variable φ, il suffit de calculer son coefficient directeur donné par $\varphi'(x_2)$.
Tout d'abord, par définition de φ, en $x_2 = \frac{1}{2}$ on a $\varphi(x_2) = y_2$. La seconde partie du théorème des fonctions implicites nous permet de calculer $\varphi'(x_2)$:

$$\begin{aligned}\varphi'(x_2) &= -\frac{\frac{\partial F}{\partial x}(x_2, \varphi(x_2))}{\frac{\partial F}{\partial y}(x_2, \varphi(x_2))} = -\frac{\frac{\partial F}{\partial x}(x_2, y_2)}{\frac{\partial F}{\partial y}(x_2, y_2)} \\ &= -\frac{\frac{\partial F}{\partial x}(\frac{1}{2}, \frac{\sqrt{3}}{6})}{\frac{\partial F}{\partial y}(\frac{1}{2}, \frac{\sqrt{3}}{6})} = -\frac{-\frac{1}{6}}{\frac{\sqrt{3}}{2}} \\ &= \frac{\sqrt{3}}{9}\end{aligned}$$

Conclusion l'équation de la tangente en Q est $y = (x - x_2)\varphi'(x_2) + y_2$, c'est donc $y = \frac{\sqrt{3}}{9}(x - \frac{1}{2}) + \frac{\sqrt{3}}{6}$, que l'on peut aussi écrire $y = \frac{\sqrt{3}}{9}(x + 1)$.

3.4. Preuve

Théorème d'inversion locale.
Soit F une fonction $\mathscr{C}^1$ de deux variables et $(x_0, y_0) \in \mathbb{R}^2$ tel que $F(x_0, y_0) = 0$ et $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$. Considérons la fonction Φ définie par

$$\Phi(x, y) = (x, F(x, y)).$$

La matrice jacobienne de Φ en (x_0, y_0) est

$$J_\Phi(x_0, y_0) = \begin{pmatrix} 1 & 0 \\ \frac{\partial F}{\partial x}(x_0, y_0) & \frac{\partial F}{\partial y}(x_0, y_0) \end{pmatrix}.$$

Par hypothèse $\frac{\partial F}{\partial y}$ ne s'annule pas en (x_0, y_0). La matrice $J_\Phi(x_0, y_0)$ est donc inversible et d'après le théorème d'inversion locale, Φ est localement inversible en (x_0, y_0) : il existe un ouvert U contenant (x_0, y_0) et un ouvert V, tels que $\Phi : U \to V$ soit un difféomorphisme. Notons $\Psi = \Phi^{-1} : V \to U$ son inverse et décomposons Ψ en $\Psi(X, Y) = (\psi_1(X, Y), \psi_2(X, Y))$.

Expression de Ψ. Comme Ψ est l'inverse de Φ alors d'une part

$$\Phi \circ \Psi(X, Y) = (X, Y),$$

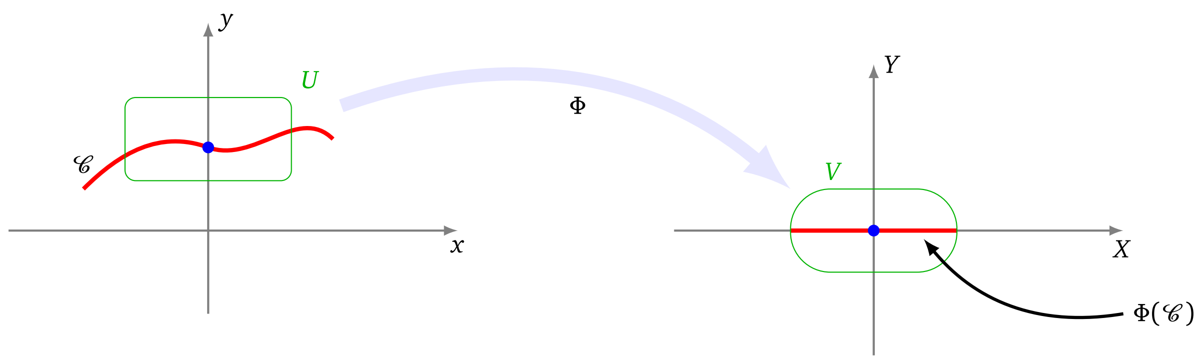
mais d'autre part :

$$\Phi \circ \Psi(X, Y) = \Phi\big(\psi_1(X, Y), \psi_2(X, Y)\big) = \big(\psi_1(X, Y), F(\psi_1(X, Y), \psi_2(X, Y))\big).$$

Ainsi :

$$\psi_1(X, Y) = X \qquad \text{et} \qquad F\big(X, \psi_2(X, Y)\big) = Y.$$

Équation $F(x, y) = 0$. Soit $(x, y) \in U$ un point vérifiant $F(x, y) = 0$. Alors son image par le difféomorphisme Φ est le point $(X, Y) \in V$ tel que $(X, Y) = \Phi(x, y) = (x, F(x, y)) = (x, 0)$. Donc $X = x$ et $Y = 0$.
Ainsi les points de U vérifiant $F(x, y) = 0$ correspondent aux points de la forme $(X, 0)$ de V. Autrement dit, le difféomorphisme transforme localement la courbe $(F(x, y) = 0)$ en un segment horizontal.



Retenons que $X = x$ et $Y = 0$ pour les points qui nous intéressent. Définissons alors $\varphi : I \subset \mathbb{R} \to \mathbb{R}$ par $\varphi(x) = \psi_2(x,0)$. C'est une fonction $\mathscr{C}^1$ (car Ψ l'est), définie sur un intervalle ouvert I contenant x_0. Reprenons le calcul (dans U et V) :

$$\begin{aligned} F(x,y)=0 &\iff \Phi(x,y)=(x,0) \\ &\iff (x,y)=\Psi(x,0) \\ &\iff (x,y)=(x,\psi_2(x,0)) \\ &\iff y=\varphi(x) \end{aligned}$$

Ainsi l'équation $F(x,y)=0$ équivaut à $y=\varphi(x)$ (dans un voisinage de (x_0,y_0)).

3.5. Cas $F(x_1,\ldots,x_n)=0$

L'étude est similaire pour les hypersurfaces de niveau en plusieurs variables, où on va pouvoir exprimer une variable en fonction des autres si la dérivée partielle correspondante n'est pas nulle.

Théorème 4.
Si $F : \mathbb{R}^n \to \mathbb{R}$ est de classe $\mathscr{C}^1$ et si $\dfrac{\partial F}{\partial x_n}(a_1,\ldots,a_n) \neq 0$ alors :

1. *La fonction implicite $x_n = \varphi(x_1,\ldots,x_{n-1})$ existe sur une boule ouverte $B \subset \mathbb{R}^{n-1}$ centrée en $(a_1,\ldots,a_{n-1})$ et un intervalle ouvert J contenant x_n avec $\varphi : B \to J$ et on a*
$$F(x_1,\ldots,x_{n-1},x_n)=0 \iff x_n=\varphi(x_1,\ldots,x_{n-1}).$$
En particulier $F(x_1,\ldots,x_{n-1},\varphi(x_1,\ldots,x_{n-1}))=0$.
2. $$\frac{\partial \varphi}{\partial x_i}(x_1,\ldots,x_{n-1},x_n)=\frac{-\frac{\partial f}{\partial x_i}(x_1,\ldots,x_{n-1},\varphi(x_1,\ldots,x_{n-1}))}{\frac{\partial f}{\partial x_n}(x_1,\ldots,x_{n-1},\varphi(x_1,\ldots,x_{n-1}))}$$

Bien sûr il existe un énoncé similaire, obtenu en remplaçant n par k, pour chaque situation où $\dfrac{\partial F}{\partial x_k}(a_1,\ldots,a_n) \neq 0$.

Prenons comme exemple le cas de $F : \mathbb{R}^3 \to \mathbb{R}$. Soit $(x_0,,y_0,z_0)$ un point où $\frac{\partial F}{\partial z}(x_0,,y_0,z_0) \neq 0$. Alors il existe un disque ouvert $D \subset \mathbb{R}^2$ centré en (x_0,y_0), un intervalle ouvert J contenant z_0 et une unique fonction $\varphi : D \to J$, $z=\varphi(x,y)$ telle que $F(x,y,\varphi(x,y))=0$ (pour tout $(x,y) \in D$ et $z \in J$).

Mini-exercices.

1. Soit $F(x,y)=xy^2+y^2+x-6y+1$ et $\mathscr{C} : (F(x,y)=0)$. Calculer $\frac{\partial F}{\partial x}$ et $\frac{\partial F}{\partial y}$. Montrer que $P_1=(-4,-1)$ et $P_2=(2,1)$ sont des points de $\mathscr{C}$ de tangence verticale. (Bonus : montrer que ce sont les seuls.) Pour (x_0,y_0) un point de $\mathscr{C}$ autre que ces deux-là, énoncer le théorème des fonctions implicites. Montrer que $Q=(1,\frac{3-\sqrt{5}}{2}) \in \mathscr{C}$ et calculer une équation de la tangente à $\mathscr{C}$ en ce point.
2. Soit $F(x,y)=xy^2-x^2+y^2-2x+3$ et $\mathscr{C} : (F(x,y)=0)$. Calculer $\frac{\partial F}{\partial x}$ et $\frac{\partial F}{\partial y}$. Trouver les points

P_1 et P_2 de $\mathscr{C}$ de tangence verticale. Pour (x_0, y_0) un point de $\mathscr{C}$ autre que ces deux-là, énoncer le théorème des fonctions implicites. Trouver les deux points Q_1 et Q_2 de $\mathscr{C}$ ayant pour abscisse $x = 3$. Calculer une équation de la tangente à $\mathscr{C}$ en chacun de ces points.

3. Soit $F(x, y, z) = \sqrt{x^2 + y^2 + z^2} - \cos(y)$. Soit $(x_0, y_0, z_0) = (0, 0, 1)$. Montrer que (x_0, y_0, z_0) est une solution de $F(x, y, z) = 0$. Calculer $\frac{\partial F}{\partial z}(x, y, z)$ et sa valeur en (x_0, y_0, z_0). Montrer que pour (x, y) proche de $(x_0, y_0) = (0, 0)$, il existe une fonction φ telle que en posant $z = \varphi(x, y)$, on ait (x, y, z) est une solution de $F(x, y, z) = 0$.

Auteurs du chapitre

Arnaud Bodin. D'après des cours de Abdellah Hanani (Lille), Goulwen Fichou et Stéphane Leborgne (Rennes), Laurent Pujo-Menjouet (Lyon). Relu par Anne Bauval et [...].

Chapitre 6

Extremums

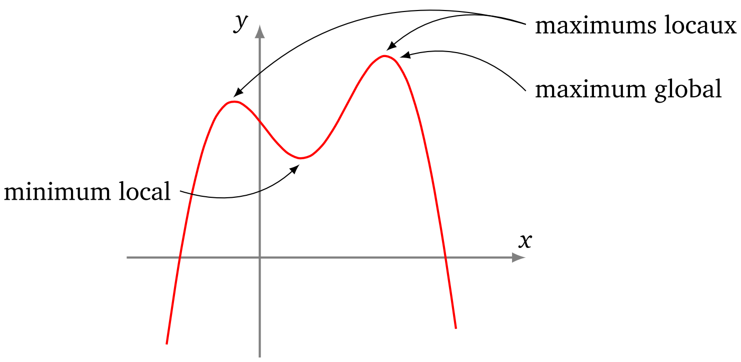
1. Motivation – Cas d'une variable

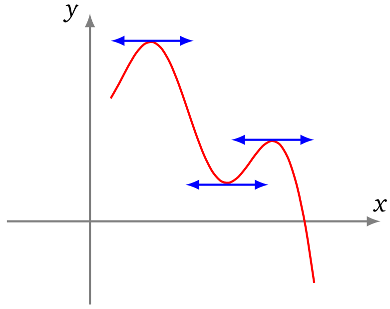
Comment trouver le maximum (ou le minimum) d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$? Ce chapitre est consacré à l'étude de l'existence des extremums. Nous apprendrons à repérer les extremums locaux (qui ne sont pas nécessairement des minimums ou maximums globaux). On terminera ce chapitre par l'étude des « extremums liés », où la recherche est limitée par une contrainte. Pour mieux comprendre ce qui se passe en plusieurs variables, on commence par revoir rapidement le cas d'une variable.

1.1. Minimum, maximum, point critique

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction d'une variable.

