FONCTIONS DE DEUX VARIABLES

33

Dans ce chapitre, on étudie les fonctions de deux variables à valeurs réelles, c'est-à-dire des fonctions définies sur $\mathbf{R}^2$, ou une partie de $\mathbf{R}^2$, à valeurs dans $\mathbf{R}$.
La plupart des concepts que l'on utilise habituellement en analyse, et notamment les notions de croissance ou de décroissance n'ont plus cours[1].
Malgré tout, nous allons généraliser la notion de continuité et celle de dérivabilité à de telles fonctions, l'un des buts étant notamment de commencer à dégager des critères pour trouver les extrema d'une telle fonction, malgré l'absence d'un analogue à la notion de tableau de variations.
Tout ce travail sera poursuivi en seconde année, et élargi au cadre de fonctions définies sur $\mathbf{R}^n$, et pas seulement sur $\mathbf{R}^2$.
Il est important d'essayer de vous faire une intuition géométrique pour les fonctions de deux variables, car nous pourrons représenter graphiquement ces fonctions[2], ce qui ne sera plus possible pour des fonctions de $n \geqslant 3$ variables.

[1] Il n'y a pas de relation d'ordre «naturelle» sur $\mathbf{R}^2$.

[2] Par des surfaces en 3D.

Un certain nombre de preuves de ce chapitre seront très techniques, faute d'un vocabulaire adéquat, vous consacrerez du temps en seconde année à développer des outils simplifiant ces preuves, et les inscrivant dans un contexte plus général.

33.1 INTRODUCTION À LA TOPOLOGIE DE $\mathbf{R}^2$

Lors de l'étude des fonctions d'une variable, nous considérions souvent des fonctions définies sur des intervalles, ou des réunions d'intervalles. Et certains résultats de calcul différentiel[3] nécessitaient de se placer sur des intervalles ouverts.
Nous définissons dans cette partie les parties ouvertes de $\mathbf{R}^2$, qui sont «le bon cadre» pour faire du calcul différentiel.

[3] Notamment celui qui affirme qu'une fonction dérivable atteint ses extrema en des points critiques.

Dans tout le chapitre, $\langle \cdot, \cdot \rangle$ désigne le produit scalaire canonique de $\mathbf{R}^2$:
$\langle (x_1, y_1), (x_2, y_2) \rangle = x_1x_2 + y_1y_2$, et $\|\cdot\|$ désigne la norme euclidienne associée : $\|(x, y)\| = \sqrt{x^2 + y^2}$.

33.1.1 Boules ouvertes, boules fermées

L'idée principale est que si $u = (x, y)$ et $v = (x', y')$ sont deux points de $\mathbf{R}^2$, alors $\|u - v\| = \sqrt{(x - x')^2 + (y - y')^2}$ représente la distance entre u et v.

Définition 33.1 – Soit $a \in \mathbf{R}^2$ et $r > 0$. On appelle **boule ouverte de centre a et de rayon r** l'ensemble

$$B_o(a, r) = \left\{ b \in \mathbf{R}^2 \mid \|a - b\| < r \right\}.$$

La boule de centre a et de rayon r est donc l'ensemble des points de $\mathbf{R}^2$ dont la distance à a est strictement inférieure à r.
C'est l'analogue, en dimension 2, de $]a - \eta, a + \eta[= \{x \in \mathbf{R} \mid |x - a| < \eta\}$, l'ensemble des réels à distance strictement inférieure à η de a.
Remarque. Il est évident que si $0 < r < r'$, alors $B_o(a, r) \subset B_o(a, r')$.

Définition 33.2 – Soit $a \in \mathbf{R}^2$, et soit $r \geqslant 0$. On appelle **boule fermée de centre a et de rayon r** l'ensemble

$$B_f(a, r) = \left\{ b \in \mathbf{R}^2 \mid \|a - b\| \leqslant r \right\}.$$

Notons qu'une boule fermée de rayon 0 n'est rien d'autre qu'un singleton : $B_f(a, 0) = \{a\}$.

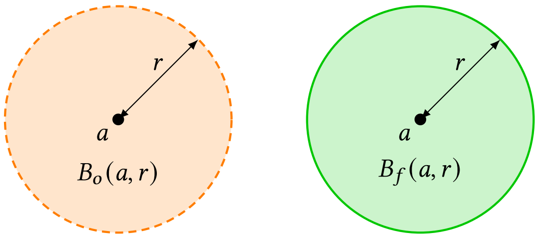


Figure 33.1 – Une boule ouverte et une boule fermée. La différence entre les deux est que les points du cercle de centre a et de rayon r (ceux qui sont exactement à distance r de a) sont dans la boule fermée, pas dans la boule ouverte.

33.1.2 Parties ouvertes, parties fermées

Nous avons utilisé à plusieurs reprises le fait que si un intervalle I de $\mathbf{R}$ est ouvert, alors pour tout $a \in I$, il existe $\eta > 0$ (dépendant de a) tel que $]a-\eta, a+\eta[\subset I$. Autrement dit, les points suffisamment proches de a sont encore dans I.
C'est cette idée qui guide la définition de partie ouverte de $\mathbf{R}^2$.

Définition 33.3 – Soit $\mathcal{O}$ une partie de $\mathbf{R}^2$. On dit que $\mathcal{O}$ est un **ouvert** de $\mathbf{R}^2$ si

$$\forall a \in \mathcal{O},\ \exists r \in \mathbf{R}_+^*,\ B_o(a,r) \subset \mathcal{O}.$$

Soit encore si

$$\forall a \in \mathcal{O},\ \exists r \in \mathbf{R}_+^*,\ \forall x \in \mathbf{R}^2,\ \|a-x\| < r \Rightarrow x \in \mathcal{O}.$$

Une première intuition est qu'une partie ouverte est une partie qui ne contient pas son bord, ou sa frontière[4].

[4] Une définition précise de ces notions sera donnée l'an prochain.

Exemples 33.4

▸ Une boule ouverte est un ouvert. En effet, considérons $a \in \mathbf{R}^2$ et $r > 0$, et prouvons que $B_o(a,r)$ est un ouvert.
Soit $x \in B_o(a,r)$, de sorte que $\|a-x\| < r$.
Soit $r_1 > 0$ tel que $\|a-x\|+r_1 < r$ (notons qu'il existe toujours un tel r_1, par exemple $r_1 = \frac{r-\|x-a\|}{2}$).
Alors $B_o(x,r_1) \subset B_o(a,r)$.
En effet, pour $y \in B_o(x,r_1)$, on a

$$\|y-a\| = \|(y-x)+(x-a)\| \leqslant \|y-x\| + \|x-a\| \leqslant r_1 + \|x-a\| < r,$$

de sorte que $y \in B_o(a,r)$.
▸ $\mathbf{R}_+^* \times \mathbf{R}_+^*$ est un ouvert de $\mathbf{R}^2$.
En effet, soit $a = (x,y) \in \mathbf{R}_+^* \times \mathbf{R}_+^*$.
Soit alors $r = \min(x,y) > 0$, et soit $u = (x',y') \in B_o(a,r)$.
Alors $|x-x'|^2 \leqslant (x-x')^2 + (y-y')^2 < r^2 \leqslant x^2$.
Donc $-x < x - x' < x$, si bien que $x' > 0$.
Et de même, on prouve que $y' > 0$, et donc que $u \in \mathbf{R}_+^* \times \mathbf{R}_+^*$, si bien que $B_o(a,r) \subset \mathbf{R}_+^* \times \mathbf{R}_+^*$.
▸ La boule fermée $B_f(0_{\mathbf{R}^2}, r)$ n'est en revanche pas un ouvert.
Si x est tel que $\|x\| = r$, alors quel que soit $r_1 > 0$, la boule $B_o(x,r_1)$ n'est pas incluse dans $B_f(0,r)$.
En effet, soit $y = \left(1+\frac{r_1}{2r}\right)x$. Alors $y \in B_o(x,r_1)$ car

$$\|x-y\| = \left\|x - \left(1+\frac{r_1}{2r}\right)x\right\| = \frac{r_1}{2r}\|x\| = \frac{r_1}{2} < r_1,$$

mais $\|0_{\mathbf{R}^2} - y\| = \left\|0_{\mathbf{R}^2} - \left(1+\frac{r_1}{2r}\right)x\right\| = \left(1+\frac{r_1}{2r}\right)\underbrace{\|x\|}_{=r} > r$, de sorte que

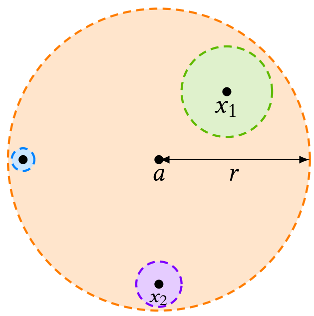


Figure 33.2– Le rayon r_1 de la «petite» boule dépend du point x.

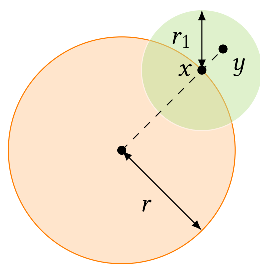


Figure 33.3– $B_o(x,r_1)$ n'est pas inclus dans $B_f(0,r)$. On constate que x et y sont colinéaires.

> $y \notin B_f(0_{\mathbf{R}^2}, r)$.
> Sur le même principe, on prouve qu'une boule fermée n'est jamais un ouvert.

Proposition 33.5 (Propriétés des ouverts) :

1. *$\emptyset$ et $\mathbf{R}^2$ sont des ouverts*
2. *si $(\mathcal{O}_i)_{i \in I}$ est une famille[5] d'ouverts de $\mathbf{R}^2$, alors $\bigcup_{i \in I} \mathcal{O}_i$ est un ouvert de $\mathbf{R}^2$.*
3. *si $\mathcal{O}_1, \dots, \mathcal{O}_n$ sont des ouverts, en nombre fini, alors $\bigcap_{i=1}^{n} \mathcal{O}_i$ est un ouvert de $\mathbf{R}^2$.*

[5] Indexée par un ensemble I potentiellement infini.

Démonstration. 1. $\emptyset$ est un ouvert puisque $\forall x \in \emptyset, \dots$ est toujours vrai.
Et $\mathbf{R}^2$ est un ouvert puisque pour tout $x \in \mathbf{R}^2$, $B_o\left(x, \frac{\pi^2}{6}\right) \subset \mathbf{R}^2$.

2. Soit $x \in \bigcup_{i \in I} \mathcal{O}_i$, et soit $i_0 \in I$ tel que $x \in \mathcal{O}_{i_0}$. Alors il existe $r > 0$ tel que $B_o(x, r) \subset \mathcal{O}_{i_0} \subset \bigcup_{i \in I} \mathcal{O}_i$.

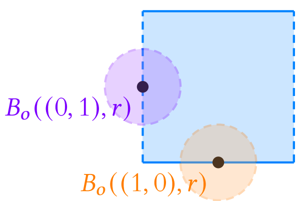
3. Soit $x \in \bigcap_{i=1}^{n} \mathcal{O}_i$. Alors pour tout $i \in [\![1, n]\!]$, il existe $r_i > 0$ tel que $B_o(x, r_i) \subset \mathcal{O}_i$.
En particulier, si on note $r = \min_{1 \leqslant i \leqslant n} r_i$, alors pour tout i, $B_o(x, r) \subset \mathcal{O}_i$.
Et donc $B_o(x, r) \subset \bigcap_{i=1}^{n} \mathcal{O}_i$.

□

La notion de partie fermée n'est pas explicitement au programme de première année. Mais puisque vous la rencontrerez en seconde année, autant s'en faire une intuition dès maintenant.

Définition 33.6 – Une partie $\mathcal{F}$ de $\mathbf{R}^2$ est un **fermé** de $\mathbf{R}^2$ si son complémentaire $\overline{\mathcal{F}}$ est un ouvert.

⚠ Les fermés ne sont donc pas les parties non ouvertes. Il existe des parties de $\mathbf{R}^2$ qui ne sont ni ouvertes ni fermées. Par exemple $A =]0, 2[\times [0, 2]$ n'est ni ouvert ni fermé.
En effet, il n'est pas ouvert car il contient $(1, 0)$ mais que pour tout $r > 0$, $B_o((0, 1), r)$ contient $(-\frac{r}{2}, 1)$ qui n'est pas dans A.
Il n'est pas non plus fermé car $(0, 1) \in \overline{A}$, mais pour tout $r > 0$, $B_o((0, 1), r)$ contient $(\frac{r}{2}, 1) \notin \overline{A}$.



Proposition 33.7 :

1. *$\emptyset$ et $\mathbf{R}^2$ sont des fermés de $\mathbf{R}^2$*
2. *si $(\mathcal{F}_i)_{i \in I}$ est une famille quelconque de fermés, alors $\bigcap_{i \in I} \mathcal{F}_i$ est un fermé*
3. *si $\mathcal{F}_1, \dots, \mathcal{F}_n$ sont des fermés, en nombre fini, alors $\bigcup_{i=1}^{n} \mathcal{F}_i$ est un fermé.*

Démonstration. 1) $\emptyset$ et $\mathbf{R}^2$ sont des ouverts, donc leurs complémentaires respectifs, qui sont $\mathbf{R}^2$ et $\emptyset$ sont des fermés.
2) Pour tout $i \in I$, $\overline{\mathcal{F}_i}$ est un ouvert. Donc $\bigcup_{i\in I} \overline{\mathcal{F}_i} = \overline{\bigcap_{i\in I} \mathcal{F}_i}$ est un ouvert, si bien que $\bigcap_{i\in I} \mathcal{F}_i$ est fermé.
3) De même, $\bigcap_{i\in I} \overline{\mathcal{F}_i}$ est ouvert, donc son complémentaire, $\bigcup_{i\in I} \mathcal{F}_i$ est fermé. □

33.2 Fonctions de deux variables

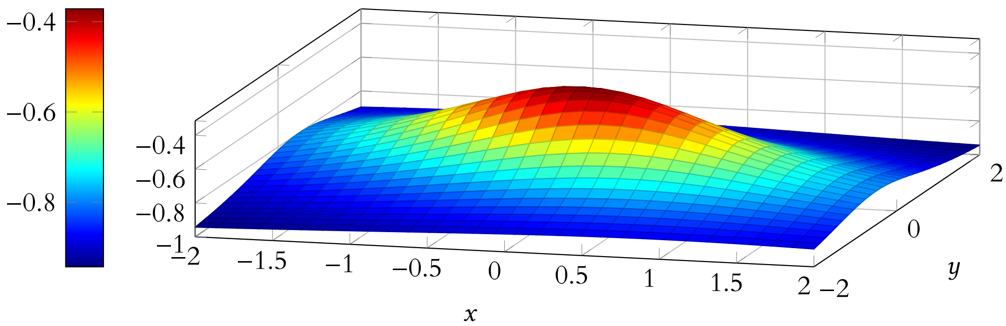
Dans la suite du chapitre, nous allons considérer des fonctions $f : A \to \mathbf{R}$ définies sur une partie A de $\mathbf{R}^2$.

Définition 33.8 – Soit $f : A \to \mathbf{R}$ une fonction définie sur une partie A de $\mathbf{R}^2$. On appelle **graphe de** f l'ensemble $\{(x, y, z) \in \mathbf{R}^3 \mid (x, y) \in A$ **et** $z = f(x, y)\}$.

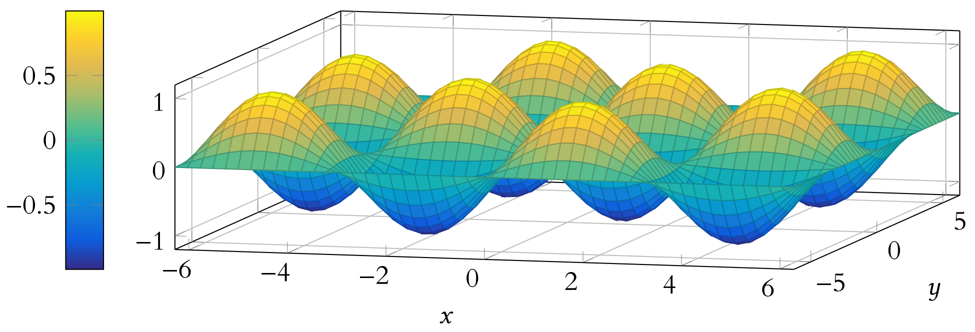
Le graphe d'une fonction est donc une surface tracée dans $\mathbf{R}^3$.

