Chapitre 18

Fonctions de plusieurs variables
Calcul et géométrie différentiels

Dans ce chapitre, p désigne un entier naturel non nul et U désigne un ouvert de $\mathbb{R}^p$. On notera $\|\cdot\|$ une norme quelconque sur $\mathbb{R}^p$ et $\mathcal{B} = (e_1, \ldots, e_p)$ la base canonique de $\mathbb{R}^p$.

Dans le chapitre **Espaces vectoriels normés**, nous nous sommes intéressés notamment à la continuité des fonctions de $\mathbb{R}^p$ dans $\mathbb{R}$. Dans ce chapitre, nous nous intéressons à l'aspect différentiel. Bien sûr, on ne peut pas procéder comme pour les fonctions de la variable réelle, car la notion de taux d'accroissement n'a pas de sens si $p \geqslant 2$. On souhaite malgré tout généraliser la notion de dérivée, qui permet notamment, si f est une fonction d'une variable, l'approximation

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

L'une des principales difficultés est la généralisation du terme $f'(a)h$ lorsque $p \geqslant 2$.

I. Introduction

Soit $f : U \to \mathbb{R}$ une fonction. Pour étudier f, une première idée peut être de se ramener à des fonctions d'une variable en considérant les fonctions obtenues à partir de f en « fixant » toutes les variables sauf une.

Plus précisément, pour étudier f au voisinage de $a = (a_1, \ldots, a_p) \in \mathbb{R}^p$, on s'intéresse aux p fonctions

$$t \mapsto f\left(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p\right)$$

pour $i \in [\![1, p]\!]$, respectivement définies sur les ensembles

$$U_{a,i} = \{t \in \mathbb{R};\ (a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p) \in U\}.$$

Elles sont parfois appelées « applications partielles » de f en a.

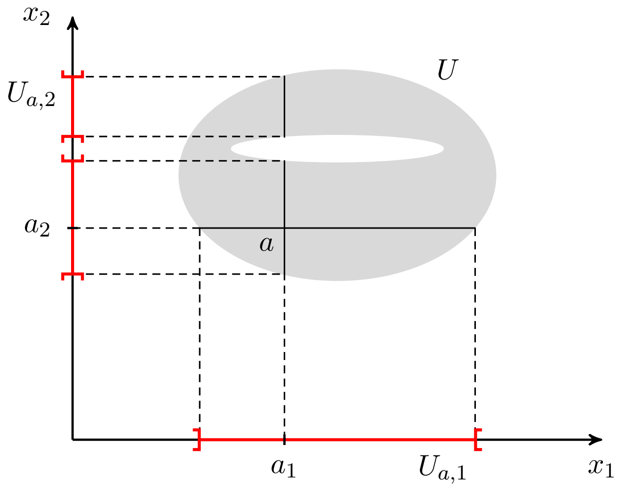
Exemple – Soient

$$f : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R}_+^* & \to & \mathbb{R} \\ (x,y) & \mapsto & e^{xy} \ln(y) \end{array} \right.$$

et $a = (1,2) \in \mathbb{R} \times \mathbb{R}_+^*$. Les deux fonctions décrites ci-dessus sont

$$t \mapsto \ln(2)\, e^{2x} \quad \text{définie sur } \mathbb{R}, \quad \text{et} \quad t \mapsto e^y \ln(y) \quad \text{définie sur } \mathbb{R}_+^*. \qquad \square$$

La figure ci-dessous montre les ensembles $U_{a,i}$ pour un choix particulier de U et a, en dimension 2.



Dans le cas général, U étant ouvert, on montre facilement que $U_{a,i}$ est toujours un ouvert de $\mathbb{R}$: soit $t_0 \in U_{a,i}$; alors

$$(a_1, \ldots, a_{i-1}, t_0, a_{i+1}, \ldots, a_p) \in U.$$

Comme U est ouvert, il existe $r > 0$ tel qu'on ait l'implication

$$[\,|t - t_0| < r \quad \text{et} \quad \forall\, k \neq i,\ |x_k - a_k| < r\,] \quad \Rightarrow \quad (x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_p) \in U.$$

En particulier, en choisissant $x_k = a_k$ pour tout $k \neq i$, on a montré que $]t_0 - r, t_0 + r[\subset U_{a,i}$, d'où le résultat.

Par un raisonnement analogue, on montre aussi le résultat suivant :

Propriété

Si f est continue en $a \in U$, alors pour tout $a \in U$ et $i \in [\![1, p]\!]$,

$$t \mapsto f\,(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p)$$

est continue en a_i : la continuité de f entraîne sa continuité « par rapport à chacune de ses variables ».