- f admet un ***maximum local*** en $x_0 \in \mathbb{R}$ s'il existe un intervalle ouvert I contenant x_0 tel que :
$$\text{pour tout } x \in I \qquad f(x) \leqslant f(x_0).$$
- f admet un ***minimum local*** en $x_0 \in \mathbb{R}$ s'il existe un intervalle ouvert I contenant x_0 tel que :
$$\text{pour tout } x \in I \qquad f(x) \geqslant f(x_0).$$
- f admet un ***extremum local*** en $x_0 \in \mathbb{R}$ si f admet un maximum local ou bien un minimum local en ce point.
- f admet un ***point critique*** en $x_0 \in \mathbb{R}$ si $f'(x_0) = 0$. Géométriquement, c'est un point de tangente horizontale.
- Proposition : si f dérivable admet un minimum local ou un maximum local en x_0, alors $f'(x_0) = 0$. Autrement dit, si x_0 est un extremum local alors c'est un point critique.
- La réciproque n'est pas toujours vraie. Par exemple, pour $f : x \mapsto x^3$, le point $x_0 = 0$ est un point critique, mais ce n'est ni un maximum local ni un minimum local (c'est un ***point d'inflexion***).

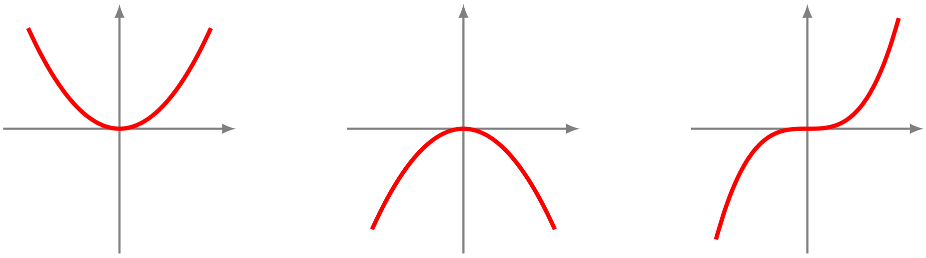




Sur la figure de gauche : des exemples de minimum local, maximum local, maximum global ; il n'y a pas de minimum global sur $\mathbb{R}$. Sur la figure de droite : un extremum local est nécessairement un point critique.

1.2. Exemples fondamentaux

- $f : x \mapsto x^2$, minimum local en 0, on a $f'(0) = 0$ et $f''(0) > 0$.
- $f : x \mapsto -x^2$, maximum local en 0, on a $f'(0) = 0$ et $f''(0) < 0$.
- $f : x \mapsto x^3$, ni minimum ni maximum local en 0, on a $f'(0) = 0$ et $f''(0) = 0$.



$f(x) = x^2$ $f(x) = -x^2$ $f(x) = x^3$

1.3. Formule de Taylor à l'ordre 2

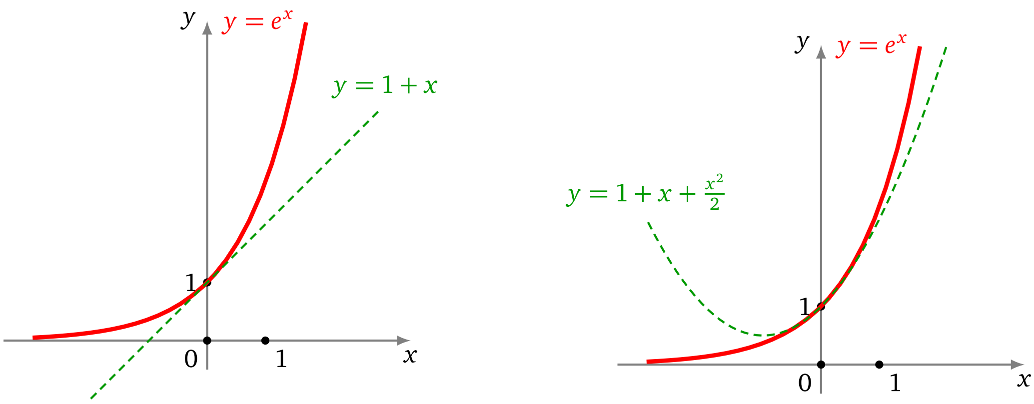
Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction d'une variable de classe $\mathscr{C}^2$.

Théorème 1 (Formule de Taylor à l'ordre 2)**.**
Pour tout $x_0 \in \mathbb{R}$, on a

$$\boxed{f(x_0+h) = f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0) + h^2\epsilon(h)}$$

où $\epsilon(h) \to 0$ lorsque $h \to 0$.

Le développement limité à l'ordre 1, $f(x_0+h) \simeq f(x_0) + hf'(x_0)$, correspond à l'approximation du graphe de f par sa tangente en x_0 (figure de gauche ci-dessous). Le développement limité à l'ordre 2, $f(x_0+h) \simeq f(x_0) + hf'(x_0) + \frac{h^2}{2}f''(x_0)$, correspond à une approximation par une parabole (figure de droite).



Choisissons pour x_0 une valeur telle que $f'(x_0) = 0$ et $f''(x_0) \neq 0$. Alors, pour h assez petit, le terme $\frac{h^2}{2}f''(x_0) + h^2\epsilon(h)$ est du même signe que $f''(x_0)$. Si par exemple $f''(x_0) > 0$, on en déduit que $f(x_0+h) \geqslant f(x_0)$ (pour h proche de 0) et donc que f admet un minimum local en x_0. Ci-dessous, on va en déduire une caractérisation des minimums et maximums.

1.4. Caractérisation des minimums et maximums

La recherche pratique des extremums locaux pour une fonction d'une variable se passe donc ainsi :

1. On recherche les points critiques donnés par l'équation $f'(x) = 0$.
2. Pour chaque point critique x_0, on calcule la dérivée seconde :
   - si $f''(x_0) > 0$, alors f admet un minimum local en x_0,
   - si $f''(x_0) < 0$, alors f admet un maximum local en x_0,
   - si $f''(x_0) = 0$, alors il faut approfondir l'étude.

Lorsque $f : [a,b] \to \mathbb{R}$ est définie sur un intervalle compact, il faudra en plus étudier le comportement de f en a et en b (c'est-à-dire au bord du domaine de définition). Comme l'ensemble de départ est compact, on a la garantie de l'existence d'extremums globaux.

2. Dérivées partielles d'ordre 2

2.1. Dérivées partielles d'ordre 2

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une application différentiable. Les deux dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont aussi des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$; supposons que ce soient aussi des applications différentiables. Alors on peut calculer les dérivées partielles de $\frac{\partial f}{\partial x}$:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \qquad \text{et} \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right).$$

On peut aussi calculer les dérivées partielles de $\frac{\partial f}{\partial y}$:

$$\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \qquad \text{et} \qquad \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right).$$

On note ces dérivées partielles :

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

Ce sont des fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$.
Plus généralement, pour $f : \mathbb{R}^n \to \mathbb{R}$, on note $\frac{\partial f}{\partial x_i} : \mathbb{R}^n \to \mathbb{R}$ les dérivées partielles d'ordre 1 ($1 \leqslant i \leqslant n$) et $\frac{\partial^2 f}{\partial x_j \partial x_i}$ les dérivées partielles d'ordre 2 ($1 \leqslant i, j \leqslant n$).

2.2. Théorème de Schwarz

Pour $f : \mathbb{R}^2 \to \mathbb{R}$, il y a quatre dérivées partielles secondes à calculer, mais en général deux d'entre elles sont égales.

Exemple 1.
Soit $f : U \to \mathbb{R}$ définie par $f(x,y) = x^2\cos(y) + \ln(x - y^2)$ sur $U = \left\{(x,y) \in \mathbb{R}^2 \mid x - y^2 > 0\right\}$. Alors :

$$\frac{\partial f}{\partial x}(x,y) = 2x\cos(y) + \frac{1}{x - y^2} \qquad \frac{\partial f}{\partial y}(x,y) = -x^2\sin(y) - \frac{2y}{x - y^2}$$

On peut maintenant dériver une nouvelle fois pour obtenir les dérivées partielles d'ordre 2 :

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}(x,y)\right) = \frac{\partial}{\partial x}\left(2x\cos(y) + \frac{1}{x - y^2}\right) = 2\cos(y) - \frac{1}{(x - y^2)^2}$$

$$\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(x,y)\right) = \frac{\partial}{\partial y}\left(2x\cos(y) + \frac{1}{x - y^2}\right) = \boxed{-2x\sin(y) + \frac{2y}{(x - y^2)^2}}$$

$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x,y)\right) = \frac{\partial}{\partial x}\left(-x^2\sin(y) - \frac{2y}{x-y^2}\right) = \boxed{-2x\sin(y) + \frac{2y}{(x-y^2)^2}}$$

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

On note sur l'exemple précédent que $\frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{\partial^2 f}{\partial x \partial y}(x,y)$. C'est un phénomène général que l'on va détailler.

Définition 1.
Une fonction $f : \mathbb{R}^n \to \mathbb{R}$ est de ***classe*** $\mathscr{C}^2$ si f est de classe $\mathscr{C}^1$ (c'est-à-dire ses dérivées partielles existent et sont continues) et si ses dérivées partielles sont aussi de classe $\mathscr{C}^1$.

Le théorème de Schwarz dit que le résultat ne dépend pas de l'ordre dans lequel on effectue les dérivations.

Théorème 2 (Théorème de Schwarz).
Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$. Pour tous $i, j \in \{1, \dots, n\}$, on a :

$$\boxed{\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)}$$

Ainsi, pour $f : \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathscr{C}^2$, on a :

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

Pour $f : \mathbb{R}^3 \to \mathbb{R}$ de classe $\mathscr{C}^2$, il y a 9 dérivées partielles d'ordre 2, mais seulement 6 calculs à faire :

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

Le contre-exemple suivant, qui peut être omis lors d'une première lecture, prouve qu'il est nécessaire d'avoir une fonction de classe $\mathscr{C}^2$. Si cette hypothèse manque alors les dérivées partielles croisées peuvent ne pas être égales.

Exemple 2.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

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

On vérifie que f est de classe $\mathscr{C}^1$ sur $\mathbb{R}^2$ et que

$$\frac{\partial f}{\partial x}(x,y) = \begin{cases} \dfrac{y^5 - x^2y^3}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

et

$$\frac{\partial f}{\partial y}(x,y) = \begin{cases} \dfrac{3x^3y^2 + xy^4}{(x^2+y^2)^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0). \end{cases}$$

Le taux d'accroissement

$$\frac{\frac{\partial f}{\partial x}(0,y) - \frac{\partial f}{\partial x}(0,0)}{y-0} = 1 \xrightarrow[y \to 0]{} 1$$

ce qui montre que $\dfrac{\partial^2 f}{\partial y \partial x}(0,0) = 1$. De même, le taux d'accroissement

$$\frac{\frac{\partial f}{\partial y}(x,0) - \frac{\partial f}{\partial y}(0,0)}{x-0} = 0 \xrightarrow[x \to 0]{} 0$$

ce qui montre que $\frac{\partial^2 f}{\partial x \partial y}(0,0) = 0$. Les dérivées partielles croisées ne sont pas égales en $(0,0)$. On en déduit que l'une (au moins) des dérivées partielles secondes $\frac{\partial^2 f}{\partial x \partial y}$ ou $\frac{\partial^2 f}{\partial y \partial x}$ n'est pas continue en $(0,0)$. Autrement dit, la fonction f n'est pas de classe $\mathscr{C}^2$ en $(0,0)$ et le théorème de Schwarz ne s'applique pas.

2.3. Hessienne

La matrice jacobienne est la matrice des dérivées partielles. La matrice hessienne est la matrice des dérivées partielles d'ordre 2.
Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction de n variables. La ***matrice hessienne*** de f en $x = (x_1, \dots, x_n)$ est la matrice $n \times n$:

$$\boxed{H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{1 \leqslant i,j \leqslant n}}$$

Pour une fonction de classe $\mathscr{C}^2$, d'après le théorème de Schwarz, c'est une ***matrice symétrique***.
Dans le cas d'une fonction de deux variables :

$$\boxed{H_f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{pmatrix}}$$

Pour trois variables, la matrice hessienne (à évaluer en (x,y,z)) vaut :

$$H_f = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z^2} \end{pmatrix}.$$

Exemple 3.
Calculons la matrice hessienne de $f(x,y) = xy^2 + x^4 - y^4$.
On calcule d'abord

$$\frac{\partial f}{\partial x}(x,y) = 4x^3 + y^2 \qquad \frac{\partial f}{\partial y}(x,y) = 2xy - 4y^3.$$

On a donc

$$H_f(x,y) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{pmatrix} = \begin{pmatrix} 12x^2 & 2y \\ 2y & 2x - 12y^2 \end{pmatrix}.$$

Mini-exercices.

1. Soit $f(x,y) = x^3 + 5x^2y - y^2$. Calculer les dérivées partielles d'ordre 1 de f. Calculer les dérivées partielles d'ordre 2 de f. Vérifier la validité du théorème de Schwarz. Calculer la matrice hessienne de f. Calculer les dérivées partielles d'ordre 3 de f.
2. Soit $f(x,y) = xe^{x^2-y^2}$. Calculer les dérivées partielles d'ordre 1 et d'ordre 2 de f. Calculer la matrice hessienne de f.
3. Soit $f(x,y,z) = xy^2 \ln(z)$. Déterminer l'ensemble de définition de f. Calculer les dérivées partielles d'ordre 1 et d'ordre 2 ainsi que la matrice hessienne de f.