Exemples 33.9

• Soit $f(x, y) = -e^{-\frac{1}{x^2+2y^2+xy+1}}$. Alors le graphe de f est



• Soit $g(x, y) = \sin(x)\sin(y)$. Alors le graphe de g est

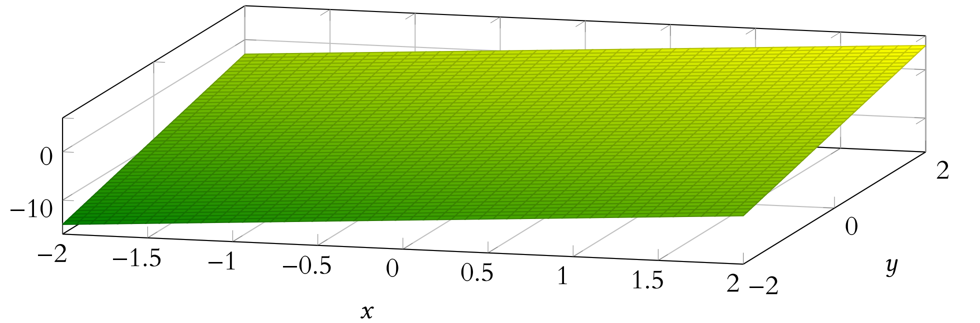


• Si l'on considère une partie de la surface du globe, et qu'on la suppose plate. Alors tout point de la surface est représenté par deux coordonnées x et y (qui sont la latitude et la longitude). Alors si l'on note $h(x, y)$ l'altitude en ce point, le graphe de la fonction h représente la surface de la Terre.

2D/3D
Bien qu'il s'agisse d'une représentation «en 3 dimensions», ces figures sont toujours représentées sur le plan de la feuille ou de l'écran. Les couleurs aident alors à se représenter une troisième dimension, les variations de couleurs représentant les variations d'altitude.

Cas particulier : si $f(x, y) = ax + by + c$, alors le graphe de f est un plan de l'espace, car c'est l'ensemble des points (x, y, z) vérifiant

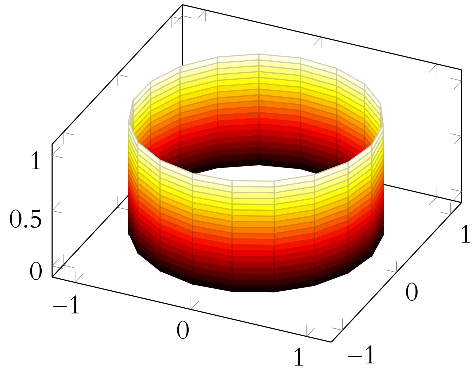
$$z = ax + by + c \Leftrightarrow z - ax - by - c = 0.$$

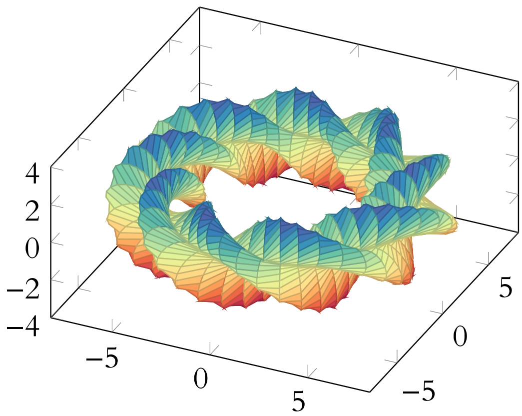


Bien entendu, toutes les «surfaces»[6] tracées dans $\mathbf{R}^3$ ne sont pas le graphe d'une fonction définie sur une partie de $\mathbf{R}^2$, en particulier, le graphe d'une fonction ne peut pas contenir deux points distincts de même abscisse et même ordonnée.

[6] Whatever it means.

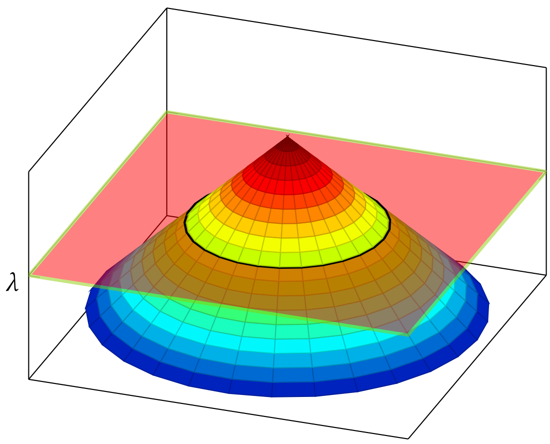
Par exemple, les surfaces suivantes ne sont pas des graphes de fonctions de deux variables.





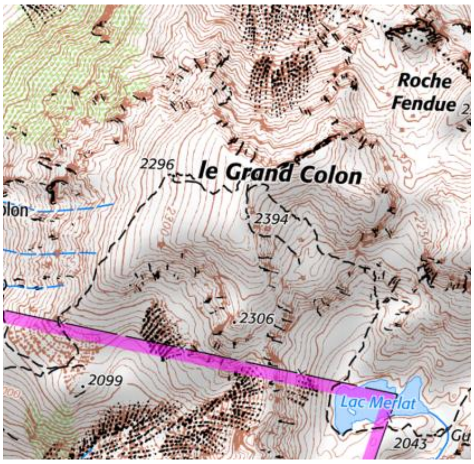
Définition 33.10 – Soit $f : A \to \mathbf{R}$. Pour tout $\lambda \in \mathbf{R}$, on appelle **ligne de niveau** λ **de** f l'ensemble $\mathscr{C}_\lambda(f) = f^{-1}(\{\lambda\}) = \{x \in A \mid f(x) = \lambda\}$.

Les lignes de niveau sont des parties de $\mathbf{R}^2$, mais la ligne de niveau λ de f est reliée à l'intersection du graphe de f avec le plan (horizontal) d'équation $z = \lambda$.
Plus précisément, cette intersection est $\{(x, y, \lambda),\ (x, y) \in \mathscr{C}_\lambda(f)\}$.

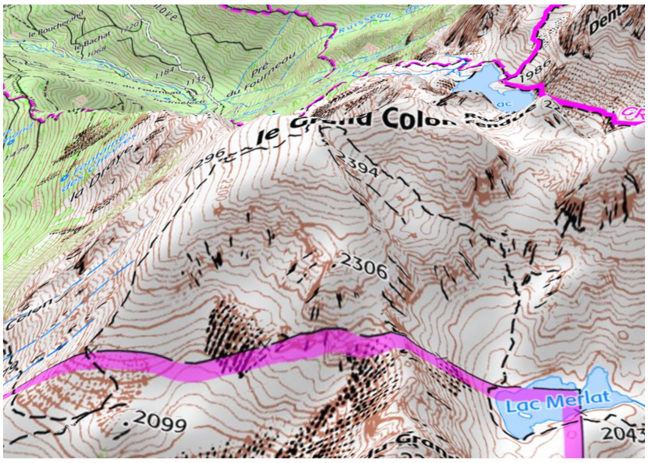


Exemples 33.11

Si f est la fonction qui à la longitude et la latitude associe l'altitude, alors les lignes de niveau représentent les points qui sont à la même altitude : si on se promène sur une ligne de niveau, on ne monte ni ne descend. Ce sont bien les lignes de niveau représentées sur les cartes topographiques.

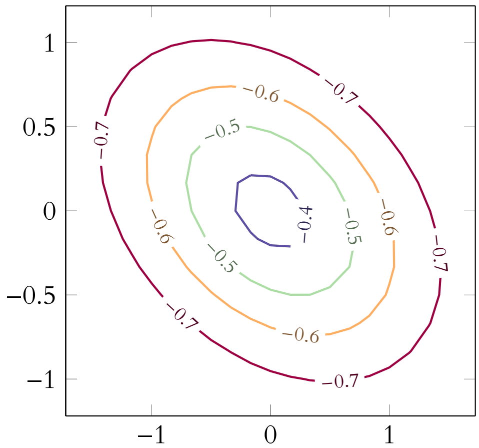


Des lignes de niveau en Belledonne

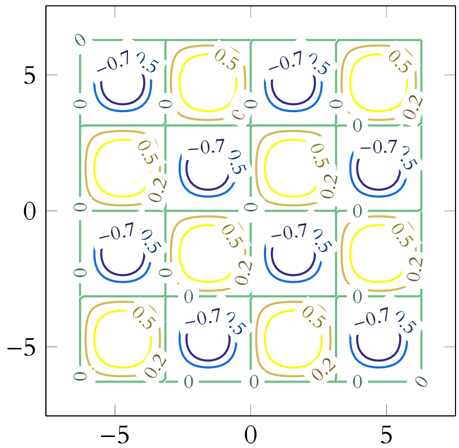


Les mêmes lignes de niveau en 3D.

Des courbes de niveau des deux fonctions f et g données précédemment sont tracées ci-dessous.



Des lignes de niveau de f.



Des lignes de niveau de g.

33.3 Fonctions continues

Définition 33.12 – Soit A une partie de $\mathbf{R}^2$, soit $f : A \to \mathbf{R}$, et soit $a \in A$. On dit que f **est continue en** a si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in A,\ \|x - a\| < \eta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Soit encore si

$$\forall \varepsilon > 0,\ \exists \eta > 0,\ \forall x \in B_o(a, \eta) \cap A,\ |f(x) - f(a)| < \varepsilon.$$

On dit que f **est continue sur** A si elle est continue en tout $a \in A$.

Remarques. ▸ Vous aurez bien sûr reconnu l'analogie avec la définition quantifiée de la continuité en a d'une fonction d'une seule variable.
L'intuition y est d'ailleurs la même : deux points suffisamment proches ont des images proches.
On pourrait de même définir une notion de limite en un point d'une fonction définie sur une partie de $\mathbf{R}^2$, vous le ferez l'an prochain.
▸ Si f est continue en a, alors elle est bornée sur un voisinage de a : il existe $r > 0$ tel que pour tout $x \in A \cap B_o(a, r)$, $|f(x)| \leqslant |f(a)| + 1$.

Exemples 33.13

▸ Une fonction constante est bien évidemment continue sur son ensemble de définition.

▸ Soit $p_1 : \left| \begin{array}{ccc} \mathbf{R}^2 & \longrightarrow & \mathbf{R} \\ (x,y) & \longmapsto & x \end{array} \right.$.

Soit alors $a = (a_1, a_2) \in \mathbf{R}^2$, et soit $\varepsilon > 0$. Alors pour tout $(x,y) \in B_o(a,\varepsilon)$, on a $|a_1 - x| < \|a - (x,y)\| < \varepsilon$, soit encore $|p_1((x,y)) - p_1(a)| < \varepsilon$, de sorte que p_1 est continue en a, et donc sur $\mathbf{R}^2$.

On prouverait de même que $p_2 : \left| \begin{array}{ccc} \mathbf{R}^2 & \longrightarrow & \mathbf{R} \\ (x,y) & \longmapsto & y \end{array} \right.$ est continue.

▸ La fonction norme, $\|\cdot\| : x \mapsto \|x\|$ est continue sur $\mathbf{R}^2$.
En effet, soit $a \in \mathbf{R}^2$, et soit $\varepsilon > 0$. Alors pour $\|x - a\| < \varepsilon$, on a

$$\big|\|x\| - \|a\|\big| \leqslant \|x - a\| < \varepsilon.$$

Ineg triang
L'inégalité triangulaire renversée se prouve de la même manière qu'elle se prouve dans $\mathbf{R}$ ou dans $\mathbf{C}$ à partir de l'inégalité triangulaire classique.

▸ Soit $f : (x,y) \mapsto \begin{cases} xy\ln(x^2+y^2) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$

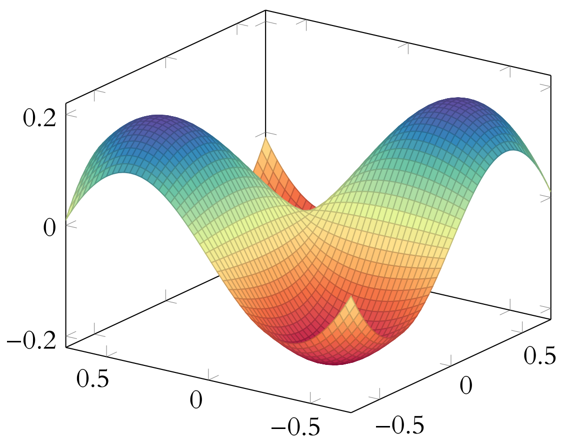
Alors pour $(x,y) \neq (0,0)$, on a

$$|f(x,y)| \leqslant \frac{|xy|}{x^2+y^2}(x^2+y^2)|\ln(x^2+y^2)| \leqslant \frac{1}{2}(x^2+y^2)|\ln(x^2+y^2)|.$$

Rappel
$|xy| \leqslant \frac{1}{2}(x^2+y^2)$.

Mais $t\ln(t) \underset{t\to 0^+}{\longrightarrow} 0$, si bien que pour $\varepsilon > 0$, il existe $\eta > 0$ tel que $|t| < \eta \Rightarrow t|\ln(t)| < \varepsilon$.

Et donc pour $(x,y) \in B_o((0,0),\sqrt{\eta})$, on a $0 \leqslant \sqrt{x^2+y^2} < \sqrt{n}$, si bien que $0 \leqslant x^2+y^2 < \eta$, et donc $|f(x,y) - f(0,0)| < \varepsilon$, si bien que f est continue en 0.



Proposition 33.14 : *Soit A une partie de $\mathbf{R}^2$. Et soient $f, g : A \to \mathbf{R}$ deux fonctions continues sur A. Alors*

1. *$\forall \lambda \in \mathbf{R}$, $\lambda f + g$ est continue sur A.*
2. *$f \times g$ est continue sur A.*
3. *Si g ne s'annule pas sur A, alors $\dfrac{1}{g}$ et $\dfrac{f}{g}$ sont continues sur A.*
4. *Si $\varphi : I \to \mathbf{R}$ est une fonction continue sur une partie I de $\mathbf{R}$ et que f est à valeurs dans I, alors $\varphi \circ f : A \to \mathbf{R}$ est continue sur A.*

Démonstration. Dans toute la suite, on fixe $a \in A$, et on prouve les continuités annoncées en a.

1. Soit $\varepsilon > 0$, et soit $\lambda \in \mathbf{R}^*$. Alors il existe $\eta_1 > 0$ et $\eta_2 > 0$ tels que pour $x \in A$, $\|x - a\| < \eta_1 \Rightarrow |f(x) - f(a)| < \frac{\varepsilon}{2\lambda}$, et $\|x - a\| < \eta_2 \Rightarrow |g(x) - g(a)| < \frac{\varepsilon}{2}$.

Alors en posant $\eta = \min(\eta_1, \eta_2)$, pour $x \in A \cap B_o(a, \eta)$,

$$|(\lambda f + g)(x) - (\lambda f + g)(a)| \leqslant |\lambda||f(x) - f(a)| + |g(x) - g(a)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Donc $\lambda f + g$ est continue en a.

2. Soit $\varepsilon > 0$, et soient alors η_1 et η_2 tels que $\forall x \in A$,

$$\|x - a\| < \eta_1 \Rightarrow |f(x) - f(a)| < \varepsilon \textbf{ et } \|x - a\| < \eta_2 \Rightarrow |g(x) - g(a)| < \frac{\varepsilon}{|f(a)| + \varepsilon}.$$

Notons alors que pour $\|x - a\| < \eta_1$, on a $|f(x)| \leqslant |f(a)| + \varepsilon$.
Alors pour $\|x - a\| < \min(\eta_1, \eta_2)$, on a

$$\begin{aligned} |f(x)g(x) - f(a)g(a)| &= |f(x)\,(g(x) - g(a)) + g(a)\,(f(x) - f(a))\,| \\ &\leqslant |f(x)||g(x) - g(a)| + |g(a)||f(x) - f(a)| \\ &\leqslant \varepsilon + |g(a)|\varepsilon \leqslant \varepsilon(1 + |g(a)|). \end{aligned}$$

Et donc fg est continue en a.