Un point très important est le fait que la réciproque est **fausse** : toutes les fonctions

$$t \mapsto f\,(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p)$$

peuvent être continues sans que f le soit. Ceci tient au fait que la continuité de f signifie que pour tout $a \in U$, $f(x) \to f(a)$ lorsque x tend vers a « de façon arbitraire ». La continuité par rapport à la variable x_i signifie que $f(x) \to f(a)$ lorsque x tend vers a en ayant la forme particulière

$$(a_1, \ldots, a_{i-1}, t, a_{i+1}, \ldots, a_p),$$

c'est-à-dire, le long de la droite $a + \mathbb{R}\, e_i$ passant par a et dirigée par e_i. Ceci est plus restrictif, même lorsque cela a lieu pour tout i.

Exemple – Par exemple, soit $f : \mathbb{R}^2 \to \mathbb{R}$ définie par

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

Les deux fonctions

$$t \mapsto f(t, 0) \quad \text{et} \quad t \mapsto f(0, t)$$

sont continues sur $\mathbb{R}$: elles sont nulles. Pourtant, f n'est pas continue en 0 car pour $t \neq 0$,

$$f(t,t) = \frac{t^2}{2t^2} = \frac{1}{2},$$

qui ne tend pas vers $f(0,0)$ lorsque $t \to 0$. En quelque sorte, il y a continuité à l'origine le long des deux axes de coordonnées, mais pas le long de la première bissectrice. $\square$

L'étude de la dépendance de f par rapport à chacune de ses variables **ne suffit donc pas** à faire l'étude de f. Sans renoncer entièrement à cette approche dans la suite, il faudra se souvenir de ce phénomène.

II. Fonctions de classe $\mathcal{C}^1$

1. Dérivées partielles

Définition – Dérivées partielles

Soient $a = (a_1, \dots, a_p) \in U$ et $i \in [\![1,p]\!]$. On dit que f admet une **dérivée partielle** en a par rapport à la i-ième variable si

$$t \mapsto f\left(a_1, \dots, a_{i-1}, t, a_{i+1}, \dots, a_p\right)$$

est dérivable en a_i, c'est-à-dire, si

$$h \mapsto \frac{1}{h}\left(f\left(a_1, \dots, a_{i-1}, a_i + h, a_{i+1}, \dots, a_p\right) - f\left(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_p\right)\right)$$

a une limite finie lorsque $h \to 0$ avec $h \neq 0$.
Dans ce cas, cette limite est notée

$$\frac{\partial f}{\partial x_i}(a) \quad \text{ou} \quad \partial_i f(a).$$

Elle est appelée **dérivée partielle** de f en a par rapport à la i-ième variable.

Remarque – On considère des fonctions définies sur un **ouvert** car cela garantit que le taux d'accroissement apparaissant dans la définition précédente est bien défini pour $h \neq 0$ assez proche de 0 (sans avoir à imposer, notamment, $h > 0$ ou $h < 0$).

Cas particulier – Pour une fonction f de deux variables, étudier l'existence de dérivées partielles en $(a,b) \in U$ revient à étudier l'existence éventuelle de limites lorsque $h \to 0$ des « taux d'accroissements partiels »

$$\frac{f(a+h,b)}{h} \quad \text{et} \quad \frac{f(a,b+h)}{h}.$$

Exemple – Soit

$$f : \left\{ \begin{array}{ccc} \mathbb{R} \times \mathbb{R}_+ & \to & \mathbb{R} \\ (x_1, x_2) & \mapsto & x_1^2 \sqrt{x_2} \end{array} \right.$$

Quel que soit $a_2 \in \mathbb{R}_+$ fixé, la fonction $x_1 \mapsto f(x_1, a_2) = \sqrt{a_2}\, x_1^2$ est dérivable en tout point a_1 de $\mathbb{R}$, et

$$\frac{\partial f}{\partial x_1}(a_1, a_2) = 2\sqrt{a_2}\, a_1.$$

En revanche, si l'on fixe $a_1 \in \mathbb{R}$, la fonction $x_2 \mapsto f(a_1, x_2) = a_1^2 \sqrt{x_2}$ est dérivable en tout point a_2 de $\mathbb{R}_+$ si $a_1 = 0$ (car c'est alors la fonction nulle), et dérivable en tout point a_2 de $\mathbb{R}_+^*$ mais pas en 0 si $a_1 \neq 0$, et dans ce cas, pour tout $a_2 \in \mathbb{R}^*$,

$$\frac{\partial f}{\partial x_2}(a_1, a_2) = \frac{a_1^2}{2\sqrt{a_2}}.$$

Définition – Fonctions dérivées partielles

Si f admet une dérivée partielle sur U (*i.e.*, en tout point de U) par rapport à la i-ième variable, alors la fonction

$$a \mapsto \frac{\partial f}{\partial x_i}(a)$$

(qui est définie sur U, à valeurs dans $\mathbb{R}$) est appelée **dérivée partielle** de f par rapport à la i-ième variable.