3. Formule de Taylor à l'ordre 2

3.1. Formule de Taylor à l'ordre 1

Nous avons déjà vu qu'une façon de dire que f est différentiable est de dire que f admet un ***développement limité à l'ordre 1***. Pour $f : \mathbb{R}^2 \to \mathbb{R}$, au point (x_0, y_0), si f est différentiable alors

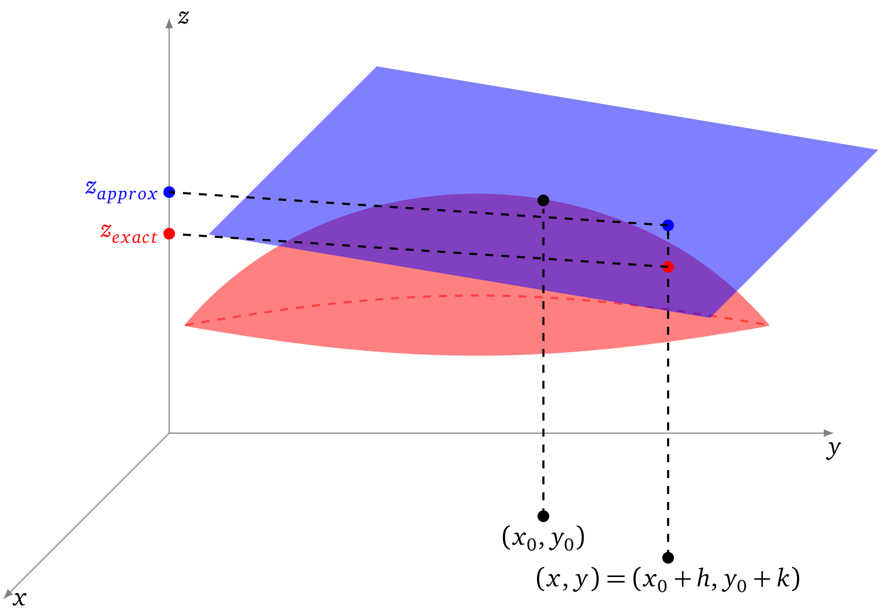
$$f(x_0+h, y_0+k) = f(x_0, y_0) + h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0) + o(\|(h,k)\|).$$

Connaissant les valeurs de f, $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ uniquement en (x_0, y_0), on en déduit une approximation de f en tout (x, y) proche de (x_0, y_0).

Le but va être d'améliorer cette approximation par un développement limité à l'ordre 2.
On rappelle la notation « petit o ».
Notation. Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction définie au voisinage de $(0,0)$. On dit que g est ***négligeable*** par rapport à $\|(x, y)\|^n$ au voisinage de $(0,0)$ et on note $g = o(\|(x, y)\|^n)$ si

$$\lim_{(x,y)\to(0,0)} \frac{g(x,y)}{\|(x,y)\|^n} = 0.$$

Interprétation géométrique.
Le plan tangent (en bleu) au graphe de f au point (x_0, y_0) est le plan qui approche le mieux le graphe de f (en rouge) autour de ce point. Pour calculer $f(x, y)$ lorsque $(x, y) = (x_0 + h, y_0 + k)$ est proche de (x_0, y_0), on remplace la valeur exacte $z_{exact} = f(x, y)$ par son approximation $z_{approx} = f(x_0, y_0) + h\frac{\partial f}{\partial x}(x_0, y_0) + k\frac{\partial f}{\partial y}(x_0, y_0)$. Le point (x, y, z_{exact}) est un point du graphe de f alors que (x, y, z_{approx}) est un point du plan tangent au graphe de f au point (x_0, y_0).



3.2. Formule de Taylor à l'ordre 2 (en 2 variables)

Théorème 3.
Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$ sur un ouvert $U \subset \mathbb{R}^2$ et soit $(x_0, y_0) \in U$. Alors :

$$\boxed{\begin{aligned} f(x_0+h,y_0+k) = f(x_0,y_0) \ &+\ h\frac{\partial f}{\partial x}(x_0,y_0)+k\frac{\partial f}{\partial y}(x_0,y_0) \\ &+\frac{1}{2}\left[h^2\frac{\partial^2 f}{\partial x^2}(x_0,y_0)+2hk\frac{\partial^2 f}{\partial x\partial y}(x_0,y_0)+k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right] \\ &+o\left(\|(h,k)\|^2\right) \end{aligned}}$$

- On dit aussi que f admet un développement limité d'ordre 2 au point (x_0,y_0).
- Attention à ne pas oublier le facteur $\frac{1}{2}$ devant les termes de degré 2, ni le facteur 2 devant le terme hk.
- On rappelle que si on choisit la norme euclidienne alors $\|(h,k)\|^2 = h^2+k^2$ et que $o\left(\|(h,k)\|^2\right)$ désigne une fonction égale à $\|(h,k)\|^2\epsilon(h,k)$ avec $\epsilon(h,k) \xrightarrow[(h,k)\to(0,0)]{} 0$.

Exemple 4.
Étudions $f(x,y)=\exp(xy^2-2)$ autour de $(x_0,y_0)=(2,1)$.
- $f(2,1)=1$.
- **DL à l'ordre 1.**

$$\frac{\partial f}{\partial x}(x,y)=y^2\exp(xy^2-2) \quad \frac{\partial f}{\partial y}(x,y)=2xy\exp(xy^2-2) \quad \text{donc} \quad \frac{\partial f}{\partial x}(2,1)=1 \quad \frac{\partial f}{\partial y}(2,1)=4$$

  Ainsi :

$$f(2+h,1+k)=1+h+4k+o\left(\|(h,k)\|\right).$$

- **DL à l'ordre 2.**

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x,y) &= y^4\exp(xy^2-2) \\ \frac{\partial^2 f}{\partial x\partial y}(x,y) &= (2y+2xy^3)\exp(xy^2-2) \\ \frac{\partial^2 f}{\partial y^2}(x,y) &= (2x+4x^2y^2)\exp(xy^2-2) \end{aligned}$$

  Donc :

$$\frac{\partial^2 f}{\partial x^2}(2,1)=1 \qquad \frac{\partial^2 f}{\partial x\partial y}(2,1)=6 \qquad \frac{\partial^2 f}{\partial y^2}(2,1)=20.$$

  Ainsi :

$$f(2+h,1+k) \ = \ 1 \quad + \quad h+4k \quad + \quad \frac{1}{2}\left(h^2+2\cdot 6\cdot hk+20k^2\right) \quad + \quad o\left(\|(h,k)\|^2\right).$$

Interprétation géométrique. Un développement limité à l'ordre 1 correspond à approcher le graphe de f par un plan tangent d'équation $z=a+bx+cy$. Un développement limité à l'ordre 2 correspond à une approximation par une surface quadratique, c'est-à-dire une surface définie par une équation de degré 2 :

$$z=a+bx+cy+dx^2+2exy+fy^2.$$

Cette approximation est meilleure mais reste valable uniquement autour du point (x_0,y_0) considéré.

Approximations numériques.
Si (x,y) est proche de (x_0,y_0) (c'est-à-dire si h et k sont petits), alors on remplace le calcul de $f(x,y)$ qui peut être compliqué par une bonne approximation donnée par le DL à l'ordre 1, ou mieux, le DL à l'ordre 2.

Exemple 5.
Sans calculatrice, évaluons $\sqrt{1+x+xy}$ avec $x=0.1$ et $y=0.2$.
Soit $f(x,y)=\sqrt{1+x+xy}=(1+x+xy)^{\frac{1}{2}}$ que l'on étudie autour de $(x_0,y_0)=(0,0)$. Alors :
- $f(0,0)=1$.

- **DL à l'ordre 1.**

$$\frac{\partial f}{\partial x}(x,y) = \frac{1+y}{2}(1+x+xy)^{-\frac{1}{2}} \qquad \frac{\partial f}{\partial y}(x,y) = \frac{x}{2}(1+x+xy)^{-\frac{1}{2}}$$

donc

$$\frac{\partial f}{\partial x}(0,0) = \frac{1}{2} \qquad \frac{\partial f}{\partial y}(0,0) = 0.$$

Ainsi, pour (x,y) proche de $(0,0)$, on a

$$f(x,y) \simeq 1 + \frac{1}{2}x.$$

- **DL à l'ordre 2.**

$$\begin{aligned}
\frac{\partial^2 f}{\partial x^2}(x,y) &= -\frac{(1+y)^2}{4}(1+x+xy)^{-\frac{3}{2}} \\
\frac{\partial^2 f}{\partial x \partial y}(x,y) &= \frac{2+x+xy}{4}(1+x+xy)^{-\frac{3}{2}} \\
\frac{\partial^2 f}{\partial y^2}(x,y) &= -\frac{x^2}{4}(1+x+xy)^{-\frac{3}{2}}
\end{aligned}$$

donc

$$\frac{\partial^2 f}{\partial x^2}(0,0) = -\frac{1}{4} \qquad \frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{1}{2} \qquad \frac{\partial^2 f}{\partial y^2}(0,0) = 0.$$

Ainsi, pour (x,y) proche de $(0,0)$, on a

$$f(x,y) \simeq \ 1 \ + \ \frac{1}{2}x \ + \ \frac{1}{2}\left(-\frac{1}{4}x^2 + xy\right).$$

- **Valeurs numériques.** Avec $x = 0.1$ et $y = 0.2$, la valeur exacte est $f(x,y) = \sqrt{1.12} = 1.05830\ldots$ Le DL à l'ordre 1 fournit l'approximation

$$f(x,y) \simeq 1 + \frac{1}{2}x = 1.05000,$$

alors que le DL à l'ordre 2 donne

$$f(x,y) \simeq 1 + \frac{1}{2}x + \frac{1}{2}\left(-\frac{1}{4}x^2 + xy\right) = 1.05875.$$

3.3. Formule de Taylor à l'ordre 2 (en n variables)

Exprimons d'abord la formule de Taylor en deux variables à l'aide de vecteurs et matrices. La matrice jacobienne est ici une matrice ligne et

$$J_f(x_0,y_0) \times \begin{pmatrix} h \\ k \end{pmatrix} = h\frac{\partial f}{\partial x}(x_0,y_0) + k\frac{\partial f}{\partial y}(x_0,y_0).$$

Remarque : on pourrait aussi utiliser le gradient de f qui est un vecteur colonne :

$$\mathrm{d}f(x_0,y_0)(h,k) = J_f(x_0,y_0) \times \begin{pmatrix} h \\ k \end{pmatrix} = \langle \operatorname{grad} f(x_0,y_0) \mid \begin{pmatrix} h \\ k \end{pmatrix} \rangle$$

On vérifie que les termes de degré 2 s'expriment à l'aide de la hessienne :

$$\frac{1}{2}(h,k)\, H_f(x_0,y_0) \begin{pmatrix} h \\ k \end{pmatrix} = \frac{1}{2}\left[h^2 \frac{\partial^2 f}{\partial x^2}(x_0,y_0) + 2hk\frac{\partial^2 f}{\partial x \partial y}(x_0,y_0) + k^2\frac{\partial^2 f}{\partial y^2}(x_0,y_0)\right]$$

où $(h,k) = \left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right)^T$ est un vecteur ligne.

De façon plus générale, pour $f : \mathbb{R}^n \to \mathbb{R}$, si on note $x_0 = (x_1,\ldots,x_n)$ et $h = (h_1,\ldots,h_n)$ (considérés comme des vecteurs colonnes), on a :

$$\boxed{f(x_0+h) = f(x_0) + J_f(x_0)h + \frac{1}{2}h^T H_f(x_0)h + o\left(\|h\|^2\right)}$$

où h^T désigne le vecteur ligne obtenu par transposition du vecteur colonne h.
On rencontre aussi fréquemment cette formule sous la forme :

$$f(x) = f(x_0) + J_f(x_0)(x-x_0) + \frac{1}{2}(x-x_0)^T H_f(x_0)(x-x_0) + o\left(\|x-x_0\|^2\right)$$

où l'on a effectué le changement de variable $x = x_0 + h$ (et donc $h = x - x_0$).