3. Pour $x \in A$, on a $\dfrac{1}{g(x)} - \dfrac{1}{g(a)} = \dfrac{g(a) - g(x)}{g(x)g(a)}$.
Mais puisque $g(a) \neq 0$, pour $\varepsilon = \frac{|g(a)|}{2}$, il existe $\eta_1 > 0$ tel que pour $\|x - a\| < \eta_1$, $|g(x) - g(a)| < \varepsilon \Rightarrow |g(x)| > \frac{|g(a)|}{2}$.
Soit alors $\varepsilon > 0$ et soit $\eta_2 > 0$ tel que pour $\|x - a\| < \eta_2$, $|g(x) - g(a)| < \varepsilon$.
Alors pour $\|x - a\| < \min(\eta_1, \eta_2)$,

$$\left|\frac{1}{g(x)} - \frac{1}{g(a)}\right| = \frac{|g(a) - g(x)|}{|g(x)g(a)|} < \frac{2}{|g(a)|^2}\varepsilon$$

et donc $\frac{1}{g}$ est continue en a.
La continuité de $\frac{f}{g}$ découle alors du point 2.

4. Notons que φ est en particulier continue en $f(a)$.
Soit $\varepsilon > 0$, et soit $\eta_1 > 0$ tel que pour $t \in I$, $|t - f(a)| < \eta_1 \Rightarrow |\varphi(t) - \varphi(f(a))| < \varepsilon$.
Soit alors $\eta > 0$ tel que pour $x \in A$, $\|x - a\| < \eta \Rightarrow |f(x) - f(a)| < \eta_1$.
Alors pour $\|x - a\| < \eta$, on a $|f(x) - f(a)| < \eta_1$, et donc $|\varphi(f(x)) - \varphi(f(a))| < \varepsilon$.
Donc $\varphi \circ f$ est continue en a.

□

Ça suffit ?
Normalement vous commencez à avoir l'habitude de ce type de raisonnement : si on arrive à rendre $|f(x) - f(a)|$ inférieur à $C\varepsilon$, avec C une constante (indépendante de x), et ce pour tout ε, alors quitte à remplacer ε par $\frac{\varepsilon}{C}$, on peut obtenir (pour x suffisamment proche de a), $|f(x) - f(a)| < \varepsilon$.

Remarques. ▶ Si on avait disposé d'une caractérisation séquentielle de la continuité[7], alors on aurait pu prendre exactement les mêmes preuves que celles qui ont été données dans le cas des fonctions de $\mathbf{R}$ dans $\mathbf{R}$.
Cela dit, toutes les preuves que nous venons de donner, qui nécessitent un peu plus de «découpage d'ε» se transposent mot à mot au cas des fonctions réelles, en remplaçant les normes par des valeurs absolues.
▶ On déduit de ce qui précède que les fonctions continues sur A sont un sous-anneau et un sous-espace vectoriel de $\mathcal{F}(A, \mathbf{R})$.

[7] Et elle existe, là encore vous en parlerez l'an prochain, mais elle nécessite de parler de limite de suites à valeurs dans $\mathbf{R}^2$.
Avec un peu d'imagination, vous devriez pouvoir deviner ce qu'on a envie d'appeler une suite convergente à valeurs dans $\mathbf{R}^2$, et énoncer un théorème de caractérisation séquentielle de la continuité.

Définition 33.15 – Une fonction **polynomiale** sur une partie A de $\mathbf{R}^2$ est une fonction pour laquelle il existe deux entiers naturels p, q et des réels $(\lambda_{i,j})_{\substack{0 \leqslant i \leqslant p \\ 0 \leqslant j \leqslant q}}$ tels que pour tout $(x, y) \in A$,

$$f(x, y) = \sum_{i=0}^{p} \sum_{j=0}^{q} \lambda_{i,j} x^i y^j.$$

Exemple 33.16

$f : (x, y) \mapsto 3x^3(y^2 + 1) + x^2 y + (y^2 + y + 1)x^7$ est une fonction polynomiale.

Il est facile de prouver que l'ensemble des fonctions polynomiales sur A est un sous-anneau et un sous-espace vectoriel de l'ensemble des fonctions définies sur A, et qu'une famille génératrice en est la famille des $(x, y) \mapsto x^i y^j$, pour $(i, j) \in \mathbf{N}^2$.

Exercice
Prouver que cette famille est une base de l'ensemble des fonctions polynomiales lorsque $A = \mathbf{R}^2$.

Proposition 33.17 : *Une fonction polynomiale est continue.*

Démonstration. Une fonction polynomiale est une combinaison linéaire de produits des fonctions $p_1 : (x, y) \mapsto x$ et $p_2 : (x, y) \mapsto y$ dont nous avons déjà dit qu'elles sont continues. □

La proposition suivante peut parfois aider à déterminer si une partie de $\mathbf{R}^2$ est ouverte ou fermée.

⚠ Attention !
On demande à ce que f soit définie sur $\mathbf{R}^2$ tout entier.

Proposition 33.18 : *Soit $f : \mathbf{R}^2 \to \mathbf{R}$ une fonction continue.*

1. *pour tout $a \in \mathbf{R}$, $f^{-1}(]-\infty, a[)$ et $f^{-1}(]a, +\infty[)$ sont des ouverts de $\mathbf{R}^2$.*
2. *pour tout $a \in \mathbf{R}$, $f^{-1}(]-\infty, a])$ et $f^{-1}([a, +\infty[)$ sont des fermés de $\mathbf{R}^2$.*

Démonstration.

1. Soit $a \in \mathbf{R}$, et soit $x \in f^{-1}(]-\infty, a[)$, c'est-à-dire tel que $f(x) < a$.
   Prenons alors $\varepsilon = a - f(x) > 0$.
   Alors il existe $\eta > 0$ tel que pour tout $y \in \mathbf{R}^2$, $\|x - y\| < \eta \Rightarrow |f(x) - f(y)| < \varepsilon$.
   Et en particulier, pour $y \in B_o(x, \eta)$, $f(y) < \varepsilon + f(x) = a$.
   Ceci signifie donc que $B_o(x, \eta) \subset f^{-1}(]-\infty, a[)$.
   Et donc $f^{-1}(]-\infty, a[)$ est ouvert.

   Le même raisonnement[8] prouve que $f^{-1}(]a, +\infty[)$ est ouvert.
2. Le complémentaire de $B = f^{-1}(]-\infty, a])$ est $f^{-1}(]a, +\infty[)$, qui est ouvert. Donc B est fermé.

□

[8] Par exemple en appliquant ce qui vient d'être prouvé à la fonction $-f$.

On en déduit notamment que $f^{-1}(\{a\}) = f^{-1}(-\infty, a]) \cap f^{-1}([a, +\infty[)$ est fermé car intersection de deux fermés, que $f^{-1}(]a, b[) = f^{-1}(]a, +\infty[) \cap f^{-1}(]-\infty, b[)$ est ouvert car intersection de deux ouverts, que $f^{-1}(\mathbf{R}^*) = f^{-1}(]-\infty, 0[) \cup f^{-1}(]0, +\infty[)$ est ouvert, etc.

On remarquera en particulier que les lignes de niveau d'une fonction continue sur $\mathbf{R}^2$ sont des fermés de $\mathbf{R}^2$.

Exemple 33.19

$\{(x, y) \in \mathbf{R}^2 \mid x^2 + 2y^2 < 3\}$ est ouvert car image réciproque de $]-\infty, 3[$ par la fonction continue[9] $(x, y) \mapsto x^2 + 2y^2$.
Et de même, $\{(x, y) \in \mathbf{R}^2 \mid x^2 + 2y^2 \leqslant 3\}$ est fermé.

[9] Car polynomiale.

33.4 Dérivées partielles, fonctions de classe $\mathscr{C}^1$

Dans toute cette partie, $\mathscr{O}$ désigne un **ouvert** de $\mathbf{R}^2$.

33.4.1 Dérivées partielles

Définition 33.20 – Soit $f : \mathcal{O} \to \mathbf{R}$, et soit $a = (a_1, a_2) \in \mathcal{O}$.
Soit alors $r > 0$ tel que $B_o(a, r) \subset \mathcal{O}$.
Alors la première fonction partielle de f en a est

$$f_1 : \left| \begin{array}{ccc}]a_1 - r, a_1 + r[& \longrightarrow & \mathbf{R} \\ x & \longmapsto & f(x, a_2) \end{array} \right.$$

De même, la seconde fonction partielle de f en a est

$$f_2 : \left| \begin{array}{ccc}]a_2 - r, a_2 + r[& \longrightarrow & \mathbf{R} \\ y & \longmapsto & f(a_1, y) \end{array} \right. .$$

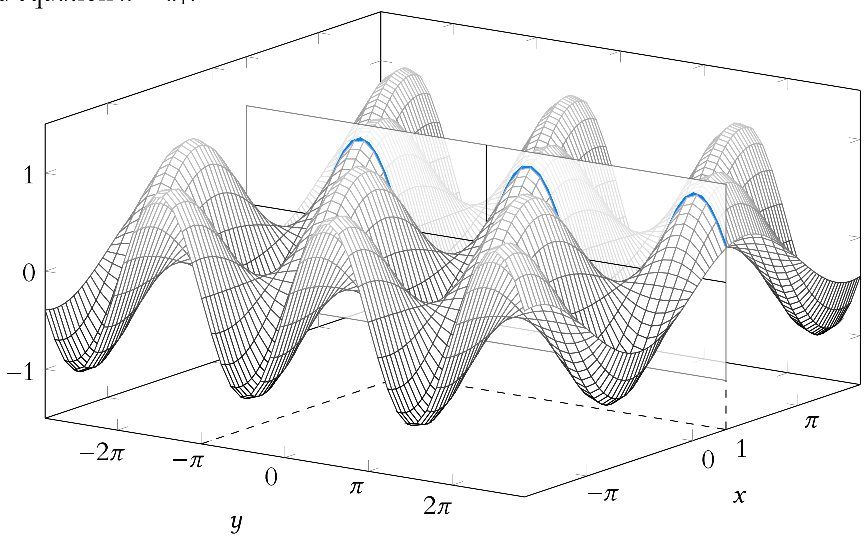
Autrement dit
f_1 est la fonction d'une seule variable obtenue en faisant varier x et en laissant y égal à a_2. Et de même pour a_1.

Remarques. ▸ Il est facile de prouver que si f est continue en a, alors f_1 est continue en a_1 et f_2 est continue en a_2.
▸ L'ensemble de définition des fonctions partielles n'est pas clairement défini puisqu'il n'y a pas unicité du r de l'énoncé. Mais ce n'est pas grave, seul leur comportement en a_1/a_2 va nous intéresser dans la suite, l'essentiel est donc qu'elles soient définies sur un voisinage de a_1/a_2.
Notons tout de même que si $B_o(a, r) \subset \mathcal{O}$, alors pour tout $x \in]a_1 - r, a_1 + r[$,
$\|(x, a_2) - (a_1, a_2)\| = |x - a_1| < r$, si bien que $(x, a_2) \in \mathcal{O}$, et donc $f(x, a_2)$ est bien défini.

Exercice
Le prouver.

Exemple 33.21

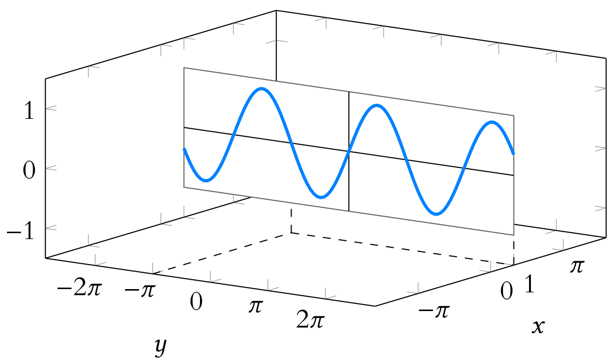
Si $a = (a_1, a_2)$. f_2 est alors la fonction $t \mapsto f(a_1, t)$.
La courbe de f_2 est alors obtenue en faisant l'intersection du graphe de f avec le plan d'équation $x = a_1$.

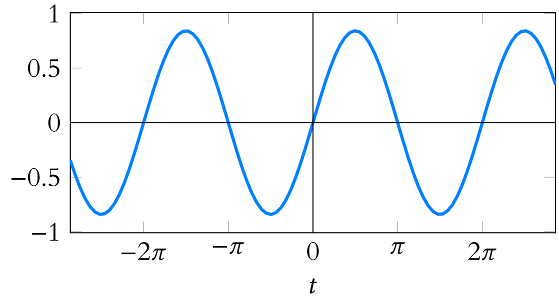


Ici, on considère la fonction

$$f(x, y) = \sin(x)\sin(y)$$

On souhaite tracer la courbe de la seconde fonction partielle de f en $a = (1, -\pi)$. Pour cela, on s'intéresse à l'intersection du graphe de f et du plan d'équation $x = 1$.





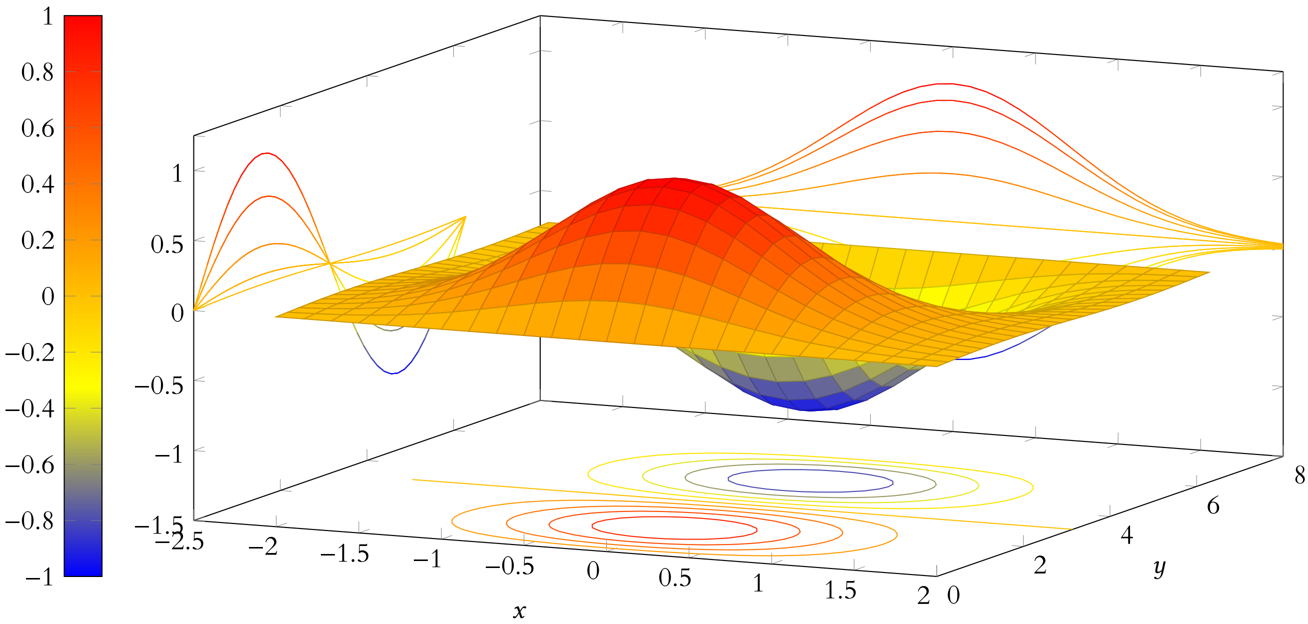


Figure 33.4 – Sur le plan du fond sont représentées des premières fonctions partielles (sections par les plans $y = \lambda$), sur le plan de gauche des secondes fonctions partielles, et sur le plan du bas, des lignes de niveau.

Définition 33.22 – Soit $f : \mathcal{O} \to \mathbf{R}$, et soit $a = (a_1, a_2) \in \mathcal{O}$. On dit que f admet une **première dérivée partielle** en a (ou dérivée par rapport à x) si f_1 est dérivable en a_1.

C'est-à-dire si $\displaystyle\lim_{h\to 0} \frac{f(a_1+h, a_2) - f(a_1, a_2)}{h}$ existe et est finie.

Lorsque c'est le cas, on note $\dfrac{\partial f}{\partial x}(a)$ cette limite.

De même, on définit la **seconde dérivée partielle de** f **en** a (ou dérivée par rapport à y) comme étant, si elle existe, la dérivée de f_2 en a_2.

On la note alors $\dfrac{\partial f}{\partial y}(a)$.

En pratique

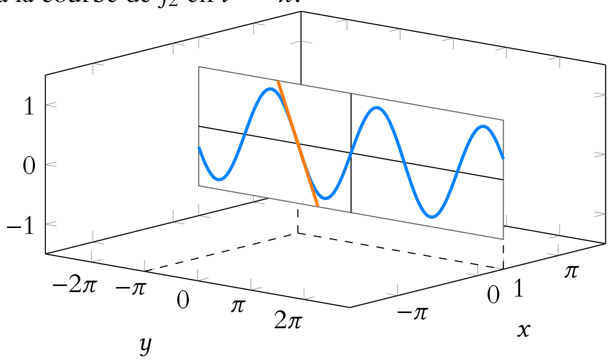
Pour étudier l'existence, et le cas échéant calculer $\frac{\partial f}{\partial x}(a)$, il suffit de dériver f_1, qui est une fonction de **la seule** variable x. Toutes les règles sur la dérivation des fonctions d'une variable sont valables, et il n'est pas nécessaire de revenir au taux d'accroissement.

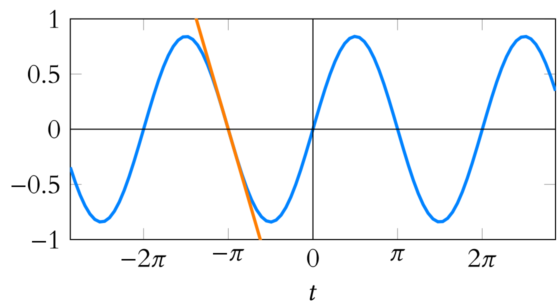
Exemple 33.23

Reprenons le cas de $f(x, y) = \sin(x)\sin(y)$ et $a = (1, -\pi)$.
Alors la seconde fonction partielle de f en a est $f_2 : t \mapsto f(1, t) = \sin(1)\sin(t)$.

Sa dérivée est $t \mapsto \sin(1)\cos(t)$ et donc $\dfrac{\partial f}{\partial y}(1, -\pi) = \sin(1)\cos(\pi) = -\sin(1)$.

Graphiquement, cela signifie que $-\sin(1)$ est le coefficient directeur de la tangente à la courbe de f_2 en $t = -\pi$.





Concrètement, pour calculer $\dfrac{\partial f}{\partial x}$, il suffit de considérer y comme une constante, et d'appliquer les règles de dérivation[10] usuelles.

[10] D'une somme, d'un produit, etc

Exemples 33.24

- Soit $f : \mathbf{R}^2 \to \mathbf{R}$ définie par $f(x,y) = xy + 2x^3 + 3y^2$.
Alors f admet des dérivées partielles, et
$$\frac{\partial f}{\partial x}(x,y) = y + 6x^2,\ \frac{\partial f}{\partial y}f(x,y) = x + 6y.$$
- Soit $g : \mathbf{R}^2 \to \mathbf{R}$ définie par $g(x,y) = xe^{x^2+y^2}$. Alors g admet des dérivées partielles
$$\frac{\partial g}{\partial x}(x,y) = e^{x^2+y^2} + 2xe^{x^2+y^2},\ \frac{\partial g}{\partial y}g(x,y) = 2xye^{x^2+y^2}.$$

Remarque. Cette notation a un inconvénient, c'est qu'elle sous-entend que les deux variables dont dépend notre fonction s'appellent x et y, dans cet ordre.
Et si on décide que notre fonction f dépend des variables r et θ (pensez à des coordonnées polaires...), il faudrait tout de même noter $\dfrac{\partial f}{\partial x}f(r,\theta)$ lorsqu'on dérive par rapport à r ??
En pratique, si le cas se présente, on notera $\dfrac{\partial f}{\partial r}(r,\theta)$ la première dérivée partielle et $\dfrac{\partial f}{\partial \theta}(r,\theta)$ la seconde.

Une notation plus simple[11] consiste à noter $\partial_1 f$ la dérivée de f par rapport à sa première variable, indépendamment du nom de celle-ci, et donc $\partial_2 f$ la dérivée par rapport à la seconde variable.

[11] Et une fois de plus, vous l'utiliserez en seconde année...

⚠ Contrairement au cas des fonctions d'une variable, pour lesquelles la dérivabilité implique la continuité, une fonction de deux variables peut avoir deux dérivées partielles en un point sans être continue en ce point.

Définition 33.25 – Soit $f : \mathcal{O} \to \mathbf{R}$ et $a \in \mathcal{O}$ tels que f possède des dérivées partielles en a.
Alors on appelle **gradient de f en a** et on note $\nabla f(a)$ le vecteur de $\mathbf{R}^2$ défini par
$$\nabla f(a) = \left(\frac{\partial f}{\partial x}(a), \frac{\partial f}{\partial y}(a)\right).$$

⚠ Attention !
Le gradient est un vecteur, c'est donc un couple de nombres, et non un seul !

Exemple 33.26

Soit $f(x,y) = x^2(1+\ln(1+y^2))$. Alors en tout point (x,y), on a
$$\frac{\partial f}{\partial x}(x,y) = 2x(1+\ln(1+y^2)),\ \frac{\partial f}{\partial y}(x,y) = \frac{2yx^2}{1+y^2}$$
et donc $\nabla f(x,y) = \left(2x(1+\ln(y^2)), \frac{2yx^2}{1+y^2}\right) \in \mathbf{R}^2$.

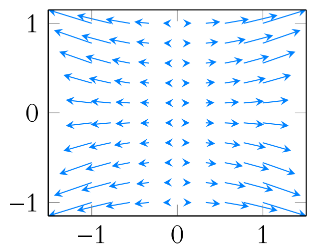


FIGURE 33.5– En chaque point du plan le gradient est un vecteur.

Nous reviendrons plus tard sur l'interprétation géométrique du gradient. Pour l'instant, disons que cela revient à se donner un vecteur en chaque point.

33.4.2 Fonctions de classe $\mathscr{C}^1$

Définition 33.27 – Soit $f : \mathscr{O} \to \mathbf{R}$. On dit que f est de classe $\mathscr{C}^1$ sur $\mathscr{O}$ si pour tout $a \in \mathscr{O}$, $\dfrac{\partial f}{\partial x}(a)$ et $\dfrac{\partial f}{\partial y}(a)$ existent, et que de plus les fonctions $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues sur $\mathscr{O}$.
On note alors $\mathscr{C}^1(\mathscr{O}, \mathbf{R})$ l'ensemble des fonctions de classe $\mathscr{C}^1$ sur $\mathscr{O}$.

Exemple 33.28

Reprenons la fonction $f : (x,y) \mapsto \begin{cases} xy\ln(x^2+y^2) & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$
Alors pour $(x,y) \neq (0,0)$, f possède des dérivées partielles en (x,y), données par

$$\frac{\partial f}{\partial x}(x,y) = y\ln(x^2+y^2) + \frac{2x^2y}{x^2+y^2} \text{ et } \frac{\partial f}{\partial y}(x,y) = x\ln(x^2+y^2) + \frac{2xy^2}{x^2+y^2}.$$

De plus, pour $x \neq 0$, $\dfrac{f(x,0)-f(0,0)}{x} = 0 \underset{x\to 0}{\longrightarrow} 0$, si bien que $\dfrac{\partial f}{\partial x}(0,0) = 0$.
Et de même $\dfrac{\partial f}{\partial y}(0,0) = 0$.
Pour $(x,y) \neq (0,0)$, on a donc

$$\left|\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(0,0)\right| \leqslant |y\ln(x^2+y^2)| + \left|\frac{2x^2y}{x^2+y^2}\right| \leqslant |y\ln(x^2+y^2)| + \frac{1}{2}|x|.$$

Si on suppose de plus que $(x,y) \in B_o((0,0),1)$, alors $y^2 \leqslant x^2+y^2 \leqslant 1$ et donc $2\ln(|y|) \leqslant \ln(x^2+y^2) \leqslant 0$, si bien que $|\ln(x^2+y^2)| \leqslant 2|\ln(|y|)|$.
Mais alors pour $\varepsilon > 0$, puisque $t\ln(t) \underset{t\to 0^+}{\longrightarrow} 0$, il existe $\eta_1 > 0$ tel que pour $|y| < \eta_1$, $|y\ln(|y|)| < \dfrac{\varepsilon}{2}$.
Soit alors $\eta = \min(\eta_1, \varepsilon)$. Alors pour $(x,y) \in B_o((0,0),\eta)$, on a $|x| < \varepsilon$ et $|y| < \eta_1$, si bien que

$$\left|\frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(0,0)\right| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

Et donc $\dfrac{\partial f}{\partial x}$ est bien continue en $(0,0)$.
On prouve de même la continuité de $\dfrac{\partial f}{\partial y}$ en $(0,0)$.

Puisque par ailleurs, par opérations usuelles sur les fonctions continues, il est clair que $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continue sur $\mathbf{R}^2 \setminus \{(0,0)\}$, elles sont continues sur $\mathbf{R}^2$, si bien que f est une fonction de classe $\mathscr{C}^1$ sur $\mathbf{R}^2$.

Proposition 33.29 : *Soient f et g deux fonctions de classe $\mathscr{C}^1$ sur $\mathscr{O}$. Alors*

1. *pour tout $\lambda \in \mathbf{R}$, $\lambda f + g$ est de classe $\mathscr{C}^1$ sur $\mathscr{O}$, et*
$$\frac{\partial(\lambda f+g)}{\partial x} = \lambda\frac{\partial f}{\partial x} + \frac{\partial g}{\partial x} \text{ et } \frac{\partial(\lambda f+g)}{\partial y} = \lambda\frac{\partial f}{\partial y} + \frac{\partial g}{\partial y}$$
2. *$f \times g$ est $\mathscr{C}^1$ sur $\mathscr{O}$ et*
$$\frac{\partial(fg)}{\partial x} = f\frac{\partial g}{\partial x} + \frac{\partial f}{\partial x}g \text{ et } \frac{\partial fg}{\partial y} = f\frac{\partial g}{\partial y} + \frac{\partial f}{\partial y}g.$$
3. *Si g ne s'annule pas, alors $\frac{1}{g}$ est $\mathscr{C}^1$ sur $\mathscr{O}$ et*
$$\frac{\partial}{\partial x}\left(\frac{1}{g}\right) = -\frac{\frac{\partial g}{\partial x}}{g^2} \text{ et } \frac{\partial}{\partial y}\left(\frac{1}{g}\right) = -\frac{\frac{\partial g}{\partial y}}{g^2}.$$

⚠ Attention !
N'apprenez en aucun cas ces formules !
1) Elles sont là essentiellement pour vous dire qu'elles existent.
2) C'est du bon sens : dériver par rapport à x, c'est fixer y, puis dériver une fonction d'une seule variable, pour lesquelles on dispose alors de formules pour la dérivée d'un produit/d'un quotient/d'une composée.

4. *Si f est à valeurs dans un intervalle I, et si $\varphi : I \to \mathbf{R}$ est une fonction de classe $\mathscr{C}^1$, alors $\varphi \circ f$ est de classe $\mathscr{C}^1$, et*

$$\frac{\partial(\varphi \circ f)}{\partial x} = \frac{\partial f}{\partial x} \times \varphi' \circ f \text{ et } \frac{\partial(\varphi \circ f)}{\partial y} = \frac{\partial f}{\partial y} \times \varphi' \circ f.$$

Démonstration. Prouvons par exemple le point 2) : soit $(x_0, y_0) \in \mathscr{O}$.
Alors la fonction $x \mapsto (fg)(x, y_0)$ est dérivable en x_0 puisque les fonctions $x \mapsto f(x, y_0)$ et $x \mapsto g(x, y_0)$ le sont.
Et alors sa dérivé en x_0 est $x \mapsto f(x_0, y_0)\dfrac{\partial g}{\partial x}(x_0, y_0) + \dfrac{\partial f}{\partial x}(x_0, y_0)g(x_0, y_0)$.
Donc $\dfrac{\partial(fg)}{\partial x}$ est définie sur $\mathscr{O}$ et égale à $f\dfrac{\partial g}{\partial x} + \dfrac{\partial f}{\partial x}g$.
Puisque $f, g, \dfrac{\partial f}{\partial x}$ et $\dfrac{\partial g}{\partial x}$ sont continues sur $\mathscr{O}$, $\dfrac{\partial(fg)}{\partial x}$ l'est aussi.
On raisonne de même pour la seconde dérivée partielle, qui existe et est continue sur $\mathscr{O}$, si bien que fg est $\mathscr{C}^1$ sur $\mathscr{O}$. □

Proposition 33.30 : *Une fonction polynomiale est de classe $\mathscr{C}^1$.*

Démonstration. Si on note $p_1 : (x, y) \mapsto x$, alors $\dfrac{\partial p_1}{\partial x}$ est constante égale à 1, et $\dfrac{\partial p_1}{\partial y}$ est constante égale à 0.
Donc ces deux dérivées partielles sont continues, si bien que p_1 est $\mathscr{C}^1$.
De même, $p_2 : (x, y) \mapsto y$ est $\mathscr{C}^1$, si bien que toute fonction polynomiale, qui est combinaison linéaire de produits de p_1 et p_2 est $\mathscr{C}^1$. □

Exemple 33.31

$f : u \mapsto \|u\|$ est de classe $\mathscr{C}^1$ sur $\mathbf{R}^2 \setminus \{0_{\mathbf{R}^2}\}$.
En effet, $\|(x, y)\| = \sqrt{x^2 + y^2}$.
Or la fonction $(x, y) \mapsto x^2 + y^2$ est polynomiale, donc $\mathscr{C}^1$ sur l'ouvert $\mathbf{R}^2 \setminus \{(0, 0)\}$, à valeurs dans $\mathbf{R}_+^*$.
La fonction $t \mapsto \sqrt{t}$ est $\mathscr{C}^1$ sur $\mathbf{R}_+^*$, et donc par composition de fonctions $\mathscr{C}^1$, $(x, y) \mapsto \sqrt{x^2 + y^2}$ est $\mathscr{C}^1$ sur $\mathbf{R}^2 \setminus \{(0, 0)\}$.

Nous pouvons même calculer ses dérivées partielles :

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

En revanche, bien que f soit définie sur tout $\mathbf{R}^2$, elle ne possède pas de dérivées partielles en $(0, 0)$, par exemple car $f_1 : x \mapsto f(x, 0) = \sqrt{x^2} = |x|$ n'y est pas dérivable.

33.4.3 Formule de Taylor–Young à l'ordre 1

Pour une fonction f d'une variable, de classe $\mathscr{C}^1$ sur un intervalle I, la formule de Taylor–Young peut s'écrire de la manière suivante : pour tout $a \in I$,

$$f(a+h) \underset{h\to 0}{=} f(a) + hf'(a) + o(h).$$

Soit encore, en revenant à la définition de o : pour tout $a \in I$, il existe une fonction ε définie sur un voisinage V de 0, telle que $\lim\limits_{h\to 0} \varepsilon(h) = 0$, telle que pour tout $h \in V$,

$$f(a+h) = f(a) + hf'(a) + h\varepsilon(h).$$

Le résultat qui va suivre est une généralisation de cette formule aux fonctions de deux variables.