Exemple 6.
Soit $f(x,y,z) = ye^{x^2} + xz^2$. Calculons le DL à l'ordre 2 en $(x_0,y_0,z_0) = (1,2,3)$ de f, pour lequel $f(x_0,y_0,z_0) = 2e+9$.

$$J_f(x,y,z) = \left(2xye^{x^2} + z^2, \;\; e^{x^2}, \;\; 2xz\right) \quad \text{donc} \quad J_f(x_0,y_0,z_0) = (4e+9, \;\; e, \;\; 6)$$

$$H_f(x,y,z) = \begin{pmatrix} (4x^2y+2y)e^{x^2} & 2xe^{x^2} & 2z \\ 2xe^{x^2} & 0 & 0 \\ 2z & 0 & 2x \end{pmatrix} \quad \text{donc} \quad H_f(x_0,y_0,z_0) = \begin{pmatrix} 12e & 2e & 6 \\ 2e & 0 & 0 \\ 6 & 0 & 2 \end{pmatrix}$$

Ainsi :

$$\begin{aligned} f(x_0+h_x, y_0+h_y, z_0+h_z) = 2e+9 & \\ +(4e+9)h_x + eh_y + 6h_z & \\ +\frac{1}{2}\left[12eh_x^2 + 4eh_xh_y + 12h_xh_z + 2h_z^2\right] & \\ +o(h_x^2+h_y^2+h_z^2) & \end{aligned}$$

Mini-exercices.

1. Soit $f(x,y) = x + 2x^2 + 3xy + 4y^2$. Calculer le DL à l'ordre 2 de f en $(0,0)$, puis le DL à l'ordre 2 de f en $(-1,3)$.
2. Soit $f(x,y) = x^3 + 2xy - y^2$. Calculer le développement limité de f à l'ordre 1 en $(x_0,y_0) = (1,2)$. En déduire l'équation du plan tangent au graphe de f en (x_0,y_0). Calculer le développement limité de f à l'ordre 2 en (x_0,y_0). En déduire l'équation de la surface quadratique approchant au mieux le graphe de f en (x_0,y_0).
3. On étudie $f(x,y) = \sqrt{x^2y}$ autour de $(x_0,y_0) = (2,4)$. Calculer le développement limité de f à l'ordre 1, et en déduire une valeur approchée de $f(1.99,4.02)$. Calculer le développement limité de f à l'ordre 2, et en déduire une meilleure valeur approchée de $f(1.99,4.02)$. Comparer avec la valeur exacte.

4. Différentielle seconde

Cette section est beaucoup plus théorique et doit être passée lors d'une première lecture.

4.1. Différentielle seconde

Soit $f : U \to \mathbb{R}$ une fonction f définie sur un ouvert $U \subset \mathbb{R}^n$.

Définition 2.
f est dite *deux fois différentiable en $x_0 \in U$* :

- si elle est différentiable dans un voisinage ouvert V_{x_0} de x_0,
- et si sa différentielle $\mathrm{d}f : V_{x_0} \to \mathscr{L}(\mathbb{R}^n, \mathbb{R})$ est différentiable en x_0.

On dit que f est *deux fois différentiable sur U* si elle est différentiable en tout point de U.

On note $\mathscr{L}(\mathbb{R}^n,\mathbb{R})$ l'ensemble des applications linéaires de $\mathbb{R}^n$ vers $\mathbb{R}$. Plus généralement, $\mathscr{L}(\mathbb{R}^n,\mathbb{R}^p)$ désigne l'ensemble des applications linéaires de $\mathbb{R}^n$ vers $\mathbb{R}^p$.
Par sa définition, la différentielle de $\mathrm{d}f$ en x, que l'on écrit $\mathrm{d}(\mathrm{d}f)(x)$, est une application linéaire de $\mathbb{R}^n$ dans $\mathscr{L}(\mathbb{R}^n,\mathbb{R})$. Autrement dit, on a

$$\mathrm{d}f : U \to \mathscr{L}(\mathbb{R}^n,\mathbb{R}) \qquad \text{et} \qquad \mathrm{d}(\mathrm{d}f) : U \to \mathscr{L}(\mathbb{R}^n,\mathscr{L}(\mathbb{R}^n,\mathbb{R})).$$

Mais elle s'identifie naturellement avec une application bilinéaire de $\mathbb{R}^n \times \mathbb{R}^n$ dans $\mathbb{R}$ grâce à la proposition suivante :

$$\mathscr{L}(\mathbb{R}^n,\mathscr{L}(\mathbb{R}^n,\mathbb{R})) \simeq \mathscr{L}(\mathbb{R}^n \times \mathbb{R}^n,\mathbb{R})$$

L'identification se fait ainsi : si $L \in \mathscr{L}(\mathbb{R}^n,\mathscr{L}(\mathbb{R}^n,\mathbb{R}))$ alors, pour $h \in \mathbb{R}^n$, $L(h) \in \mathscr{L}(\mathbb{R}^n,\mathbb{R})$, $k \in \mathbb{R}^n \mapsto L(h)(k) \in \mathbb{R}$. L'application $(h,k) \mapsto L(h)(k)$ (que l'on note $L(h,k)$) est une application linéaire en h et en k, donc bilinéaire en (h,k).

Définition 3.
La ***différentielle seconde*** d'une fonction $f : U \subset \mathbb{R}^n \to \mathbb{R}$ deux fois différentiable est l'application

$$\begin{array}{rccl} \mathrm{d}^2 f : & U & \to & \mathscr{L}(\mathbb{R}^n \times \mathbb{R}^n,\mathbb{R}) \\ & x & \mapsto & \mathrm{d}^2 f(x) \end{array}$$

définie par

$$\mathrm{d}^2 f(x)(h,k) = \mathrm{d}(\mathrm{d}f)(x)(h)(k) \qquad \text{pour tout } (h,k) \in \mathbb{R}^n \times \mathbb{R}^n.$$

Donnons ici les différentielles d'ordre 2 pour deux types de fonctions classiques : les applications affines et les applications quadratiques.

- Une application affine $f : x \mapsto \ell(x) + b$ avec $\ell \in \mathscr{L}(\mathbb{R}^n,\mathbb{R})$ et $b \in \mathbb{R}$ est deux fois différentiable et sa différentielle seconde est identiquement nulle. C'est la version étendue du fait que la dérivée seconde de la fonction d'une variable $x \mapsto ax + b$ est nulle.
- Une application quadratique $f : x \mapsto \phi(x,x)$ avec $\phi \in \mathscr{L}(\mathbb{R}^n \times \mathbb{R}^n,\mathbb{R})$ est deux fois différentiable et sa différentielle seconde est constante, et même égale à 2ϕ si ϕ est symétrique. C'est la version étendue du fait que la dérivée seconde de la fonction d'une variable $x \mapsto ax^2 + bx + c$ est constante.

4.2. Théorème de Schwarz – Hessienne

Le théorème de Schwarz sur l'égalité des dérivées croisées se reformule ainsi.

Théorème 4 (Théorème de Schwarz)**.**
Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$. Alors, pour tout $x \in U$, $\mathrm{d}^2 f(x)$ est une application bilinéaire **symétrique**. *Autrement dit, pour tout $(h,k) \in \mathbb{R}^n \times \mathbb{R}^n$, on a*

$$\mathrm{d}^2 f(x)(h,k) = \mathrm{d}^2 f(x)(k,h).$$

Soit $f : U \to \mathbb{R}$ une fonction deux fois différentiable sur un ouvert $U \subset \mathbb{R}^n$. Soit $(e_1,\ldots,e_n)$ la base canonique de $\mathbb{R}^n$. Alors, pour tout $x \in U$, pour tous $i,j \in \{1,\ldots,n\}$, on a

$$\mathrm{d}^2 f(x)(e_i,e_j) = \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

On rappelle que la matrice hessienne de f en x est la matrice des dérivées partielles secondes :

$$H_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x) \right)_{1 \leqslant i,j \leqslant n}$$

Par bilinéarité, si h et k sont deux vecteurs de $\mathbb{R}^n$ de coordonnées respectives $(h_1,\ldots,h_n)$ et $(k_1,\ldots,k_n)$ (vus comme des vecteurs colonnes), alors

$$\mathrm{d}^2 f(x)(h,k) = h^T \cdot H_f(x) \cdot k = \sum_{i=1}^{n} \sum_{j=1}^{n} h_i k_j \frac{\partial^2 f}{\partial x_i \partial x_j}(x).$$

Autrement dit, $H_f(x)$ est la matrice de la forme bilinéaire $\mathrm{d}^2 f(x)$ par rapport à la base canonique de $\mathbb{R}^n$. Le théorème de Schwarz assure de plus que la matrice hessienne est symétrique si f est de classe $\mathscr{C}^2$ sur U.

4.3. Formule de Taylor

Soit $f : U \to \mathbb{R}$ une fonction différentiable sur un ouvert $U \subset \mathbb{R}^n$. La formule de Taylor à l'ordre 1 est :

$$f(x+h) = f(x) + \mathrm{d}f(x)(h) + o(\|h\|)$$

car on rappelle que $\mathrm{d}f(x)(h) = J_f(x) \cdot h$.
Pour une fonction f deux fois différentiable, la ***formule de Taylor à l'ordre 2***, dite aussi ***développement limité à l'ordre 2***, est :

$$\boxed{f(x+h) = f(x) + \mathrm{d}f(x)(h) + \frac{1}{2}\mathrm{d}^2 f(x)(h,h) + o\left(\|h\|^2\right)}$$

Plus généralement, en itérant les différentielles :

Théorème 5 (Formule de Taylor à l'ordre p)**.**
Si U est un ouvert de $\mathbb{R}^n$ et si $f : U \to \mathbb{R}$ est une fonction p fois différentiable en $x \in U$ alors

$$f(x+h) = f(x) \ + \sum_{k=1}^{p} \frac{1}{k!}\mathrm{d}^k f(x)(h^{[k]}) \ + \ o(\|h\|^p)$$

où on a noté $h^{[k]} = (h, h, \ldots, h) \in \mathbb{R}^k$.

5. Minimum et maximum : cas de deux variables

5.1. Maximum, minimum

Soit $f : U \to \mathbb{R}$ une fonction de deux variables, où U est un ouvert de $\mathbb{R}^2$.

Définition 4.
On dit que f admet un ***maximum local*** (resp. ***minimum local***) en $(x_0, y_0) \in U$ s'il existe un disque ouvert $D \subset U$, centré en (x_0, y_0), tel que :

$$\forall (x,y) \in D \qquad f(x,y) \leqslant f(x_0, y_0)$$

(resp. $f(x,y) \geqslant f(x_0, y_0)$).
On dit que f admet un ***extremum local*** en (x_0, y_0) si elle y admet un maximum local ou un minimum local.

5.2. Point critique

On suppose f de classe $\mathscr{C}^2$ sur un ouvert U, c'est-à-dire que ses dérivées partielles jusqu'à l'ordre 2 existent et sont continues.

Proposition 1.
Si f admet un extremum local en (x_0, y_0) d'un ouvert U, alors

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

Démonstration. La fonction d'une variable $x \mapsto f(x, y_0)$ admet un extremum local en x_0 sur un ouvert de $\mathbb{R}$, donc sa dérivée, qui est $\frac{\partial f}{\partial x}(x, y_0)$, s'annule en x_0. On fait de même avec $y \mapsto f(x_0, y)$. □

Autrement dit, si f possède un minimum ou maximum local en un point, alors le gradient de f est le vecteur nul en ce point. Les points de U où le gradient de f s'annule sont appelés ***points critiques*** de f. Le résultat précédent dit que les extremums d'une fonction sur un ouvert ne peuvent se produire qu'en un point critique. La réciproque est fausse.

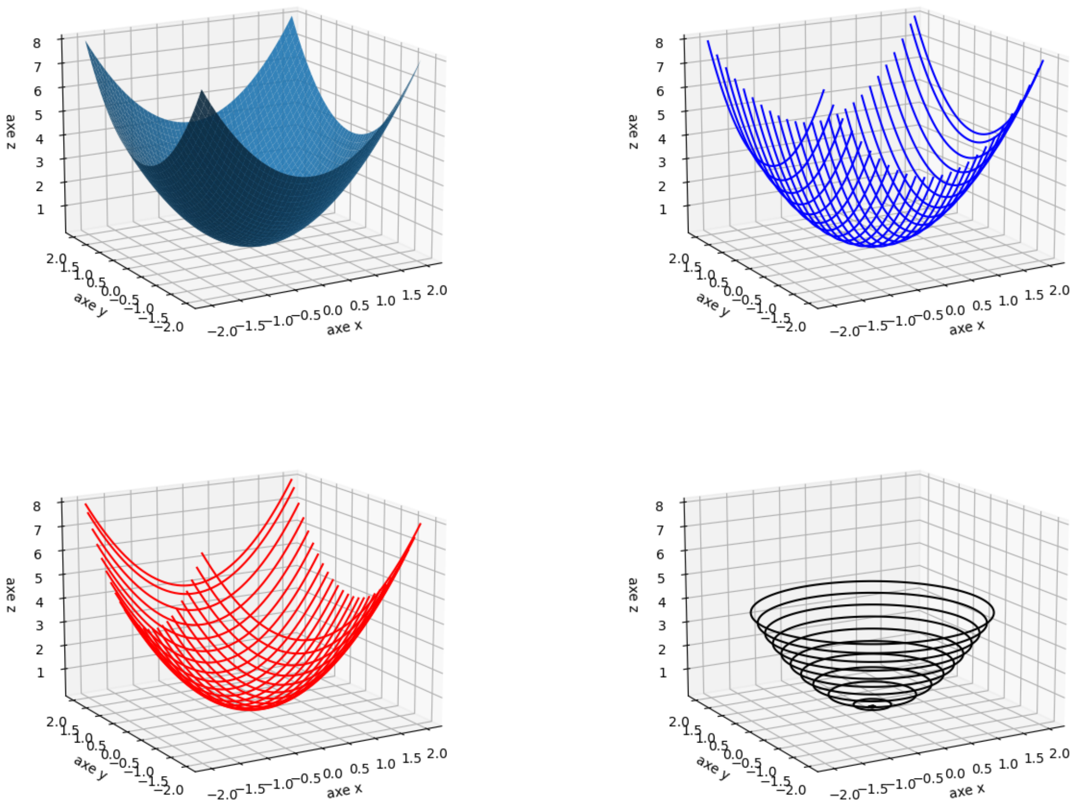
Par définition, un point critique qui n'est ni un maximum local ni un minimum local est nommé ***point-selle*** (ou ***point-col***).