> **Théorème 33.32 :** *Soit $f : \mathscr{O} \to \mathbf{R}$ une fonction de classe $\mathscr{C}^1$ sur un ouvert de $\mathbf{R}^2$, et soit $a \in \mathscr{O}$.*
> *Notons $r > 0$ tel que $B_o(a,r) \subset \mathscr{O}$.*
> *Alors il existe une fonction $\varepsilon : B_o(0_{\mathbf{R}^2}, r) \to \mathbf{R}$, vérifiant $\varepsilon(0_{\mathbf{R}^2}) = 0$, continue en $0_{\mathbf{R}^2}$ et telle que pour tout $h = (h_1, h_2) \in B_o(0_{\mathbf{R}^2}, r)$,*
>
> $$f(a+h) = f(a) + \langle \nabla f(a), h\rangle + \|h\|\varepsilon(h) = f(a) + \frac{\partial f}{\partial x}(a)h_1 + \frac{\partial f}{\partial y}(a)h_2 + \|h\|\varepsilon(h).$$

Démonstration. Soit donc $a = (a_1, a_2) \in \mathscr{O}$, et $r > 0$ comme dans l'énoncé.
Notons que quitte à ajouter une constante[12] à f, on peut supposer que $f(a) = 0$.

On définit alors une fonction ε sur $B_o(0_{\mathbf{R}^2}, r)$ en posant, pour tout h tel que $\|h\| < r$:

$$\varepsilon(h) = \begin{cases} \dfrac{1}{\|h\|}\left(f(a+h) - \langle \nabla f(a), h\rangle\right) & \text{si } h \neq (0,0) \\ 0 & \text{sinon} \end{cases}$$

Il est alors évident que ε vérifie l'égalité de l'énoncé, il s'agit donc essentiellement de prouver que ε est continue en $(0,0)$.

Soit donc $\varepsilon_0 > 0$. Par continuité des dérivées partielles de f, il existe η_1, η_2 tels que $\|h\| < \eta_1 \Rightarrow \left\|\frac{\partial f}{\partial x}(a+h) - \frac{\partial f}{\partial x}(a)\right\| < \varepsilon_0$ et $\|h\| < \eta_2 \Rightarrow \left\|\frac{\partial f}{\partial y}(a+h) - \frac{\partial f}{\partial y}(a)\right\| < \varepsilon_0$.
Posons donc $\eta = \min(\eta_1, \eta_2, r)$, et soit $h = (h_1, h_2) \in \mathbf{R}^2$ tel que $\|h\| < \eta$. Alors

$$\begin{aligned}
|f(a+h) - \langle \nabla f(a), h\rangle| &= \left| f(a_1+h_1, a_2+h_2) - \frac{\partial f}{\partial x}(a)h_1 - \frac{\partial f}{\partial y}(a)h_2 \right| \\
&\leqslant \left| f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2) - \frac{\partial f}{\partial x}(a)h_1 + f(a_1, a_2+h_2) - \frac{\partial f}{\partial y}(a)h_2 \right| \\
&\leqslant \left| f(a_1+h_1, a_2+h_2) - f(a_1, a_2+h_2) - \frac{\partial f}{\partial x}(a)h_1 \right| + \left| f(a_1, a_2+h_2) - f(a_1, a_2) - \frac{\partial f}{\partial y}(a)h_2 \right| \\
&\leqslant \left| \int_{a_1}^{a_1+h_1} \frac{\partial f}{\partial x}(u, a_2+h_2)\,du - \int_{a_1}^{a_1+h_1} \frac{\partial f}{\partial x}(a)\,dv \right| + \left| \int_{a_2}^{a_2+h_2} \frac{\partial f}{\partial y}(a_1, v)\,dv - \int_{a_2}^{a_2+h_2} \frac{\partial f}{\partial y}(a)\,dv \right| \\
&\leqslant \left| \int_{a_1}^{a_1+h_1} \left(\frac{\partial f}{\partial x}(u, a_2+h_2) - \frac{\partial f}{\partial x}(a) \right) du \right| + \left| \int_{a_2}^{a_2+h_2} \left(\frac{\partial f}{\partial y}(a_1, v) - \frac{\partial f}{\partial y}(a) \right) dv \right|.
\end{aligned}$$

Mais pour tout u compris entre a_1 et $a_1 + h_1$, on a

$$\|(u, a_2+h_2) - a\| = \|(u - a_1, h_2)\| = \sqrt{(u-a_1)^2 + h_2^2} \leqslant \sqrt{h_1^2 + h_2^2} = \|h\| < \eta$$

si bien que $\left|\frac{\partial f}{\partial x}(u, a_2+h_2) - \frac{\partial f}{\partial x}(a)\right| < \varepsilon_0$.
Et donc par inégalité triangulaire[13],

$$\left| \int_{a_1}^{a_1+h_1} \left(\frac{\partial f}{\partial x}(u, a_2+h_2) - \frac{\partial f}{\partial x}(a) \right) du \right| < |h_1|\varepsilon_0.$$

[12] Ce qui ne change pas ses dérivées partielles, donc ne change pas son gradient.

[13] En commençant éventuellement par remettre les bornes dans le bon sens.

On prouve exactement sur le même principe que

$$\left|\int_{a_2}^{a_2+h_2}\left(\frac{\partial f}{\partial y}(a_1,v)-\frac{\partial f}{\partial y}(a)\right)\,dv\right| \leqslant |h_2|\varepsilon_0.$$

Et donc $|\varepsilon(h)| \leqslant |\frac{|h_1|+|h_2|}{\|h\|}\varepsilon_0 \leqslant 2\varepsilon_0$.

Détails
On a $|h_1| \leqslant \|h\|$.

Autrement dit, nous venons de prouver que pour tout $\varepsilon_0 > 0$, il existe $\eta > 0$ tel que pour $\|h\| < \eta$, $|\varepsilon(h)| < \varepsilon_0$.
Sachant que $\varepsilon((0,0)) = 0$, c'est bien là la définition de la continuité de ε en $(0,0)$.

□

Remarques. Le $\|h\|\varepsilon(h)$ est l'analogue du $o(h)$ que l'on aurait écrit pour une fonction d'une seule variable. D'ailleurs vous emploierez des notations similaires l'an prochain.
Pour (x,y) «proche» de a, en notant $h = (x,y) - (a_1,a_2)$, on a donc

$$f(x,y) \simeq f(a) + \frac{\partial f}{\partial x}(a)(x-a_1) + \frac{\partial f}{\partial y}(a)(y-a_2).$$

Donc au voisinage de a, le graphe de f «ressemble» à l'ensemble des points $(x,y,z) \in \mathbf{R}^3$ tels que $z = f(a) + \frac{\partial f}{\partial x}(a)(x-a_1) + \frac{\partial f}{\partial y}(a)(y-a_2)$.
C'est là l'équation d'un plan de l'espace, que l'on appelle le plan tangent au graphe f en a.

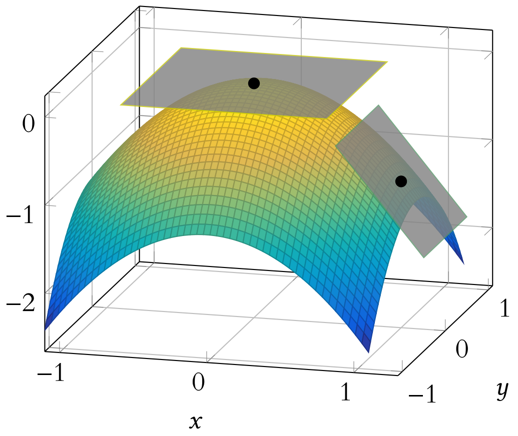


FIGURE 33.6– Deux plans tangents

Corollaire 33.33 – *Si f est $\mathscr{C}^1$ sur un ouvert $\mathscr{O}$, alors f est continue sur $\mathscr{O}$.*

Démonstration. Soit $a \in \mathscr{O}$, soit ε définie sur $B_o(0_{\mathbf{R}^2}, r)$ comme ci-dessus, et soit $\varepsilon_0 > 0$.

Puisque $h \mapsto \|h\|\varepsilon(h)$ est continue en $(0,0)$, car produit de fonctions qui le sont, il existe $\eta_1 > 0$ tel que $\|h\| < \eta_1 \Rightarrow |\|h\|\varepsilon(h)| < \varepsilon_0$.

De plus, par l'inégalité de Cauchy-Schwarz, on a $|\langle \nabla f(a), h\rangle| \leqslant \|\nabla f(a)\| \cdot \|h\|$, si bien que pour $\|h\| \leqslant \dfrac{\varepsilon_0}{1+\|\nabla f(a)\|}$, $|\langle \nabla f(a), h\rangle| \leqslant \varepsilon_0$.
Notons alors $\eta = \min\left(\eta_1, \frac{\varepsilon_0}{1+\|\nabla f(a)\|}\right)$.
Alors pour $x \in B_o(a,\eta)$, si on pose $h = x - a$, alors $\|x\| < \eta$, si bien que

Remarque
Le $1+\dots$ est là uniquement pour ne pas avoir à traiter séparément les cas où $\|\nabla f(a)\| = 0$.

$$|f(x) - f(a)| = |f(a+h) - f(a)| \leqslant |\langle \nabla f(a), h\rangle| + |\|h\|\varepsilon(h)| \leqslant 2\varepsilon_0.$$

Et donc nous avons bien prouvé que f est continue en a. □

33.4.4 Dérivation selon un vecteur

Par définition, les dérivées partielles nous donnent la dérivée des fonctions d'une variable que l'on obtient lorsqu'on se «rapproche» d'un point $a \in \mathbf{R}^2$ parallèlement aux axes (l'axe des abscisses pour la première dérivée partielle, l'axe des ordonnées pour la seconde).
Plus généralement, que peut-on dire de la dérivabilité de $t \mapsto f(a+tu)$ où u est un vecteur quelconque de $\mathbf{R}^2$. C'est-à-dire lorsqu'on se «rapproche» de a suivant une droite non nécessairement verticale ou horizontale ?
C'est ce à quoi répond le théorème suivant, et on constate qu'il suffit de connaître les deux dérivées partielles pour répondre à cette question.

Proposition 33.34 : *Soit $f : \mathcal{O} \to \mathbf{R}$ une fonction de classe $\mathcal{C}^1$ sur un ouvert $\mathcal{O}$ de $\mathbf{R}^2$, et soit $u = (u_1, u_2) \in \mathbf{R}^2$, que l'on supposera non nul[14].*
Soit $a \in \mathcal{O}$, et soit $r > 0$ tel que $B_o(a,r) \subset \mathcal{O}$. Alors la fonction $g : t \mapsto f(a+tu)$ est définie sur $\left]-\frac{\|u\|}{r}, \frac{\|u\|}{r}\right[$, et elle est $\mathcal{C}^1$ sur cet intervalle, avec

$$g' : t \mapsto \langle \nabla f(a+tu), u\rangle.$$

En particulier, $g'(0) = \langle \nabla f(a), u\rangle = \frac{\partial f}{\partial x}(a)u_1 + \frac{\partial f}{\partial y}(a)u_2$.
*La quantité $\langle \nabla f(a), u\rangle$ est appelée **dérivée de f en a selon le vecteur u.***

[14] Le cas $u = (0,0)$ étant trivial.

Remarque. Le graphe de la fonction g est la section du graphe de f par le plan vertical contenant la droite D.
Par exemple, si $a = (0.5, -0.5)$, $u = (-1, 1)$, pour $f(x,y) = x^3 - 4xy^2$, on obtient la figure suivante :

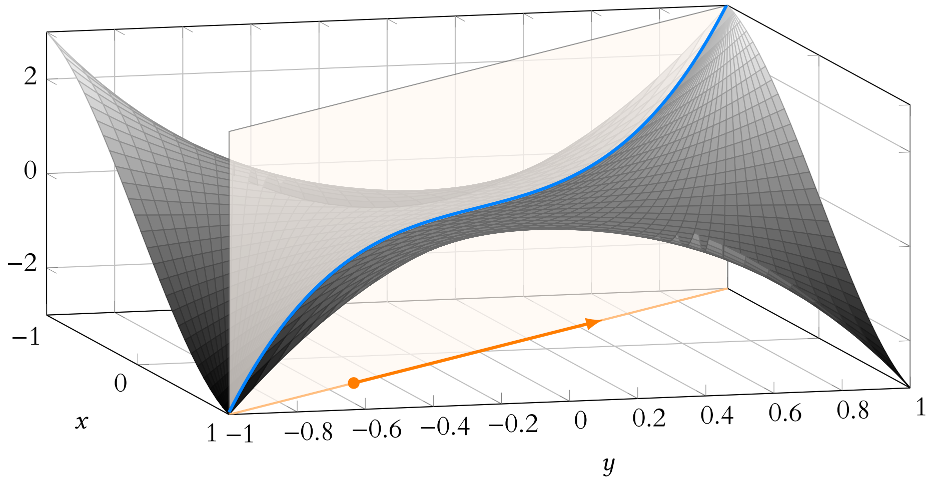


FIGURE 33.7 – En orange la droite D passant par a et de direction u et en bleu, la section du graphe de f par le plan vertical contenant D.

Démonstration. Soient $t, h \in \mathbf{R}$, avec $h \neq 0$, «suffisamment petits[15]».
Prouvons que $\lim\limits_{h\to 0} \frac{g(t+h)-g(t)}{h}$ existe. On a

$$g(t+h) = f(a+(t+h)u) = f((a+tu)+hu).$$

[15] Pour que les quantités en jeu soient bien définies.

Par la formule de Taylor, il existe une fonction ε définie au voisinage de $0_{\mathbf{R}^2}$, nulle en $0_{\mathbf{R}^2}$ et continue en $0_{\mathbf{R}^2}$ telle que

$$g(t+h) = f(a+tu) + \langle \nabla f(a+tu), hu\rangle + \|hu\|\varepsilon(hu) = g(t) + h\langle \nabla f(a+tu), u\rangle + |h|\cdot\|u\|\varepsilon(hu).$$

On en déduit que

$$\frac{g(t+h)-g(t)}{h} = \langle \nabla f(a+tu), u\rangle + \frac{|h|}{h}\|u\|\varepsilon(hu).$$

Mais ε est continue en $0_{\mathbf{R}^2}$, et $\varepsilon(0) = 0$, donc $\lim\limits_{h\to 0} |\varepsilon(h)| = 0$, si bien que

$$\lim_{h\to 0} \frac{|h|}{h}\varepsilon(hu) = 0.$$

Ainsi,

$$\lim_{h\to 0} \frac{g(t+h)-g(t)}{h} = \langle \nabla f(a+tu), u\rangle.$$

On en déduit que g est dérivable en t et

$$g'(t) = \langle \nabla f(a+tu), u\rangle = \frac{\partial f}{\partial x}(a+tu)u_1 + \frac{\partial f}{\partial y}(a+tu)u_2.$$

Remarque
En pratique, et faute de définitions adéquates, il nous faudrait epsiloner pour prouver ceci.

Signel
$\frac{|h|}{h}$ vaut 1 si $h > 0$ et -1 si $h < 0$.

Enfin, puisque $\dfrac{\partial f}{\partial x}$ et $\dfrac{\partial f}{\partial y}$ sont continues, $t \mapsto \dfrac{\partial f}{\partial x}(a+tu)$ et $t \mapsto \dfrac{\partial f}{\partial y}(a+tu)$ sont continues sur $\left]-\frac{\|u\|}{r}, \frac{\|u\|}{r}\right[$, et donc g' est continue sur $\left]-\frac{\|u\|}{r}, \frac{\|u\|}{r}\right[$. Par conséquent, g est de classe $\mathscr{C}^1$ sur $\left]-\frac{\|u\|}{r}, \frac{\|u\|}{r}\right[$. □

On note que lorsqu'on multiplie un vecteur u par λ (et donc qu'on multiplie sa norme par $|\lambda|$), alors la dérivée de $t \mapsto f(a+tu)$ est également multipliée par λ.
Pourtant, u et λu ont la même direction, et on a envie que cette dérivée représente «la variation de f dans la direction de u».
Afin de se débarrasser de cette dépendance en la norme de u, on peut considérer un vecteur u unitaire[16], et alors la dérivée directionnelle de f en a dans la direction de u est la dérivée de $t \mapsto f(a+tu)$, c'est-à-dire $\langle \nabla f(a), u\rangle$.
Notons que pour $u \neq (0,0)$, il existe encore deux vecteurs unitaires de même direction[17] que u, à savoir $\frac{u}{\|u\|}$ et son opposé.