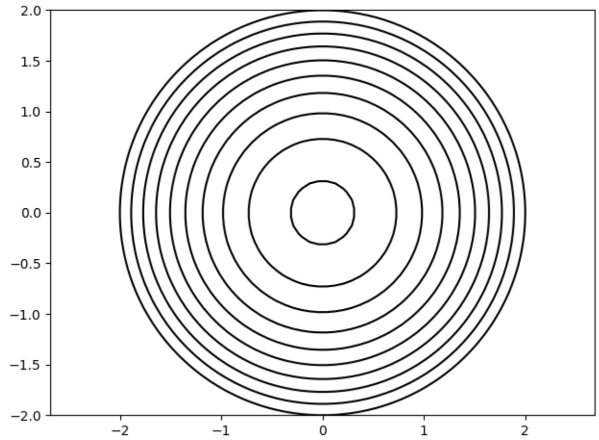
5.3. Exemples fondamentaux

Exemple 7.
$f(x,y) = x^2 + y^2$. C'est un exemple de minimum local atteint en $(0,0)$.

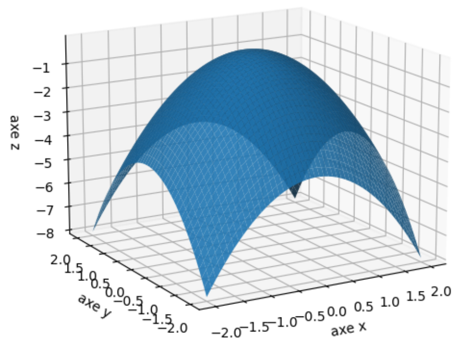
- Les tranches sont des paraboles.
- Les lignes de niveau sont des ellipses (en fait ici des cercles).
- Le graphe est donc un ***paraboloïde elliptique***.

Ci-dessous : (a) la surface, (b) les tranches avec x constant, (c) les tranches avec y constant, (d) les courbes de niveau, (e) les lignes de niveau dans le plan.





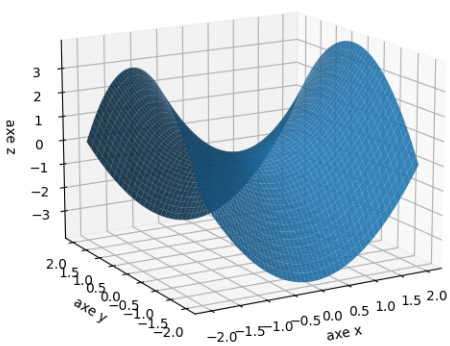
Exemple 8.
$f(x,y) = -x^2 - y^2$. C'est un exemple de maximum local atteint en $(0,0)$.

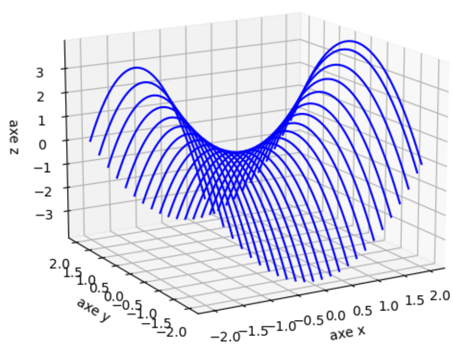


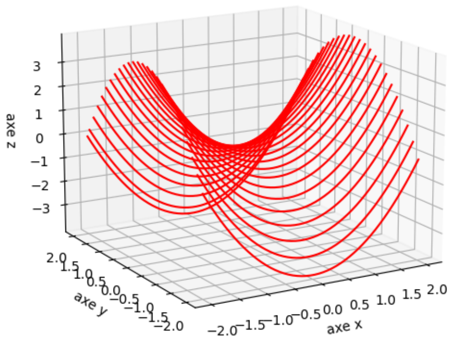
Exemple 9.
$f(x,y) = x^2 - y^2$. C'est un exemple de point-selle en $(0,0)$.

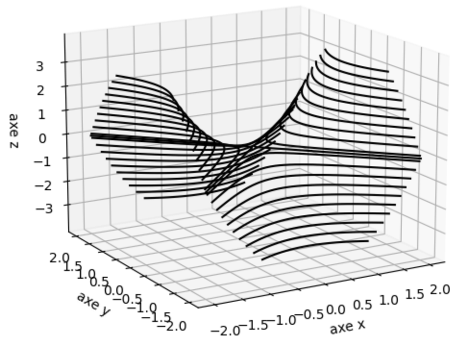
- Les tranches sont des paraboles, tournées vers le haut ou vers le bas selon la direction de la tranche.
- Les lignes de niveau sont des hyperboles.
- Le graphe est donc un ***paraboloïde hyperbolique*** que l'on appelle aussi la ***selle de cheval***.
- Un autre nom pour cette surface est un ***col*** (en référence à un col en montagne). En effet, le point $(0,0,0)$ est le point de passage le moins haut pour passer d'un versant à l'autre de la montagne.

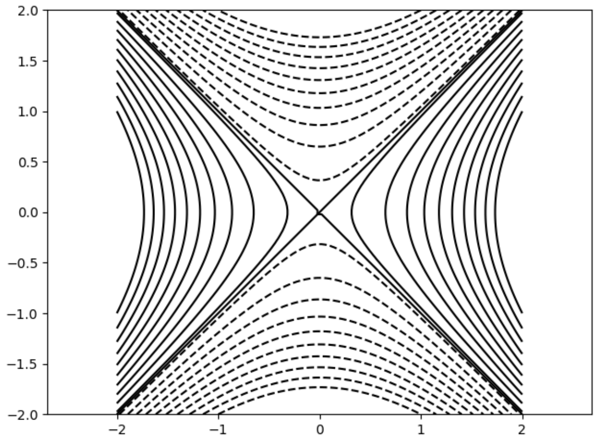
Ci-dessous : (a) la surface, (b) les tranches avec x constant, (c) les tranches avec y constant, (d) les courbes de niveau, (e) les lignes de niveau dans le plan (en pointillé les lignes de niveau négatif).











5.4. Caractérisation des minimums et maximums

Pour une fonction $f : \mathbb{R}^2 \to \mathbb{R}$, nous utiliserons la notation de Monge, qui fournit un critère simple pour détecter un minimum ou un maximum local.

Théorème 6 (Critère de Monge).
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathscr{C}^2$ et soit (x_0, y_0) un point critique de f. On pose

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

Alors :
- *si $rt - s^2 > 0$ et $r > 0$, alors (x_0, y_0) est un minimum local de f,*
- *si $rt - s^2 > 0$ et $r < 0$, alors (x_0, y_0) est un maximum local de f,*
- *si $rt - s^2 < 0$, alors (x_0, y_0) n'est ni un minimum local ni un maximum local : c'est un point-selle,*
- *si $rt - s^2 = 0$, on ne peut pas conclure directement (il faut approfondir l'étude).*

Remarque : $rt - s^2$ est le déterminant de la matrice hessienne en (x_0, y_0), $H_f(x_0, y_0) = \begin{pmatrix} r & s \\ s & t \end{pmatrix}$.

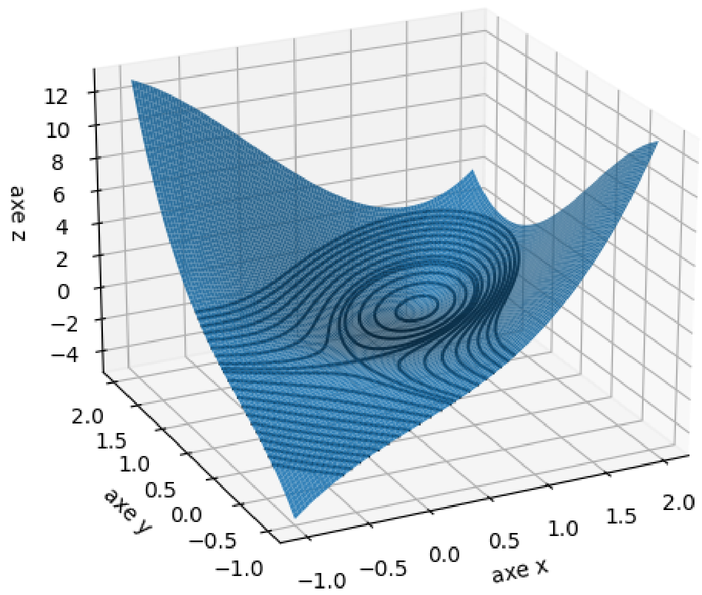
Exemple 10.
Reprenons les trois exemples fondamentaux, pour vérifier notre critère.

1. Soit $f(x, y) = x^2 + y^2$. Le point $(0, 0)$ est l'unique point critique de f. On calcule
$$r = \frac{\partial^2 f}{\partial x^2}(0,0) = 2 \qquad s = \frac{\partial^2 f}{\partial x \partial y}(0,0) = 0 \qquad t = \frac{\partial^2 f}{\partial y^2}(0,0) = 2.$$
Ainsi, $rt - s^2 = 4$ avec $r > 0$, et donc $(0, 0)$ est bien un minimum local de f.
2. Exercice : faire le même travail avec $f(x, y) = -x^2 - y^2$.

3. Soit $f(x,y) = x^2 - y^2$. On trouve un seul point critique : $(0,0)$. On calcule $H_f(0,0) = \left(\begin{smallmatrix} 2 & 0 \\ 0 & -2 \end{smallmatrix}\right)$. Cette fois, $r = 2$, $s = 0$, $t = -2$. Ainsi, $rt - s^2 = -4 < 0$ et donc $(0,0)$ correspond bien à un point-selle.

5.5. Autres exemples

Exemple 11.
Soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par $f(x,y) = x^3 + y^3 - 3xy$.



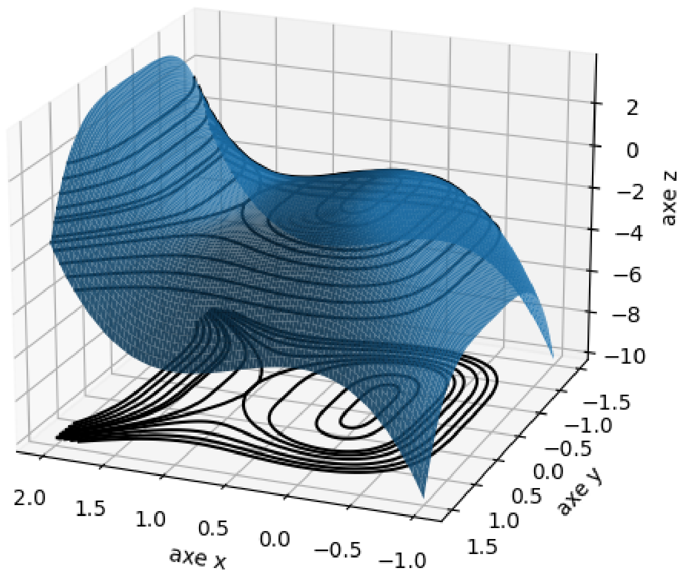
- **Dérivées partielles.**
$$\frac{\partial f}{\partial x}(x,y) = 3x^2 - 3y \qquad \frac{\partial f}{\partial y}(x,y) = 3y^2 - 3x$$
- **Points critiques.** Ce sont les points où $\frac{\partial f}{\partial x}(x,y) = 0$ et $\frac{\partial f}{\partial y}(x,y) = 0$ en même temps. On a donc simultanément $x^2 = y$ et $y^2 = x$ (ce qui implique $x, y \geqslant 0$). D'où $x^4 = y^2 = x$ dont les solutions positives sont $x = 0$ (et alors $y = 0$) et $x = 1$ (et alors $y = 1$). Ainsi, les points critiques sont $(0,0)$ et $(1,1)$.
- **Dérivées partielles secondes.**
$$\frac{\partial^2 f}{\partial x^2}(x,y) = 6x \qquad \frac{\partial^2 f}{\partial x \partial y}(x,y) = -3 \qquad \frac{\partial^2 f}{\partial y^2}(x,y) = 6y$$
- **Étude en** $(0,0)$**.**
$$H_f(0,0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$
C'est-à-dire $r = 0$, $s = -3$, $t = 0$, donc $rt - s^2 = -9 < 0$ et donc $(0,0)$ est un point-selle.
- **Étude en** $(1,1)$**.**
$$H_f(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$
C'est-à-dire $r = 6$, $s = -3$, $t = 6$, donc $rt - s^2 = 27 > 0$ avec $r > 0$ et donc $(1,1)$ est un point de minimum de f (c'est un minimum local et pas global).