[16] Rappelons que cela signifie que $\|u\| = 1$.

[17] c'est-à-dire colinéaires à u.

Par l'inégalité de Cauchy-Schwarz, pour u unitaire, on a donc

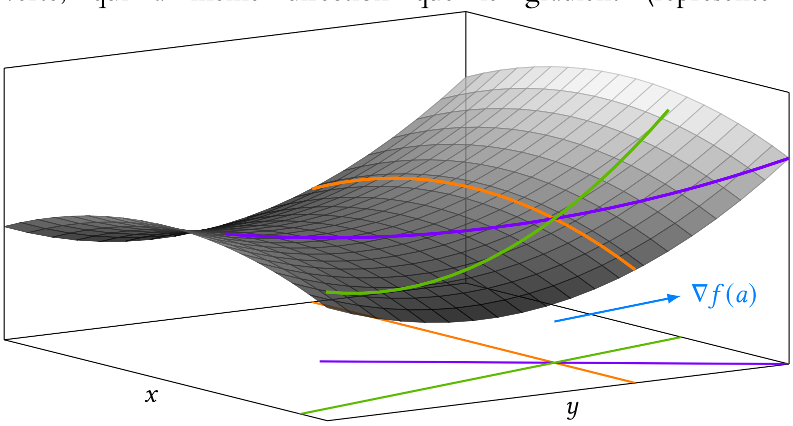
$$|\langle \nabla f(a), u\rangle \leqslant \|\nabla f(a)\| \underbrace{\|u\|}_{=1} = \|\nabla f(a)\|.$$

Et de plus, il y a égalité si et seulement si u et $\nabla f(a)$ sont colinéaires.
Autrement dit, $\langle \nabla f(a), u\rangle$ est maximal si et seulement si u est colinéaire à $\nabla f(a)$.
Donc la direction de $\nabla f(a)$ indique la direction de «la plus grande pente», et son sens est celui de la montée.

Détails
Dans Cauchy-Schwarz, on a égalité $\langle x, y\rangle = \|x\|\|y\|$ (sans valeur absolue sur le produit scalaire) si et seulement si x et y sont colinéaires et de même sens, c'est-à-dire s'il existe $\lambda \in \mathbf{R}_+$ tel que $x = \lambda y$.

Exemple 33.35

Pour la fonction suivante sont représentées plusieurs courbes dans différentes directions, passant toutes par le même point.
Celle de ces courbes qui possède la pente la plus importante est la verte, qui a même direction que le gradient (représenté en bleu).



Remarque : sur cet exemple, on constate que le sens du gradient indique la direction de la «montée», et non de la «descente». Ce fait est toujours vérifié.

Pour reprendre l'exemple où notre fonction représente l'altitude d'un point du globe, en tout point la direction du gradient indique la direction dans laquelle il faut marcher pour gagner le plus rapidement de l'altitude. Et une goutte d'eau posée en un point a partira dans la direction de $-\nabla f(a)$, la direction où elle descend le plus vite.

33.4.5 Dérivée d'une composée

Donnons nous deux fonctions d'une variable $x : t \mapsto x(t)$ et $y : t \mapsto y(t)$, définies sur un intervalle I, ainsi qu'une fonction $f : \mathbf{R}^2 \to \mathbf{R}$.
On peut alors considérer la fonction $g : t \mapsto f(x(t), y(t))$, qui est donc une fonction de I dans $\mathbf{R}$. Lorsque t parcourt I, le point de coordonnées $(x(t), y(t))$ parcourt une courbe $\mathscr{C}$ tracée dans $\mathbf{R}^2$.

Remarque
Une telle courbe est appelée un arc paramétré, leur étude a désormais complètement disparu des programmes de prépa.

Et donc le point de coordonnées $(x(t), y(t), g(t))$ se «promène» le long d'une courbe tracée sur le graphe de f.

Prenons par exemple $\begin{cases} x(t) = \cos^3 t \\ y(t) = \sin^3 t \end{cases}$.

La courbe parcourue par le point $(x(t), y(t))$ est appelée une astroïde et a l'allure ci-contre.

Lorsque t varie, le point de coordonnées $(x(t), y(t), f(x(t), y(t)))$ parcourt donc une courbe tracée sur la surface représentant f.

La fonction $g : t \mapsto f(x(t), y(t))$ représente l'altitude d'un point parcourant cette courbe au cours du temps.

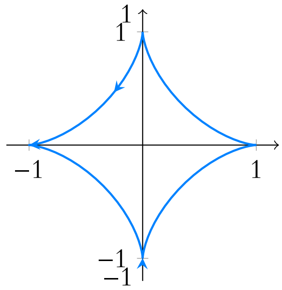
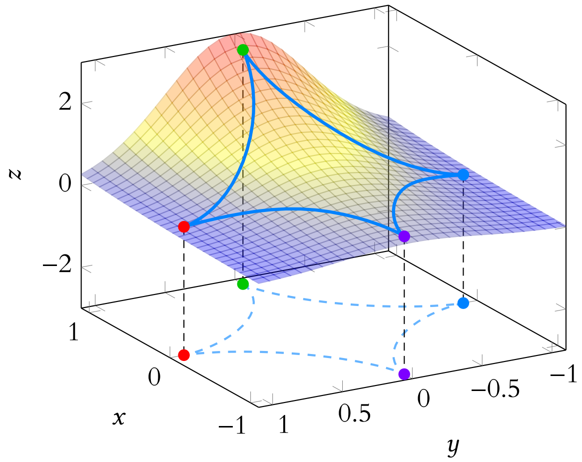
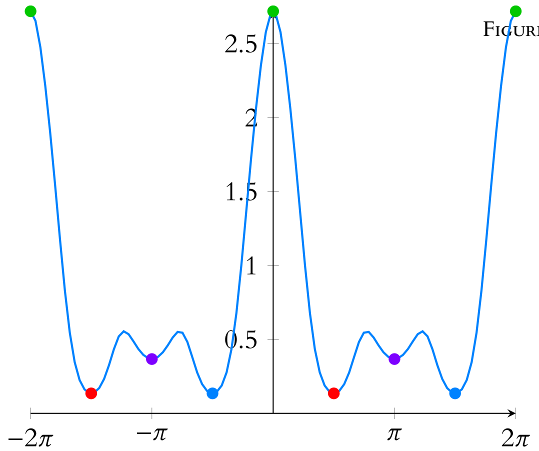


FIGURE 33.8– L'astroïde.





Ce que dit la proposition suivante, c'est qu'alors la dérivée de g en t_0 est égale à la dérivée directionnelle de f en $(x(t_0), y(t_0))$ dans la direction de $(x'(t_0), y'(t_0))$.

Or ce vecteur n'est rien d'autre qu'un vecteur directeur de la tangente à la courbe $\mathscr{C}$ en $(x(t_0), y(t_0))$.

Proposition 33.36 : *Soit $\mathcal{O}$ un ouvert de $\mathbf{R}^2$, et $f : \mathcal{O} \to \mathbf{R}$ une fonction de classe $\mathscr{C}^1$. Soient $x, y : t \mapsto \mathbf{R}$ deux fonctions de classe $\mathscr{C}^1$ sur un intervalle ouvert I de $\mathbf{R}$ telles que $\forall t \in I$, $(x(t), y(t)) \in \mathcal{O}$.*
Alors la fonction $g : t \mapsto f(x(t), y(t))$ est de classe $\mathscr{C}^1$ sur I, et

$$\forall t \in I,\ g'(t) = \frac{\partial f}{\partial x}(x(t), y(t))x'(t) + \frac{\partial f}{\partial y}(x(t), y(t))y'(t).$$

Démonstration. Soit $t_0 \in I$ fixé.

Par la formule de Taylor-Young, il existe ε définie sur une boule ouverte de rayon $r > 0$ centrée en $0_{\mathbf{R}^2}$, nulle en $0_{\mathbf{R}^2}$ et continue en $0_{\mathbf{R}^2}$ telle que pour tout $u \in B_o(0_{\mathbf{R}^2}, r)$,

$$f((x(t_0), y(t_0)) + u) = f(x(t_0), y(t_0)) + \langle \nabla f(x(t_0), y(t_0)), u\rangle + \|u\|\varepsilon(u).$$

Pour $h \in I$, notons $u_h = (x(t_0 + h) - x(t_0), y(t_0 + h) - y(t_0))$, de sorte que

$$\begin{aligned} f(x(t_0 + h), y(t_0 + h)) &= f((x(t_0), y(t_0)) + u_h) \\ &= f(x(t_0), y(t_0)) + \langle \nabla f(x(t_0), y(t_0)), u_h\rangle + \|u_h\|\varepsilon(u_h). \end{aligned}$$

On a donc[18] $x(t_0+h) \underset{h\to 0}{=} x(t_0)+x'(t_0)h+o(h)$ et de même $y(t_0+h) \underset{h\to 0}{=} y(t_0)+y'(t_0)h+o(h)$.

Autrement dit, il existe deux fonctions $\varepsilon_1, \varepsilon_2$, nulles en 0, et de limites nulles en 0 telles que

$$x(t_0 + h) - x(t_0) = x'(t_0)h + h\varepsilon_1(h) \text{ et } y(t_0 + h) - y(t_0) = y'(t_0)h + h\varepsilon_2(h).$$

On a donc $u_h = h \cdot (x'(t_0) + \varepsilon_1(h), y'(t_0) + \varepsilon_2(h))$, si bien que

$$\|u_h\| = |h|\sqrt{(x'(t_0) + \varepsilon_1(h))^2 + (y'(t_0) + \varepsilon_2(h))^2} \underset{h\to 0}{\longrightarrow} 0.$$

Par ailleurs, $\dfrac{\|u_h\|}{h} \underset{h\to 0}{\longrightarrow} \sqrt{x'(t_0)^2 + y'(t_0)^2}$, limite que nous noterons A.

Il existe donc $\eta_1 > 0$ tel que pour $|h| < \eta_1$, $\dfrac{\|u_h\|}{|h|} < A + 1$.

[18] C'est Taylor-Young pour des fonctions d'une variable.

Par ailleurs, pour $\varepsilon_0 > 0$, il existe $\eta_2 > 0$ tel que pour $\|v\| < \eta_2$, $|\varepsilon(v)| < \dfrac{\varepsilon_0}{A+1}$.
Mais alors il existe $\eta_3 > 0$ tel que $|h| < \eta_3 \Rightarrow \|u_h\| < \eta_2$, et donc $|\varepsilon(u_h)| < \dfrac{\varepsilon_0}{A+1}$.
Et donc pour $|h| < \min(\eta_1, \eta_3)$, $\dfrac{\|u_h\|}{|h|}\varepsilon(u_h) < (A+1)\dfrac{\varepsilon_0}{A+1}$.
Autrement dit, $\dfrac{\|u_h\|}{|h|}\varepsilon(u_h) \underset{h\to 0}{\longrightarrow} 0$. Et donc

> **Vocabulaire**
> Il est clair qu'il serait souhaitable d'avoir un vocabulaire plus adapté pour ces preuves, et une notion plus générale de limite pour se contenter de dire «quand h est petit, alors u_h est petit, et donc $\varepsilon(u_h)$ est petit».

$$\left|\frac{g(t_0+h)-g(t_0)}{h} - \left(\frac{\partial f}{\partial x}(x(t_0),y(t_0))x'(t_0) + \frac{\partial f}{\partial y}(x(t_0),y(t_0))y'(t_0)\right)\right|$$
$$= \left|\frac{\partial f}{\partial x}(x(t_0),y(t_0))\varepsilon_1(h) + \frac{\partial f}{\partial y}(x(t_0,y(t_0)))\varepsilon_2(h) + \frac{\|u_h\|}{|h|}\varepsilon(u_h)\right| \underset{h\to 0}{\longrightarrow} 0.$$

C'est donc que g est dérivable en t_0, avec

$$g'(t_0) = \frac{\partial f}{\partial x}(x(t_0),y(t_0))x'(t_0) + \frac{\partial f}{\partial y}(x(t_0),y(t_0))y'(t_0).$$

Il faudrait travailler encore un peu[19] pour établir la continuité de g', mais disons qu'elle vient de la continuité de $\dfrac{\partial f}{\partial x}$, de x, de y, de x' et de y'. □

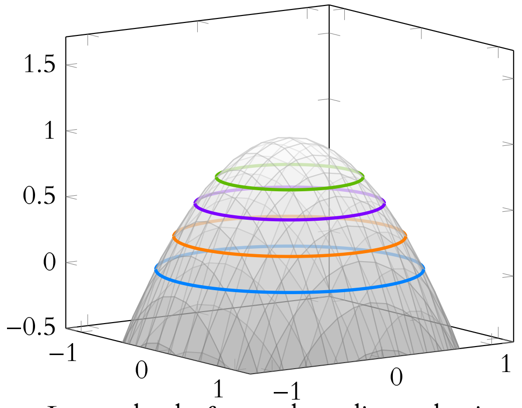
[19] Et notamment epsiloner faute de résultats de cours sur le sujet.

Exemple 33.37 Le gradient est orthogonal aux lignes de niveau.

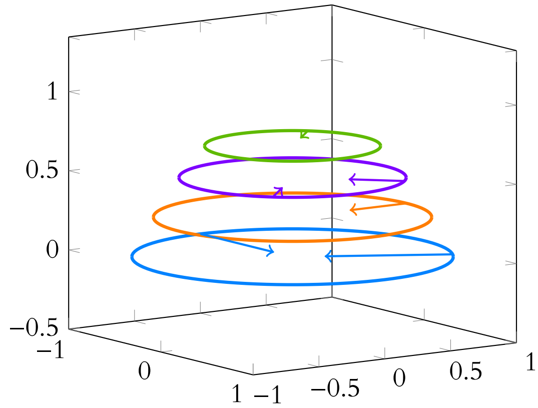
Soit f une fonction de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{O}$, et soit $\lambda \in \mathbf{R}$.
Soit $a \in \mathscr{O}$ tel que $f(a) = \lambda$, c'est-à-dire un point de la ligne de niveau λ de f.
On suppose qu'il existe une boule $B_o(a,r)$, un intervalle ouvert I contenant 0 et deux fonctions x, y de classe $\mathscr{C}^1$ sur I telles que :

1. $a = (x(0), y(0))$
2. $\forall t \in I$, $(x(t), y(t)) \in B_o(a,r)$
3. $\forall (u,v) \in B_o(a,r)$, $f(u,v) = \lambda \Leftrightarrow \exists t \in I,\ (u,v) = (x(t),y(t))$.

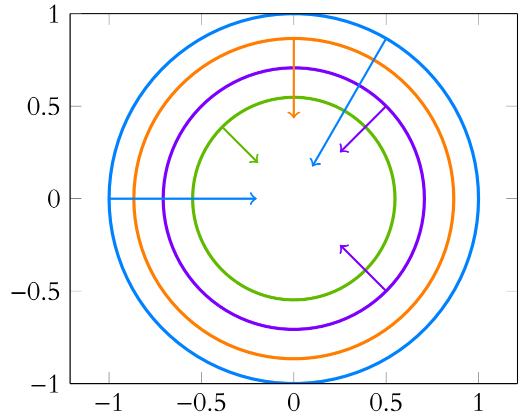
Autrement dit, au voisinage de a, la courbe de niveau λ de f est bien «une courbe».
Alors la tangente à la courbe $(x(t), y(t))$ a pour vecteur directeur $(x'(0), y'(0))$.
Mais par la proposition précédente, la fonction $g : t \mapsto f(x(t), y(t))$ vérifie $g'(0) = \dfrac{\partial f}{\partial x}(a)x'(0) + \dfrac{\partial f}{\partial y}(a)y'(0)$.
Puisque par ailleurs cette fonction est constante sur I égale à λ, on a $g'(0) = 0$.
Autrement dit, le vecteur $\nabla f(a)$ est orthogonal au vecteur $(x'(0), y'(0))$.
Et donc le gradient de f en a est orthogonal à la ligne de niveau λ.