Voici un exemple où le critère ne permet pas de conclure. Il faut terminer l'étude à la main.

Exemple 12.
Soit $f(x,y) = 2x^3 - y^4 - 3x^2$. On trouve deux points critiques : $(0,0)$ et $(1,0)$. Par ailleurs :

$$H_f(0,0) = \begin{pmatrix} -6 & 0 \\ 0 & 0 \end{pmatrix} \qquad \text{et} \qquad H_f(1,0) = \begin{pmatrix} 6 & 0 \\ 0 & 0 \end{pmatrix}.$$

On ne peut pas conclure car le déterminant $rt - s^2$ est nul. On étudie chaque cas à la main.



- En $(0,0)$. Écrivons $f(x,y) = x^2(2x-3) - y^4$. Pour $|x| \leqslant 1$, on a $2x - 3 \leqslant 0$, et donc
$$f(x,y) = x^2(2x-3) - y^4 \leqslant 0.$$
Comme $f(0,0) = 0$, alors f admet un maximum local au point $(0,0)$.
- En $(1,0)$. Tout d'abord, on se limite aux points de la forme $(1,y)$ (autour de $y_0 = 0$) :
$$f(1,y) = -1 - y^4 \leqslant -1 = f(1,0)$$
Ensuite, on se limite aux points de la forme $(x,0)$ (autour de $x_0 = 1$, par exemple pour x tel que $|x-1| \leqslant 1$) :
$$f(x,0) = (x-1)^2(2x+1) - 1 \geqslant -1 = f(1,0)$$
Donc, en $(1,0)$, ce n'est ni un minimum ni un maximum : c'est un point-selle.

5.6. Valeurs propres de la hessienne

Ce paragraphe peut être omis à la première lecture.
Nous reformulons le critère précédent en termes de valeurs propres de la matrice hessienne. Les valeurs propres de la matrice $H_f(x_0,y_0)$ sont les racines du polynôme caractéristique qui est défini par $\chi(\lambda) = \det\big(H_f(x_0,y_0) - \lambda I\big)$.

Théorème 7.
Soient U un ouvert de $\mathbb{R}^2$, f une fonction de classe $\mathscr{C}^2$ sur U et $(x_0,y_0) \in U$ un point critique de f.

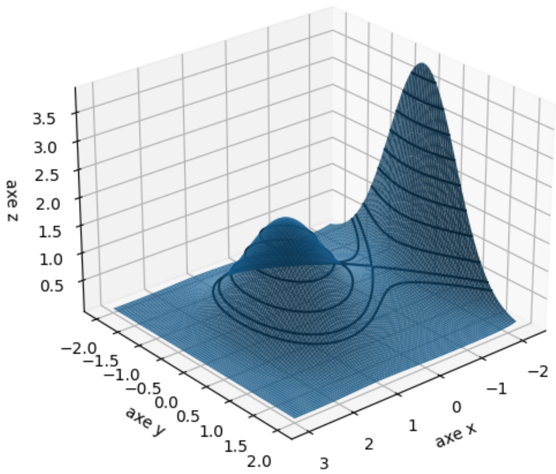
1. *Si $H_f(x_0,y_0)$ a ses deux valeurs propres strictement positives, alors f présente un minimum en (x_0,y_0).*
2. *Si $H_f(x_0,y_0)$ a ses deux valeurs propres strictement négatives, alors f présente un maximum en (x_0,y_0).*
3. *Si $H_f(x_0,y_0)$ a deux valeurs propres de signes opposés, alors f ne présente pas d'extremum en (x_0,y_0) :*

c'est un point-selle.

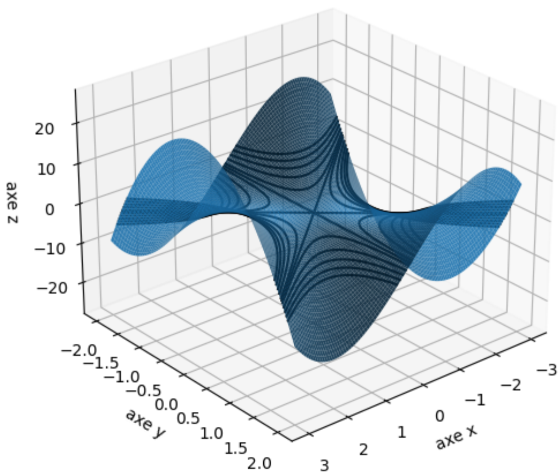
4. *Dans les autres cas, on ne peut rien dire (tout peut arriver).*

Mini-exercices.

1. Soit $f(x,y) = \exp(-\frac{1}{3}x^3 + x - y^2)$. Déterminer les deux points critiques de f et la nature (minimum/maximum/point-selle) de chacun d'entre eux.



2. Soit $f(x,y) = x^3 - 3xy^2$. Déterminer le point critique de f et sa nature. Le graphe de f s'appelle une « selle de singe ».



3. Soit $f(x,y) = x^4 + y^4 - 2x^2$. Déterminer les trois points critiques de f. Le critère de Monge permet-il de conclure ? Déterminer quand même la nature de chacun de ces points critiques.

6. Minimum et maximum : cas général

Nous étendons les résultats précédents dans deux directions : tout d'abord pour les fonctions avec un nombre quelconque de variables, puis pour la recherche de minimums/maximums sur un compact.

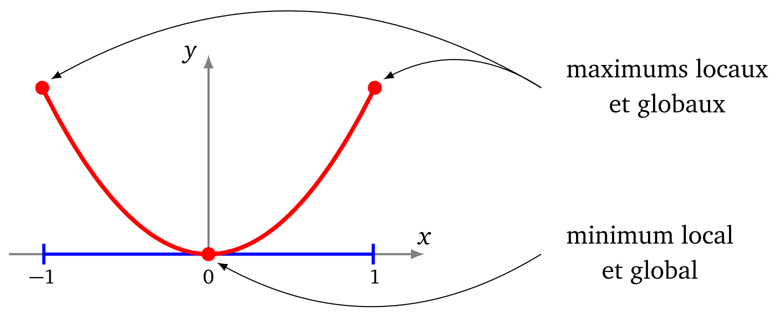
6.1. Cas de $\mathbb{R}^n$

Soit $n \geqslant 2$. Étendons les définitions au cas d'une fonction $f : U \to \mathbb{R}$, où U est un ouvert de $\mathbb{R}^n$.

- f admet un ***maximum local*** (resp. ***minimum local***) en $x_0 \in U$ s'il existe une boule ouverte $B \subset U$ centrée en x_0 telle que, pour tout $x \in B$, on ait $f(x) \leqslant f(x_0)$ (resp. $f(x) \geqslant f(x_0)$).
- f admet un ***extremum local*** en x_0 si elle y admet un maximum local ou un minimum local.
- x_0 est un ***point critique*** de f si le gradient de f s'annule en x_0, ou autrement dit si $\frac{\partial f}{\partial x_i}(x_0) = 0$ pour tout $i = 1, \ldots, n$.
- **Proposition.** Si f admet un extremum local en $x_0 \in U$ et si le gradient est défini en ce point, alors x_0 est un point critique de f.
- Ainsi, on cherche les extremums locaux parmi les points critiques de f.

6.2. Extremum sur un compact

Pour une fonction d'une variable $f : [a, b] \to \mathbb{R}$ continue sur un intervalle compact, on sait que f est bornée et atteint ses bornes, c'est-à-dire qu'elle atteint un minimum global et atteint aussi un maximum global. Ce minimum ou ce maximum peut être atteint au « bord » du segment, c'est-à-dire en a ou en b, ou bien à « l'intérieur » $]a, b[$. Il faut faire attention à bien distinguer l'étude à l'intérieur et au bord. En effet, à l'intérieur, un extremum est un point critique de f, alors que sur le bord ce n'est plus nécessairement le cas. Garder en mémoire l'exemple $f(x) = x^2$ sur $[-1, 1]$: la dérivée s'annule en $x = 0$ (minimum local et global) mais un maximum (local et global) est atteint en $x = 1$ et $x = -1$ alors que la dérivée ne s'y annule pas (explication : $+1$ et -1 sont au bord de l'intervalle).



Cette propriété se généralise aux fonctions de plusieurs variables, continues sur un ensemble compact. (Nous l'admettons.)

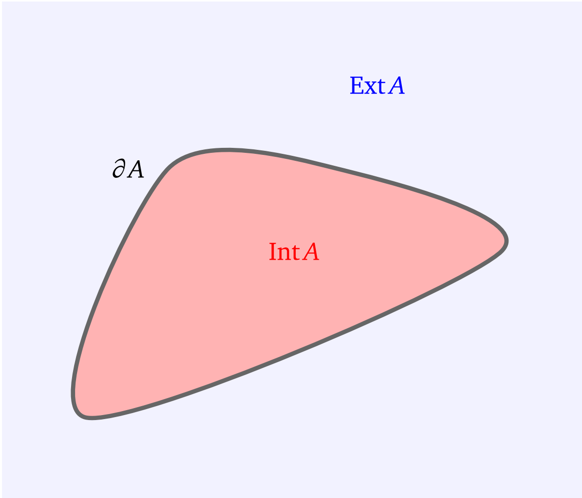
> **Proposition 2.**
> *Soit $f : K \to \mathbb{R}$ une fonction continue sur un ensemble compact $K \subset \mathbb{R}^n$. Alors f est bornée et atteint ses extremums sur K.*

La marche à suivre pour étudier les extremums d'une fonction différentiable sur un compact $K \subset \mathbb{R}^n$ est la suivante :

1. Rechercher les points critiques dans l'intérieur de K.
2. Déterminer si chacun de ces points critiques est un maximum local, un minimum local ou ni l'un ni l'autre. (Dans le cas $n = 2$, on utilise les dérivées partielles secondes. Il existe des généralisations de ces formules à n quelconque.)

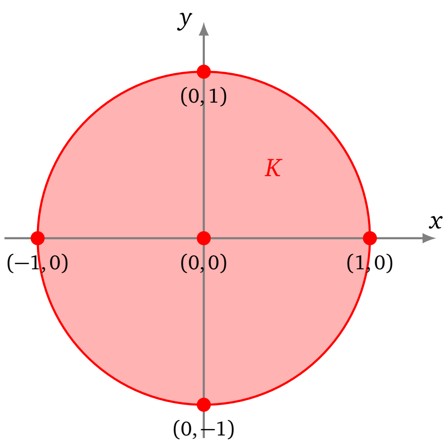
3. Chercher si f admet des extremums sur le bord de K (le critère des points critiques n'est plus valable).
4. Comparer les minimums pour déterminer le – ou les – minimums globaux ; idem pour les maximums.

On rappelle qu'un ensemble $A \subset \mathbb{R}^n$ découpe l'espace en trois parties disjointes : son intérieur $\mathrm{Int}A$, son bord ∂A (ou frontière), et son extérieur $\mathrm{Ext}A$.