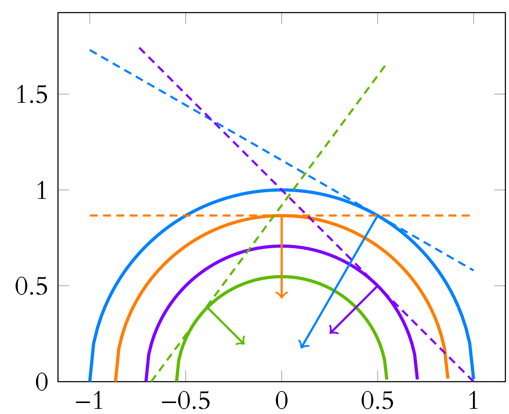
Le graphe de f et quelques lignes de niveau



Quelques lignes de niveau et le gradient en quelques points, vus dans $\mathbf{R}^3$



Lignes de niveau et gradients vus dans $\mathbf{R}^2$.



Les tangentes aux lignes de niveau sont orthogonales au gradient.

Proposition 33.38 (Règle de la chaîne[20]) : *Soit $f : \mathscr{O} \to \mathbf{R}$ une fonction de classe $\mathscr{C}^1$, et soient φ, ψ deux fonctions de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{O}'$ telles que $\forall (u,v) \in \mathscr{O}'$, $(\varphi(u,v), \psi(u,v)) \in \mathscr{O}$.*
Ainsi, la fonction $g : (u,v) \mapsto f(\varphi(u,v), \psi(u,v))$ est bien définie sur $\mathscr{O}'$.
Elle y est alors $\mathscr{C}^1$, et ses dérivées partielles sont données par

$$\frac{\partial g}{\partial x}(u,v) = \frac{\partial f}{\partial x}(\varphi(u,v),\psi(u,v))\frac{\partial \varphi}{\partial x}(u,v) + \frac{\partial f}{\partial y}(\varphi(u,v),\psi(u,v))\frac{\partial \psi}{\partial x}(u,v)$$
$$\frac{\partial g}{\partial y}(u,v) = \frac{\partial f}{\partial x}(\varphi(u,v),\psi(u,v))\frac{\partial \varphi}{\partial y}(u,v) + \frac{\partial f}{\partial y}(\varphi(u,v),\psi(u,v))\frac{\partial \psi}{\partial y}(u,v).$$

Dit autrement, on a donc

$$\partial_1 g(u,v) = \partial_1 f(\varphi(u,v),\psi(u,v))\partial_1\varphi(u,v) + \partial_2 f(\varphi(u,v),\psi(u,v))\partial_1\psi(u,v)$$

$$\partial_2 g(u,v) = \partial_1 f(\varphi(u,v),\psi(u,v))\partial_2\varphi(u,v) + \partial_2 f(\varphi(u,v),\psi(u,v))\partial_2\psi(u,v)$$

[20] *Chain rule derivation* en anglais.

Démonstration. Soit $(u_0, v_0) \in \mathscr{O}'$.
Appliquons la proposition précédente à $x(t) = \varphi(u_0+t, v_0)$ et $y(t) = \psi(u_0+t, v_0)$.
Ce sont alors deux fonctions de classe $\mathscr{C}^1$ au voisinage de 0 avec $x'(t) = \dfrac{\partial \varphi}{\partial x}(u_0+t, v_0)$ et $y'(t) = \dfrac{\partial \psi}{\partial x}(u_0+t, v_0)$.
Donc la fonction $t \mapsto g(u_0+t, v_0)$ est dérivable, de dérivée égale à

$$t \mapsto \frac{\partial f}{\partial x}(\varphi(u_0+t,v_0),\psi(u_0+t,v_0))\frac{\partial \varphi}{\partial x}(u_0,v_0) + \frac{\partial f}{\partial y}(\varphi(u_0+t,v_0),\psi(u_0+t,v_0))\frac{\partial \psi}{\partial x}(u_0+t,v).$$

Et donc sa dérivée en 0, qui est $\dfrac{\partial g}{\partial x}(\varphi(u_0,v_0),\psi(u_0,v_0))$ est

$$\frac{\partial f}{\partial x}(\varphi(u_0,v_0))\frac{\partial \varphi}{\partial x}(u_0,v_0)+\frac{\partial f}{\partial y}(\varphi(u_0,v_0))\frac{\partial \psi}{\partial x}(u_0,v_0).$$

On prouverait de même la formule pour l'autre dérivée partielle.

Enfin la continuité de ces dérivées partielle découle de celles de φ, ψ, et de leurs dérivées partielles. □

Exemple 33.39 Résolution d'une équation aux dérivées partielles à l'aide des coordonnées polaires

Cherchons toutes les fonctions de classe $\mathscr{C}^1$ sur $\mathbf{R}\times\mathbf{R}_+^*$ telles que

$$\forall (x,y)\in\mathbf{R}\times\mathbf{R}_+^*,\ y\frac{\partial f}{\partial x}(x,y)-x\frac{\partial f}{\partial y}(x,y)=f(x,y). \qquad (\star)$$

Pour cela nous allons procéder à un changement de variable, en notant que tout $(x,y)\in\mathbf{R}\times\mathbf{R}_+^*$ s'écrit de manière unique $(x,y)=(r\cos(\theta),r\sin(\theta))$, avec $(r,\theta)\in\mathbf{R}_+^*\times]0,\pi[$
Pour $(r,\theta)\in\mathbf{R}_+^*\times]0,\pi[$, posons $g(r,\theta)=f(r\cos(\theta),r\sin(\theta))$.
Autrement dit, nous sommes dans le cadre d'application de la proposition précédente, avec $\varphi(r,\theta)=r\cos(\theta)$ et $\psi(r,\theta)=r\sin(\theta)$.
Ainsi, g est de classe $\mathscr{C}^1$ sur $\mathbf{R}_+^*\times]0,\pi[$, avec

Polaires
Nous «passons» donc en coordonnées polaires, que vous utilisez en physique. Disons que pour se donner un point dans le plan, il suffit de se donner un complexe, et donc son module et son argument.

$$\begin{aligned}\frac{\partial g}{\partial r}(r,\theta)&=\frac{\partial f}{\partial x}(r\cos(\theta),r\sin(\theta))\frac{\partial\varphi}{\partial r}(r,\theta)+\frac{\partial f}{\partial y}(r\cos(\theta),r\sin(\theta))\frac{\partial\psi}{\partial r}(r,\theta)\\&=\frac{\partial f}{\partial x}(r\cos(\theta),r\sin(\theta))\cos(\theta)+\frac{\partial f}{\partial y}(r\cos(\theta),r\sin(\theta))\sin(\theta)\end{aligned}$$

et

$$\begin{aligned}\frac{\partial g}{\partial \theta}(r,\theta)&=\frac{\partial f}{\partial x}(r\cos(\theta),r\sin(\theta))\frac{\partial\varphi}{\partial \theta}(r,\theta)+\frac{\partial f}{\partial y}(r\cos(\theta),r\sin(\theta))\frac{\partial\psi}{\partial \theta}(r,\theta)\\&=\frac{\partial f}{\partial x}(r\cos(\theta),r\sin(\theta))\times(-r\sin(\theta))+\frac{\partial f}{\partial y}(r\cos(\theta),r\sin(\theta))\times(r\cos(\theta)).\end{aligned}$$

Donc f est solution de $(\star)$ si et seulement si pour tout (r,θ),

$$\frac{\partial g}{\partial\theta}(r,\theta)=-f(r\cos(\theta),r\sin(\theta))=-g(r,\theta). \qquad (\star\star)$$

Nous allons donc plutôt chercher les fonctions qui satisfont cette condition[21].
Donc si g satisfait $(\star\star)$, pour tout $r\in\mathbf{R}_+^*$ $\theta\mapsto g(r,\theta)$ est solution de l'équation différentielle $y'(\theta)=-y(\theta)$.
Donc pour tout $r\in\mathbf{R}_+^*$, il existe $C(r)$ tel que $g(r,\theta)=C(r)e^{-\theta}$.
Mais alors $C: r\mapsto e^{\frac{\pi}{2}}g\left(r,\frac{\pi}{2}\right)$ est une fonction de classe $\mathscr{C}^1$ sur $\mathbf{R}_+^*$.
En effet, sa dérivée est $r\mapsto e^{\frac{\pi}{2}}\dfrac{\partial g}{\partial r}\left(r,\frac{\pi}{2}\right)$ qui est continue puisque $\dfrac{\partial g}{\partial r}$ est continue.

Et inversement, s'il existe C de classe $\mathscr{C}^1$ sur $\mathbf{R}_+^*$, et si $g(r,\theta)=C(r)e^{-\theta}$, alors $\dfrac{\partial g}{\partial\theta}(r,\theta)=-C(r)e^{-\theta}=-g(r,\theta)$.
Donc les fonctions $g:\mathbf{R}_+^*\times]0,\pi[$ qui satisfont $(\star\star)$ sont les $(r,\theta)\mapsto C(r)e^{-\theta}$, avec $C\in\mathscr{C}^1(\mathbf{R}_+^*,\mathbf{R})$.
Ne reste alors qu'à revenir à notre fonction f de départ, en «inversant» le changement de variable d'origine : pour $(x,y)\in\mathbf{R}\times\mathbf{R}_+^*$ et $(r,\theta)\in\mathbf{R}_+^*\times]0,\pi[$, on a $(x,y)=(r\cos(\theta),r\sin(\theta))$ si et seulement si

$$r=\sqrt{x^2+y^2}\ \textbf{et}\ \theta=\operatorname{Arccos}\left(\frac{x}{\sqrt{x^2+y^2}}\right).$$

[21] Qui ne concerne plus qu'une seule dérivée partielle : en gros, il s'agit d'une équation différentielle du premier ordre.

Ainsi, $g(r,\theta) = f(r\cos(\theta), r\sin(\theta)) \Leftrightarrow f(x,y) = g\left(\sqrt{x^2+y^2}, \text{Arccos}\left(\frac{x}{\sqrt{x^2+y^2}}\right)\right)$.
Et donc f est solution de $(\star)$ si et seulement si il existe $C \in \mathscr{C}^1(\mathbf{R}_+^*, \mathbf{R})$ telle que

$$\forall (x,y) \in \mathbf{R}\times\mathbf{R}^*,\ f(x,y) = C\left(\sqrt{x^2+y^2}\right) e^{-\text{Arccos}\left(\frac{x}{\sqrt{x^2+y^2}}\right)}.$$

Notons qu'il y a donc beaucoup de telles fonctions, et que pour garantir l'unicité d'une solution à l'équation, il faut des conditions initiales assez fortes, bien plus que la connaissance de f en un point.

33.5 Extrema et points critiques

33.5.1 Extrema globaux, extrema locaux

Définition 33.40 – Soit $f : A \to \mathbf{R}$ une fonction définie sur une partie A de $\mathbf{R}^2$, et soit $a \in A$.
On dit que f possède un **maximum** (resp. un **minimum**) en a si

$$\forall x \in A,\ f(x) \leqslant f(a) \quad (\text{resp.} f(x) \geqslant f(a).$$

Exemple 33.41

Soit $f : (x,y) \mapsto x^4 - 4x^2y + 5y^2 - 2y + 2$.
Alors pour tout $(x,y) \in \mathbf{R}^2$, $f(x,y) = \left(x^2-2y\right)^2 + y^2 - 2y + 2 = (x^2-2y)^2 + (y-1)^2 + 1 \geqslant 1$.
De plus, on a $f(x,y) = 1$ si et seulement si $\begin{cases} x^2 - 2y = 0 \\ y - 1 = 0 \end{cases} \Leftrightarrow (x,y) = (\sqrt{2}, 1)$ **ou** $(x,y) = (-\sqrt{2}, 1)$.
Donc f possède un minimum atteint uniquement en $(\pm\sqrt{2}, 1)$.

Définition 33.42 – Soit $f : A \to \mathbf{R}$ une fonction définie sur une partie A de $\mathbf{R}^2$, et soit $a \in A$.
On dit que f possède un **maximum local** (resp. un **minimum local**) en a si

$$\exists r > 0,\ \forall x \in A \cap B_o(a,r),\ f(x) \leqslant f(a) \quad (\text{resp.} f(x) \geqslant f(a)).$$

Bien entendu, si f possède un maximum (resp. un minimum) en a, alors f possède un maximum (resp. minimum) local en a, la réciproque étant fausse.

Exemple 33.43

Soit $f : (x,y) \mapsto x^2(1+y)^3 + y^2$.
Alors pour $(x,y) \in B_o((0,0),1)$, on a $|x| \leqslant \|(x,y)\| \leqslant 1$, si bien que $(1+x)^3 \geqslant 0$.
Et donc $f(x,y) \geqslant 0 = f(0,0)$.
Ainsi, f admet un minimum local en $(0,0)$.

Pour autant, il ne s'agit pas d'un minimum global puisque

$$f(1,x) = (1+x)^3 + x^2 \underset{x\to-\infty}{\sim} x^3 \underset{x\to-\infty}{\longrightarrow} -\infty,$$

donc f prend des valeurs strictement inférieures à $0 = f(0,0)$.

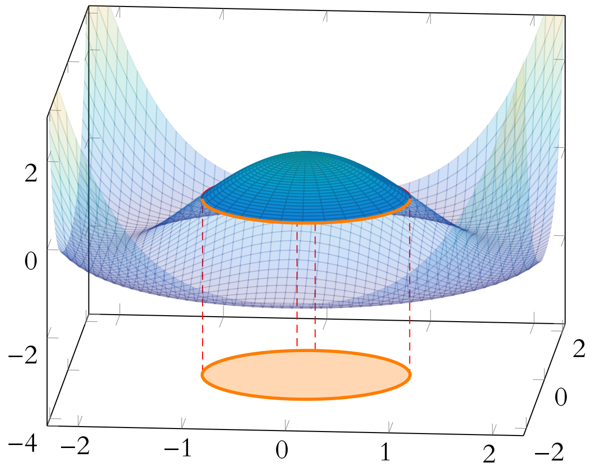


FIGURE 33.9 – Un maximum local en $(0,0)$: pour tout $x \in B_o((0,0),1)$, on a $f(x) \leqslant f(0,0)$. Notons qu'il ne s'agit pas d'un maximum global : en dehors de $B_o((0,0),1)$, f prend des valeurs supérieures à $f(0,0)$.

33.5.2 Fonctions de classe $\mathscr{C}^1$ sur un ouvert

Définition 33.44 – Soit $\mathscr{O}$ un ouvert de $\mathbf{R}^2$, et soit $f : \mathscr{O} \to \mathbf{R}$ une fonction de classe $\mathscr{C}^1$ sur $\mathscr{O}$.
Un point $a \in \mathscr{O}$ est appelé un **point critique de** f si $\nabla f(a) = (0,0)$, soit encore si $\dfrac{\partial f}{\partial x}(a) = \dfrac{\partial f}{\partial y}(a) = 0$.

Proposition 33.45 : *Soit $f : \mathscr{O} \to \mathbf{R}$ une fonction de classe $\mathscr{C}^1$ sur un ouvert $\mathscr{O}$ de $\mathbf{R}^2$. Si f possède un extremum local en $a \in \mathscr{O}$, alors a est un point critique de f.*

Démonstration. Supposons par exemple que f possède un maximum local en a.
Soit alors $r > 0$ tel que $\forall x \in B_o(x,r) \cap \mathscr{O}$, $f(x) \leqslant f(a)$.
Quitte à diminuer r, on peut supposer que $B_o(x,r) \subset \mathscr{O}$.
Soit alors $u \in \mathbf{R}^2$, et soit $f_u : \left| \begin{array}{ccc} \left]-\frac{\|u\|}{r}, \frac{\|u\|}{r}\right[& \longrightarrow & \mathbf{R} \\ t & \longmapsto & f(a+tu) \end{array} \right.$.
Alors nous avons déjà prouvé que f_u est dérivable, et que $f_u'(0) = \langle \nabla f(a), u \rangle$.
Par ailleurs, par définition d'un extremum local, f_u possède un maximum local en 0.
Mais alors cela signifie[22] que sa dérivée en 0 est nulle, c'est-à-dire que $\langle \nabla f(a), u \rangle = 0$.
Ceci étant valable pour tout $u \in \mathbf{R}^2$, c'est en particulier vrai pour $u = \nabla f(a)$, si bien que $\|\nabla f(a)\|^2 = 0 \Leftrightarrow \nabla f(a) = (0,0)$.
Et donc a est un point critique de f. □

[22] C'est un résultat bien connu pour les fonctions d'une variable sur un intervalle ouvert.