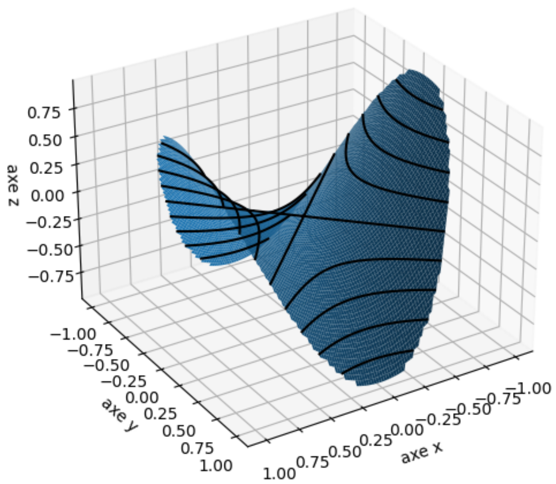
Exemple 13.
Déterminons les extremums de f sur K où :

$$f(x,y) = x^2 - y^2 \qquad \text{et} \qquad K = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}.$$



On procède de la manière suivante :

1. On cherche les points critiques et les extremums locaux dans $\mathrm{Int}K$. On trouve un seul point critique, en $(0,0)$. Mais, en $(0,0)$, f a un point-selle. La fonction n'a donc pas d'extremum à l'intérieur de K. Mais comme K est compact et f est continue sur K, f est bornée sur K et atteint ses bornes sur K. Ce sera donc sur le bord de K.
2. On analyse f sur ∂K. Une possibilité ici est de paramétrer le bord de K : c'est le cercle de rayon 1 centré en $(0,0)$ que l'on paramètre par $(\cos t, \sin t)$, pour $t \in [0, 2\pi]$. On obtient : $f(\cos t, \sin t) = \cos^2 t - \sin^2 t = \cos(2t)$. On peut alors étudier les variations de cette fonction. On obtient qu'elle est maximum, égale à 1, lorsque $2t \equiv 0 \pmod{2\pi}$, et minimum, égale à -1, lorsque $2t \equiv \pi \pmod{2\pi}$. La fonction f atteint donc son maximum 1 aux points $(1,0)$ et $(-1,0)$ de K, et son minimum -1 aux points $(0,1)$ et $(0,-1)$.



Mini-exercices.

1. Déterminer les extremums (locaux et globaux) de $f(x,y) = x^2 + y^2$ sur $C = [-1,1] \times [-1,1]$. Même question avec $f(x,y) = xy$.
2. Déterminer les extremums (locaux et globaux) de $f(x,y) = x^3 - y^2$ sur $D = \{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 \leqslant 1\}$.
3. Déterminer les extremums (locaux et globaux) de $f(x,y) = y\cos(x)$ sur $B = \mathbb{R} \times [0,1]$.

7. Extremums liés

Dans cette section, nous allons nous intéresser à la recherche des extremums sous certaines contraintes.

Vous savez que le mont Blanc culmine à 4807 mètres. Mais si vous randonnez autour du mont Blanc, quelle sera votre altitude maximale selon le parcours que vous effectuez ?

7.1. Extremums liés dans le plan

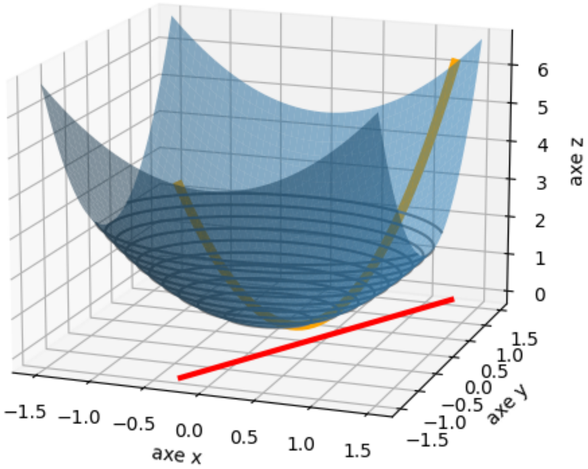
Commençons par les fonctions de deux variables.

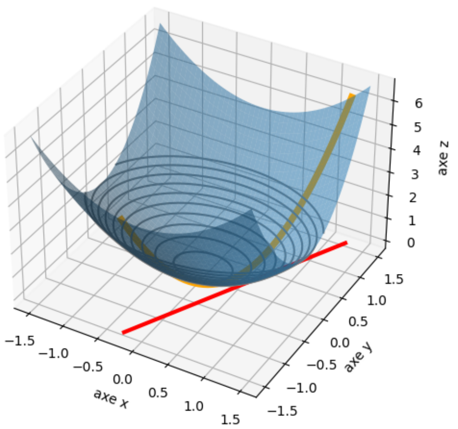
Problème : trouver le maximum (ou le minimum) d'une fonction $f(x,y)$ sachant que (x,y) doit vérifier $g(x,y) = 0$.

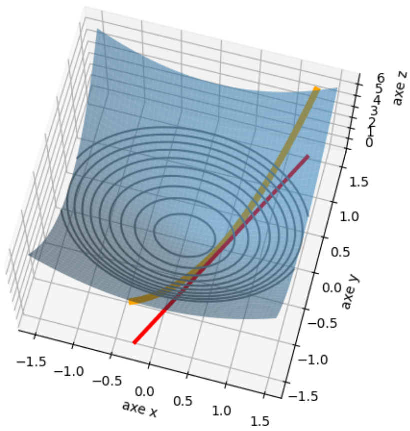
Exemple : trouver le minimum de $f(x,y) = x^2 + 2y^2$ avec la contrainte $y = 2x - 1$ (c'est-à-dire $g(x,y) = 0$ avec $g(x,y) = y - 2x + 1$ par exemple).

Clairement, le minimum global de f est 0, atteint en $(0,0)$, mais ce point ne vérifie pas la contrainte $g(x,y) = 0$.

Géométriquement, il s'agit de trouver le minimum de $f(x,y)$ en se limitant aux points (x,y) du plan situés sur la droite d'équation $y = 2x - 1$. Sur les figures ci-dessous : la surface bleue est le graphe de f, la droite rouge est l'ensemble $(g(x,y) = 0)$. La contrainte impose de ne considérer que des points (x,y,z) au-dessus de cette droite : ils sont donc ici sur la courbe orange tracée sur la surface. Trouver le minimum de f avec cette contrainte, c'est trouver le point (x,y,z) de la courbe orange ayant une altitude z minimale.





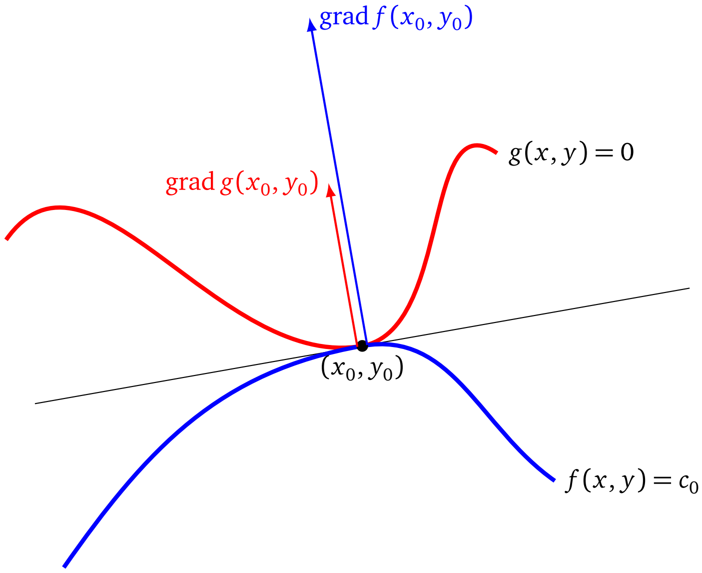


Théorème 8 (Extremums liés – Deux variables)**.**
Soient $f, g : U \to \mathbb{R}$ des fonctions de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^2$. Soit $(x_0, y_0) \in U$ tel que f soumise à la contrainte $g(x, y) = 0$ admette un extremum au point (x_0, y_0) et $\operatorname{grad} g(x_0, y_0) \neq (0, 0)$. Alors il existe un nombre réel λ tel que $\operatorname{grad} f(x_0, y_0) = \lambda \operatorname{grad} g(x_0, y_0)$. Autrement dit, on a :

$$\begin{cases} g(x_0, y_0) & = & 0 \\ \dfrac{\partial f}{\partial x}(x_0, y_0) & = & \lambda \dfrac{\partial g}{\partial x}(x_0, y_0) \\ \dfrac{\partial f}{\partial y}(x_0, y_0) & = & \lambda \dfrac{\partial g}{\partial y}(x_0, y_0) \end{cases}$$

Noter qu'on peut aussi considérer la contrainte $g(x, y) = c$ (où c est une constante), qui se ramène au cas énoncé dans le théorème en considérant $g(x, y) - c = 0$.

Interprétation géométrique. En un extremum lié (x_0, y_0), les gradients de f et g sont colinéaires. Cela signifie que la courbe $(f(x, y) = c_0)$ (avec $c_0 = f(x_0, y_0)$) et la courbe $(g(x, y) = 0)$ ont la même tangente au point (x_0, y_0). Autrement dit, ces deux courbes sont tangentes entre elles.



Remarque sur les équations. Le théorème des extremums liés en dimension n quelconque conduit à un système de n équations à n inconnues. Mais attention ! Les équations ne sont pas linéaires. Ainsi, le système obtenu est généralement très difficile voire impossible à résoudre, même pour des situations simples. C'est pourquoi les fonctions que l'on étudiera dans la suite sont toutes très simples, afin de pouvoir résoudre les équations.

Exemple 14.
Reprenons l'exemple illustré précédemment : trouver le minimum de $f(x, y) = x^2 + 2y^2$ avec la contrainte $y = 2x - 1$.
Géométriquement, on cherche le point situé sur la droite d'équation $y = 2x - 1$ qui minimise la fonction f.
On pose $g(x, y) = y - 2x + 1$.

$$\operatorname{grad} f(x, y) = \begin{pmatrix} 2x \\ 4y \end{pmatrix} \qquad \operatorname{grad} g(x, y) = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$$

En un extremum lié, les gradients sont colinéaires : $\operatorname{grad} f(x, y) = \lambda \operatorname{grad} g(x, y)$. On obtient deux équations :

$$2x = -2\lambda \tag{1}$$

$$4y = \lambda \tag{2}$$

auxquelles on rajoute la contrainte $g(x, y) = 0$:

$$y = 2x - 1 \tag{3}$$

Par (1), on obtient $x = -\lambda$, et par (2), on trouve $y = \frac{\lambda}{4}$, qu'on injecte dans (3) :

$$\frac{\lambda}{4} = -2\lambda - 1.$$

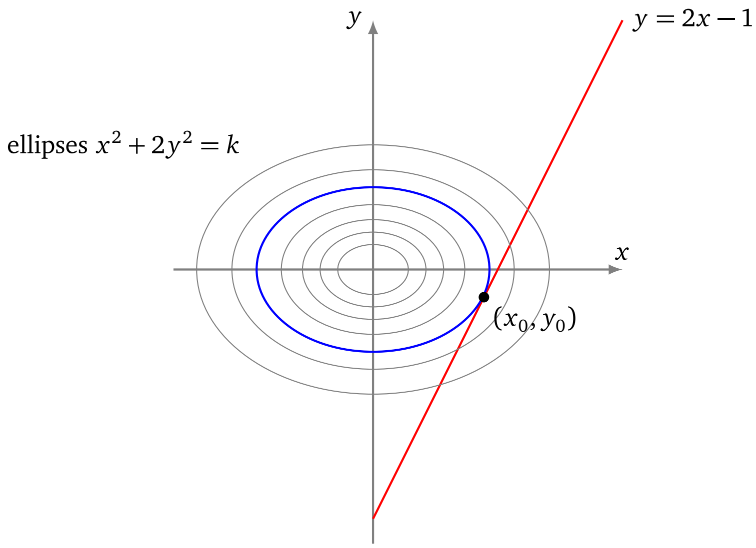
Ainsi,

$$\lambda_0 = -\frac{4}{9}.$$

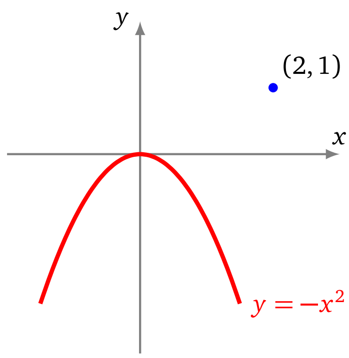
On obtient donc la solution :

$$(x_0, y_0) = \left(\frac{4}{9}, -\frac{1}{9}\right).$$

Sur $(g(x, y) = 0)$, la fonction f est positive et tend vers $+\infty$ lorsque $\|(x, y)\|$ tend vers $+\infty$, donc f admet au moins un minimum. Conclusion : f admet un unique minimum sur $(g(x, y) = 0)$, et ce minimum est atteint en (x_0, y_0).
Géométriquement, les courbes $(f(x, y) = k)$ sont des ellipses, que l'on fait grandir jusqu'à obtenir l'ellipse tangente à la droite. Le point de contact est le point recherché.



Exemple 15.
Quel est le point de la parabole définie par $y = -x^2$ le plus proche du point $(2,1)$?