⚠ Comme pour les fonctions d'une variable, la réciproque est fausse, en un point critique, une fonction n'admet pas forcément d'extremum local.
L'exemple le plus classique est celui de la fonction $f : (x,y) \mapsto x^2 - y^2$.
Alors $(0,0)$ est un point critique de f.
Pourtant, pour tout $r > 0$, $\left(\frac{r}{2}, 0\right) \in B_o((0,0),r)$ et $f\left(\frac{r}{2},0\right) = \frac{r^2}{4} > 0 = f(0,0)$, si bien que f n'admet pas de maximum local en $(0,0)$.
Et de même, $\left(0,\frac{r}{2}\right) \in B_o((0,0,),r)$ et $f\left(0,\frac{r}{2}\right) = -\frac{r^2}{4} < 0$, si bien que f n'admet pas de minimum local en $(0,0)$.

⚠ Attention !
La négation de
$$\exists r > 0,\ \forall x \in B_o(a,r) \dots$$
est bien
$$\forall r > 0,\ \exists x \in B_o(a,r) \dots$$
Et donc pour prouver que a n'est pas un extremum local, il faut prouver que dans **toute** boule centrée en a, f prend des valeurs plus grandes et des valeurs plus petites que $f(a)$.

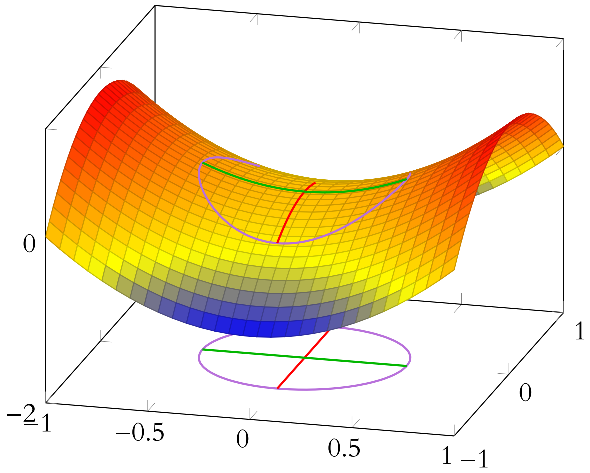


FIGURE 33.10 – $f(x,y) = x^2 - y^2$ n'a pas d'extremum local en son unique point critique.

Exemple 33.46

► Soit $f : (x,y) \mapsto x^3 + y^3 - 3xy + 1$.
Alors f est de classe $\mathscr{C}^1$ sur $\mathbf{R}^2$ car polynomiale.
On a alors, pour tout $(x,y) \in \mathbf{R}^2$, $\nabla f(x,y) = (3x^2 - 3y, 3y^2 - 3x)$.
Donc (x,y) est un point critique de f si et seulement si

$$\begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 = y \\ y^2 = x \end{cases} \Leftrightarrow \begin{cases} x^2 = y \\ x^4 = x \end{cases} \Leftrightarrow (x,y) = (0,0) \textbf{ ou } (x,y) = (1,1).$$

Aucun de ces deux points critiques n'est un extremum global puisque pour tout $x \in \mathbf{R}$, $f(x,0) = x^3 + 1 \underset{x \to +\infty}{\longrightarrow} +\infty$, si bien que f n'est pas majorée.
De même, $f(x,0) \underset{x \to -\infty}{\longrightarrow} -\infty$, donc f n'est pas minorée.

Pour $x \in \mathbf{R}$, $f(x,0) - f(0,0) = x^3$ est du signe de x.
Donc pour tout $r > 0$, $\left(\frac{r}{2}, 0\right) \in B_o((0,0),r)$, et $f\left(\frac{r}{2}, 0\right) > f(0,0)$.
Donc f ne possède pas de maximum local en $(0,0)$.
De même, $\left(\frac{-r}{2}, 0\right) \in B_o((0,0),r)$ et $f\left(-\frac{r}{2}, 0\right) < f(0,0)$, donc f ne possède pas de minimum local en $(0,0)$.
Par ailleurs, $f(1,1) = 0$, et pour $(u,v) \in \mathbf{R}^2$, on a

> **Astuce**
> Étudier $f(x,y)$ au voisinage de $(1,1)$, c'est étudier $f(1+u, 1+v)$ pour (u,v) au voisinage de $(0,0)$.

$$f(1+u, 1+v) = (1+u)^3 + (1+v)^3 - 3(1+u)(1+v) + 1 = u^2(3+u) + v^2(3+v) - 3uv.$$

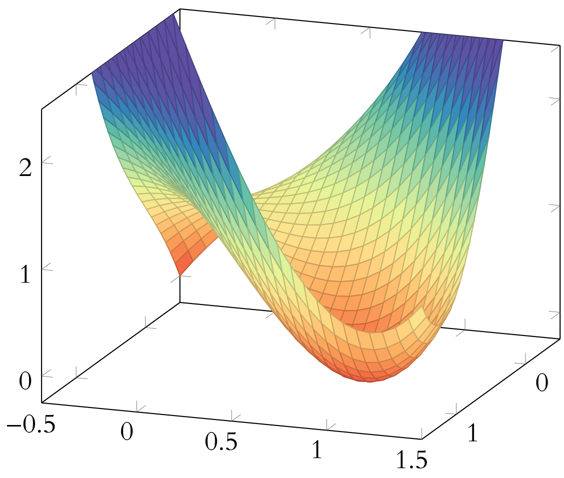
Mais alors si $(u,v) \in B_o((0,0),1)$, alors $|u| \leqslant \sqrt{u^2+v^2} < 1$, si bien que $3 + u \geqslant 2$.
Et de même $3 + v \geqslant 2$.
Et par ailleurs, $|uv| \leqslant \dfrac{u^2+v^2}{2}$, si bien que

$$f(1+u, 1+v) \geqslant 2u^2 + 2v^2 - \frac{3}{2}(u^2+v^2) \geqslant \frac{u^2+v^2}{2} \geqslant 0.$$

Et alors pour tout $(x,y) \in B_o((1,1),1)$, si on pose $(u,v) = (x-1, y-1)$, on a alors $(u,v) \in B_o((0,0),1)$, et donc

$$f(x,y) = f(1+u, 1+v) \geqslant 0 = f(1,1).$$

Donc f possède un minimum local en $(1,1)$.



33.5.3 Fonctions continues sur un fermé borné

Définition 33.47 – Soit $A \subset \mathbf{R}^2$. On dit que A est **bornée** s'il existe $M \geqslant 0$ tel que $\forall a \in A, \|a\| \leqslant M$.

Autrement dit, une partie est bornée si et seulement si elle est incluse dans une boule centrée en $0_{\mathbf{R}^2}$.

Proposition 33.48 : *Soit $A \subset \mathbf{R}^2$. Alors A est bornée si et seulement si il existe $M > 0$ tel que $\forall (x, y) \in A$, $|x| \leqslant M$* **et** *$|y| \leqslant M$.*

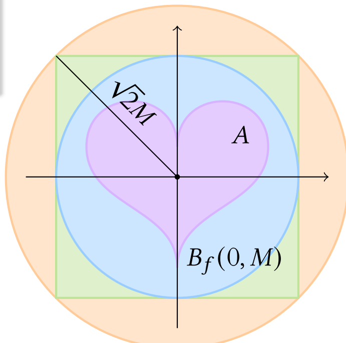


Figure 33.11– L'idée de la preuve est essentiellement de dire qu'une boule est incluse dans un cube, lui-même inclus dans une boule plus grande.

Démonstration. $\Rightarrow$ Si A est bornée, soit $M > 0$ tel que $A \subset B_f(0_{\mathbf{R}^2}, M)$.
Alors pour $(x, y) \in A$, on a $|x| \leqslant \|(x, y)\| \leqslant M$, et de même $|y| \leqslant M$.

$\Leftarrow$ Supposons qu'il existe $M > 0$ tel que pour tout $(x, y) \in A$, $|x| \leqslant M$ et $|y| \leqslant M$.
Alors pour $(x, y) \in A$, $\|(x, y)\| = \sqrt{x^2 + y^2} \leqslant \sqrt{M^2 + M^2} \leqslant \sqrt{2}M$.
Et donc $A \subset B_o(0_{\mathbf{R}^2}, \sqrt{2}M)$. □

Encore une fois, le théorème qui suit est hors programme en sup, mais il est intéressant de comprendre en quoi il généralise le théorème des bornes atteintes, et il sera (très largement) généralisé en seconde année.

Proposition 33.49 : *Soit $\mathscr{F}$ une partie fermée et bornée de $\mathbf{R}^2$, et soit $f : \mathscr{F} \to \mathbf{R}$ une fonction continue.*
Alors f possède un maximum et un minimum.

Démonstration. Notons $M = \sup\{f(u),\ u \in \mathscr{F}\}$, éventuellement égal à $+\infty$.
Par caractérisation séquentielle des bornes supérieures, il existe une suite $(u_n) \in \mathscr{F}^{\mathbf{N}}$ telle que $f(u_n) \underset{n\to+\infty}{\longrightarrow} M$.
Notons A un réel positif tel que pour tout $x \in \mathscr{F}$, $\|x\| \leqslant A$.
Alors, pour tout $u = (x_u, y_u) \in \mathscr{F}$, on a $|x_u|^2 \leqslant x_u^2 + y_u^2 \leqslant A^2$, si bien que $|x_u| \leqslant A$.
Et de même, $|y_u| \leqslant A$.

Pour tout $n \in \mathbf{N}$, notons $u_n = (x_n, y_n)$. Alors la suite (x_n) est une suite réelle bornée, si bien qu'on peut en extraire une suite convergente : il existe une extractrice φ telle que $(x_{\varphi(n)})_n$ converge vers un réel x.
Puis alors la suite $(y_{\varphi(n)})$ est encore bornée, et on peut de nouveau en extraire une suite convergente : il existe une extractrice ψ telle que $(y_{\varphi(\psi(n))})$ converge vers un réel y.

Prouvons alors que $(x,y) \in \mathcal{F}$. Supposons par l'absurde que $(x,y) \in \overline{\mathcal{F}}$. Puisque $\mathcal{F}$ est fermé, il existe $r > 0$ tel que $B_o((x,y),r) \subset \overline{\mathcal{F}}$.
Mais pour n suffisamment grand, $|x_{\varphi(\psi(n))} - x| < \frac{r}{\sqrt{2}}$ et de même, $|y_{\varphi(\psi(n))} - y| < \frac{r}{\sqrt{2}}$.
Donc $\|u_{\varphi(\psi(n))} - (x,y)\| < \sqrt{\left(\frac{r}{\sqrt{2}}\right)^2 + \left(\frac{r}{\sqrt{2}}\right)^2} = r$, si bien que $u_{\varphi(\psi(n))} \in B_o((x,y),r)$ et donc $u_{\varphi(\psi(n))} \notin \mathcal{F}$, ce qui contredit la définition de u_n.
Donc $(x,y) \in \mathcal{F}$.

Considérons à présent $\varepsilon > 0$. Par définition de la continuité de f en (x,y), il existe donc $\eta > 0$ tel que pour tout $u \in \mathcal{F}$, $\|u - (x,y)\| < \eta \Rightarrow |f(u) - f(x,y)| < \varepsilon$.
Or, comme précédemment, si n est suffisamment grand, $\|u_{\varphi(\psi(n))} - (x,y)\| < \eta$.
Et donc $|f(u_n) - f(x,y)| < \varepsilon$.
Mais lorsque $n \to +\infty$, $f(u_n) \xrightarrow[n\to+\infty]{} M$.
Ce qui prouve déjà que $M \in \mathbf{R}$, et donc que f est majorée, mais en plus que $|M - f(x,y)| \leqslant \varepsilon$, et ce pour tout $\varepsilon > 0$.
On en déduit donc que $M = f(x,y)$, et donc que f possède un maximum atteint en (x,y).

Le raisonnement est le même pour l'existence d'un minimum[23]. □

[23] Appliquer ce qui précède à $-f$.

Exemple 33.50

Soit $f : (x,y) \mapsto \sqrt{1+x^2+y^2} + (x+y)^2 - 1$, et soit $\mathcal{D} = \{(x,y) \in \mathbf{R}^2 \mid x^2+y^2 \leqslant 3\}$.
Alors $\mathcal{D}$ est un fermé, puisque c'est l'image réciproque par la fonction continue $(x,y) \mapsto x^2+y^2$ de $]-\infty, 3]$.
Il est borné, puisque c'est une boule fermée centrée en $(0,0)$: $\mathcal{D} = B_0((0,0)), \sqrt{3})$.
Puisque f est continue sur $\mathcal{D}$, elle y possède un maximum et un minimum.
Mais comment les déterminer ?

Commençons par noter que pour tout $(x,y) \in \mathcal{D}$, $f(x,y) \geqslant 0 = f(0,0)$, et donc f atteint son minimum en $(0,0)$.

Notons $\mathcal{D}_0 = B_o\left((0,0), \sqrt{3}\right) = \left\{(x,y) \in \mathbf{R}^2 \mid x^2+y^2 < 3\right\}$, qui cette fois est un ouvert puisque c'est une boule ouverte.
Alors sur l'ouvert $\mathcal{D}_0$, f est $\mathcal{C}^1$.
Donc si elle atteint son maximum ou son minimum sur $\mathcal{D}$ en un point de $\mathcal{D}_0$, il s'agira d'un extremum de f sur $\mathcal{D}_0$, nécessairement atteint en un point critique.
On a

$$\frac{\partial f}{\partial x}(x,y) = \frac{x}{\sqrt{1+x^2+y^2}} + 2(x+y) \text{ et } \frac{\partial f}{\partial y}(x,y) = \frac{y}{\sqrt{1+x^2+y^2}} + 2(x+y).$$

Et donc $(x,y) \in \mathcal{D}_0$ est un point critique de f si et seulement si

$$\begin{cases} \dfrac{x}{\sqrt{1+x^2+y^2}} + 2(x+y) = 0 \\ \dfrac{y}{\sqrt{1+x^2+y^2}} + 2(x+y) = 0 \end{cases} \Leftrightarrow \begin{cases} x = y \\ x + 4x\sqrt{1+2x^2} = 0 \end{cases} \Leftrightarrow \begin{cases} x = y \\ x\left(1+\sqrt{1+2x^2}\right) = 0 \end{cases} \Leftrightarrow x = y = 0.$$

Donc $(0,0)$ est le seul point critique de f, et nous savons déjà qu'il s'agit du point où f atteint son minimum.
Et donc nécessairement, le maximum de f sur $\mathcal{D}$ n'est pas atteint sur $\mathcal{D}_0$, mais donc sur un point de «bord[24]» de $\mathcal{D}$, c'est-à-dire de $\mathcal{D} \setminus \mathcal{D}_0$.
Soit donc $(x,y) \in \mathcal{D} \setminus \mathcal{D}_0$ de sorte que $x^2+y^2 = 3$. Alors

$$f(x,y) = \sqrt{1+3} + (x+y)^2 - 1 = x^2 + 2xy + y^2 + 1 = 4 + 2xy \leqslant 4 + (x^2+y^2) = 7.$$

Et de plus, il y a égalité si et seulement si $x = y$.

Et puisque $x^2+y^2 = 3$, on a donc $f(x,y) = 7$ si et seulement si $(x,y) = \left(\sqrt{\frac{3}{2}}, \sqrt{\frac{3}{2}}\right)$

[24] Qui est ici un cercle.

Égalité
Pour déterminer le cas d'égalité dans l'inégalité

$$ab \leqslant \frac{a^2+b^2}{2}$$

il faut se souvenir de sa preuve : on développe $(a-b)^2 \geqslant 0$.

ou $(x, y) = \left(-\sqrt{\frac{3}{2}}, -\sqrt{\frac{3}{2}}\right)$.

Et donc le minimum de f sur $\mathscr{D}$ est 0, atteint uniquement en $(0, 0)$, et le maximum est 7, atteint uniquement en les deux points précédemment cités.