On pose $f(x,y) = (x-2)^2 + (y-1)^2$ qui mesure la distance (au carré) entre le point (x,y) et le point $(2,1)$. La contrainte est donnée par $g(x,y) = 0$ où $g(x,y) = y + x^2$.

$$\operatorname{grad} f(x,y) = \begin{pmatrix} 2(x-2) \\ 2(y-1) \end{pmatrix} \qquad \operatorname{grad} g(x,y) = \begin{pmatrix} 2x \\ 1 \end{pmatrix}.$$

En un extremum lié, les gradients sont colinéaires : $\operatorname{grad} f(x,y) = \lambda \operatorname{grad} g(x,y)$. Donc :

$$2(x-2) = 2\lambda x \tag{1}$$

$$2(y-1) = \lambda \tag{2}$$

On rajoute la contrainte $g(x,y) = 0$:

$$y + x^2 = 0 \tag{3}$$

Par (1), on a $x = \frac{2}{1-\lambda}$, et par (2), on a $y = \frac{\lambda}{2} + 1$, qu'on injecte dans (3) :

$$\frac{\lambda}{2} + 1 + \frac{4}{(1-\lambda)^2} = 0.$$

Cette dernière équation est équivalente à :

$$(\lambda+2)(\lambda-1)^2 + 8 = 0,$$

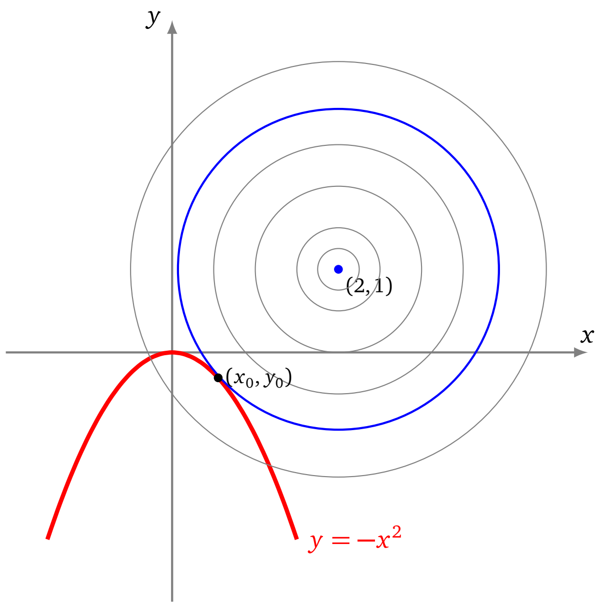
qui sous forme développée est $\lambda^3 - 3\lambda + 10 = 0$. C'est une équation de degré 3, qui ici admet une seule solution réelle (loin d'être évidente) :

$$\lambda_0 = \frac{-1 - \left(\sqrt[3]{5 - 2\sqrt{6}}\right)^2}{\sqrt[3]{5 - 2\sqrt{6}}}$$

On obtient donc la solution :

$$(x_0, y_0) = \left(\frac{2}{1-\lambda_0}, \frac{\lambda_0}{2} + 1 \right).$$

Approximations numériques : $\lambda_0 \simeq -2.613$ et $(x_0, y_0) \simeq (0.554, -0.306)$.
Interprétation géométrique. On fait grandir un cercle centré à l'origine jusqu'à ce qu'il soit tangent à la parabole. Le point de contact obtenu correspond à la distance minimale. En augmentant encore le rayon, on pourrait obtenir d'autres extremums liés (mais ce n'est pas le cas ici).



7.2. Extremums liés avec une seule contrainte

Passons au cas d'un nombre quelconque de variables, mais toujours avec une seule contrainte. Il s'agit de trouver les extremums de $f(x)$ lorsque x appartient à une hypersurface définie par $g(x) = 0$.
Exemple : trouver le minimum de $f(x, y, z) = x^2 + y^2 + z^2$ avec la contrainte $x + 2y + 3z = 1$.

> **Théorème 9** (Extremums liés – Une contrainte).
> *Soient $f, g : U \to \mathbb{R}$ des fonctions de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^n$. Soit $x_0 \in U$ tel que f soumise à la contrainte $g(x) = 0$ admette un extremum au point x_0 et $\operatorname{grad} g(x_0) \neq 0$. Alors il existe un nombre réel λ tel que $\operatorname{grad} f(x_0) = \lambda \operatorname{grad} g(x_0)$.*

Le réel λ est appelé ***multiplicateur de Lagrange***.

Remarque.
Si x_0 est un extremum lié, on a $\operatorname{grad} f(x_0)$ colinéaire à $\operatorname{grad} g(x_0)$. La réciproque n'est pas vraie. Nous avons une condition nécessaire mais pas suffisante. C'est l'équivalent de la nullité de la dérivée pour les extremums libres : en un extremum libre la dérivée est nulle, mais la dérivée peut être nulle sans que la fonction ait un extremum (penser à $x \mapsto x^3$ en $x = 0$).

Exemple 16.
Trouver le minimum de $f(x, y, z) = x^2 + y^2 + z^2$ avec la contrainte $x + 2y + 3z = 1$.
Géométriquement, on cherche le point, situé sur le plan d'équation $x + 2y + 3z = 1$, qui est le plus proche de l'origine.
On pose $g(x, y, z) = x + 2y + 3z - 1$.

$$\operatorname{grad} f(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \qquad \operatorname{grad} g(x, y, z) = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

En un extremum lié, les gradients sont colinéaires : $\operatorname{grad} f(x,y,z) = \lambda \operatorname{grad} g(x,y,z)$. On obtient trois équations

$$2x = \lambda \tag{1}$$
$$2y = 2\lambda \tag{2}$$
$$2z = 3\lambda \tag{3}$$

auxquelles on rajoute la contrainte $g(x,y,z)=0$:

$$x+2y+3z=1 \tag{4}$$

Par (1), on a $x=\frac{1}{2}\lambda$, par (2), on a $y=\lambda$, et par (3), on a $z=\frac{3}{2}\lambda$, qu'on injecte dans (4) :

$$\frac{1}{2}\lambda+2\lambda+\frac{9}{2}\lambda=1.$$

Ainsi,

$$\lambda_0=\frac{1}{7} \qquad \text{et} \qquad (x_0,y_0,z_0)=\left(\frac{1}{14},\frac{1}{7},\frac{3}{14}\right).$$

Sur $(g(x,y,z)=0)$, la fonction f est positive et tend vers $+\infty$ lorsque $\|(x,y,z)\|$ tend vers $+\infty$, donc f admet au moins un minimum. Conclusion : f admet un unique minimum sur $(g(x,y,z)=0)$, et ce minimum est atteint en (x_0,y_0,z_0).

Exemple 17.
Trouver les extremums de $f(x,y,z)=x+2y+3z$ avec la contrainte $x^2+y^2+z^2=1$.
Géométriquement, on cherche les points situés sur la sphère de rayon 1 centrée à l'origine qui minimisent ou maximisent la fonction f. Comme la sphère est un ensemble compact (et f est continue), on sait que le minimum et le maximum seront atteints.
On pose $g(x,y,z)=x^2+y^2+z^2-1$.

$$\operatorname{grad} f(x,y,z)=\begin{pmatrix}1\\2\\3\end{pmatrix} \qquad \operatorname{grad} g(x,y,z)=\begin{pmatrix}2x\\2y\\2z\end{pmatrix}.$$

En un extremum lié, on a $\operatorname{grad} f(x,y,z)=\lambda \operatorname{grad} g(x,y,z)$:

$$1=2\lambda x \tag{1}$$
$$2=2\lambda y \tag{2}$$
$$3=2\lambda z \tag{3}$$

On a de plus $g(x,y,z)=0$:

$$x^2+y^2+z^2=1 \tag{4}$$

D'une part, grâce à (1), (2), (3) :

$$(2\lambda x)^2+(2\lambda y)^2+(2\lambda z)^2=1^2+2^2+3^2=14.$$

D'autre part, grâce à (4) :

$$(2\lambda x)^2+(2\lambda y)^2+(2\lambda z)^2=4\lambda^2.$$

Ainsi,

$$\lambda=\pm\frac{1}{2}\sqrt{14}.$$

On obtient donc deux solutions :

$$(x_1,y_1,z_1)=\left(\frac{1}{\sqrt{14}},\frac{2}{\sqrt{14}},\frac{3}{\sqrt{14}}\right) \qquad \text{et} \qquad (x_2,y_2,z_2)=\left(-\frac{1}{\sqrt{14}},-\frac{2}{\sqrt{14}},-\frac{3}{\sqrt{14}}\right).$$

L'une doit correspondre au maximum de f (c'est clairement (x_1,y_1,z_1)) et l'autre au minimum de f (c'est (x_2,y_2,z_2)).

7.3. Extremums liés avec plusieurs contraintes

Généralisons la situation précédente en présence de k contraintes.

Théorème 10 (Extremums liés – Plusieurs contraintes)**.**
Soient $f : U \to \mathbb{R}$ et $g_1, \ldots, g_k : U \to \mathbb{R}$ des fonctions de classe $\mathscr{C}^1$ sur un ouvert U de $\mathbb{R}^n$.
Supposons que les vecteurs $\operatorname{grad} g_1(x), \ldots, \operatorname{grad} g_k(x)$ soient linéairement indépendants, pour tout x de l'ensemble S défini par $(g_1(x) = 0, \ldots, g_k(x) = 0)$.
Si f admet un extremum lié sur S en x_0 alors le vecteur $\operatorname{grad} f(x_0)$ est combinaison linéaire des vecteurs $\operatorname{grad} g_i(x_0)$: il existe $\lambda_1, \ldots, \lambda_k$ tels que

$$\operatorname{grad} f(x_0) = \lambda_1 \operatorname{grad} g_1(x_0) + \cdots + \lambda_k \operatorname{grad} g_k(x_0).$$

Les réels $\lambda_1, \ldots, \lambda_k$ sont appelés ***multiplicateurs de Lagrange***.

Exemple 18.
Trouver le point (x, y, z) le plus proche de l'origine vérifiant $z^2 = x^2 + y^2$ et $x + y - 2z + 1 = 0$.
On pose $f(x, y, z) = x^2 + y^2 + z^2$, $g_1(x, y, z) = x^2 + y^2 - z^2$ et $g_2(x, y, z) = x + y - 2z + 1$.

$$\operatorname{grad} f(x,y,z) = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \qquad \operatorname{grad} g_1(x,y,z) = \begin{pmatrix} 2x \\ 2y \\ -2z \end{pmatrix} \qquad \operatorname{grad} g_2(x,y,z) = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$$

En un extremum lié, $\operatorname{grad} f(x, y, z)$ est combinaison linéaire de $\operatorname{grad} g_1(x, y, z)$ et $\operatorname{grad} g_2(x, y, z)$:

$$\operatorname{grad} f(x,y,z) = \lambda \operatorname{grad} g_1(x,y,z) + \mu \operatorname{grad} g_2(x,y,z).$$

On obtient trois équations

$$2x = 2\lambda x + \mu \tag{1}$$
$$2y = 2\lambda y + \mu \tag{2}$$
$$2z = -2\lambda z - 2\mu \tag{3}$$

auxquelles on rajoute les contraintes $g_1(x, y, z) = 0$ et $g_2(x, y, z) = 0$:

$$z^2 - x^2 - y^2 = 0 \tag{4}$$
$$x + y - 2z + 1 = 0 \tag{5}$$

Nous obtenons donc 5 équations (non linéaires !) avec 5 inconnues. Après calculs, les solutions sont :

$$(x_1, y_1, z_1) = \left(\frac{1+\sqrt{2}}{2}, \frac{1+\sqrt{2}}{2}, \frac{2+\sqrt{2}}{2} \right) \qquad \text{et} \qquad (x_2, y_2, z_2) = \left(\frac{1-\sqrt{2}}{2}, \frac{1-\sqrt{2}}{2}, \frac{2-\sqrt{2}}{2} \right).$$

Le point (x_1, y_1, z_1) est un maximum lié de f, et (x_2, y_2, z_2) est un minimum lié. Géométriquement, $z^2 = x^2 + y^2$ est un cône, que l'on intersecte par le plan $x + y - 2z + 1 = 0$ pour obtenir une ellipse de l'espace (donc un ensemble compact). Nous avons calculé le point (x_2, y_2, z_2) de cette ellipse le plus proche de l'origine.

Mini-exercices.

1. Quelles sont les dimensions du rectangle de périmètre 16 qui a l'aire la plus grande possible ?
2. Trouver le point du cercle de rayon 1 centré à l'origine le plus éloigné du point $(3, 4)$.
3. Calculer la distance entre le point (x_0, y_0, z_0) et le plan d'équation $ax + by + cz + d = 0$.
4. Le cylindre infini d'équation $x^2 + y^2 = 4$ et le plan $x + 2y + z = 2$ se coupent en une ellipse. Quel est le point de cette ellipse le plus éloigné de l'origine ?

Auteurs du chapitre

Arnaud Bodin. D'après des cours de Abdellah Hanani (Lille), Goulwen Fichou et Stéphane Leborgne (Rennes), Laurent Pujo-Menjouet (Lyon). Relu par Anne Bauval et Vianney Combet.

Version 1.0 – Octobre 2022