2. Classe $\mathcal{C}^1$

Définition

Soit $f : U \to \mathbb{R}$ une fonction.

On dit que f est de classe $\mathcal{C}^1$ sur U si f admet des dérivées partielles sur U par rapport à toutes ses variables, et si toutes ces dérivées partielles sont continues sur U.

Attention ! Si f est de classe $\mathcal{C}^1$ sur U, alors pour tout $a \in U$, pour tout $i \in [\![1,p]\!]$,

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

est de classe $\mathcal{C}^1$ sur $U_{a,i}$. La réciproque est fausse, le même contre-exemple que dans le cas de la continuité le prouve.

Théorème (admis : démonstration non exigible)

Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Alors f admet en tout point $a \in U$ le développement limité à l'ordre 1

$$f(a+h) = f(a) + \frac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \frac{\partial f}{\partial x_p}(a)\, h_p + o(\|h\|),$$

lorsque $h = (h_1, \dots, h_p) \to (0, \dots, 0)$.

Remarque – La notation précédente signifie que l'on peut écrire, pour h tel que $a + h \in U$,

$$f(a+h) = f(a) + \frac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \frac{\partial f}{\partial x_p}(a)\, h_p + \|h\|\, \varepsilon(h),$$

où ε a pour limite 0 en $(0, \dots, 0)$.

Définition – Différentielle

Soient $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $a \in U$. On appelle **différentielle** de f en a la forme linéaire

$$\mathrm{d}f(a) : \left\{ \begin{array}{rcl} \mathbb{R}^p & \to & \mathbb{R} \\ (h_1, \dots, h_p) & \mapsto & \dfrac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \dfrac{\partial f}{\partial x_p}(a)\, h_p \end{array} \right.$$

L'image de $h \in \mathbb{R}^p$ par l'application $\mathrm{d}f(a)$ sera notée $\mathrm{d}f(a) \cdot h$.

Attention ! Dans le cadre précédent, pour tout $a \in U$, $\mathrm{d}f(a)$ est elle-même une **application** (en l'occurrence une forme linéaire) définie sur $\mathbb{R}^p$ tout entier.

Remarque – Le théorème précédent se réinterprète ainsi : si $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^1$, alors pour tout $a \in U$,

$$f(a+h) \underset{\|h\| \to 0}{=} f(a) + \mathrm{d}f(a) \cdot h + o(\|h\|).$$

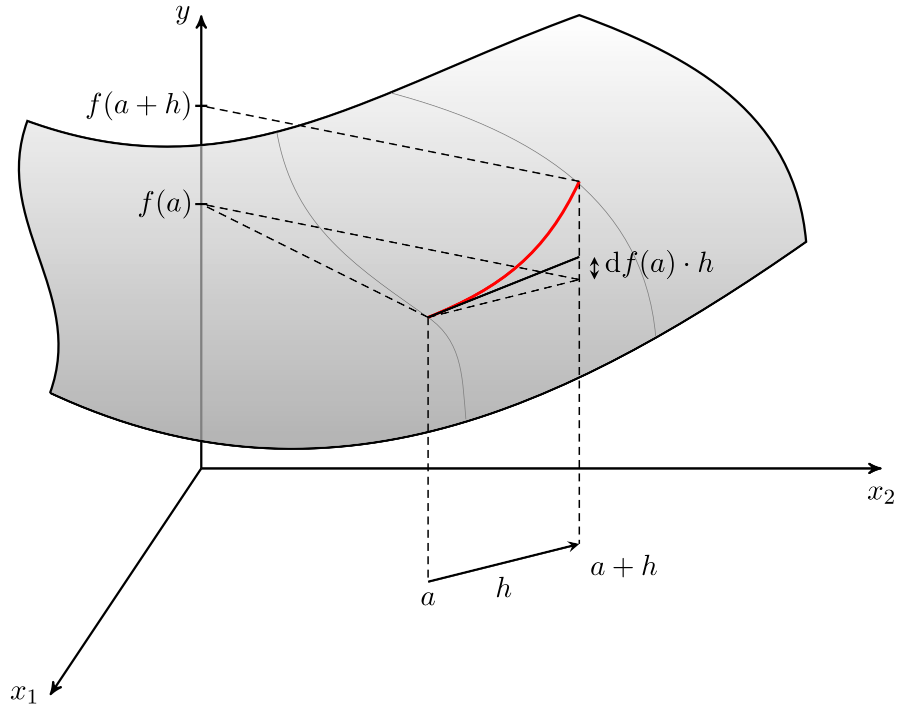
Ceci est bien sûr à mettre en relation avec le développement limité

$$g(a+h) \underset{h\to 0}{=} g(a) + g'(a)h + o(h)$$

pour une fonction $g : I \subset \mathbb{R} \to \mathbb{R}$ de classe $\mathcal{C}^1$. Ici, le terme

$$\mathrm{d}f(a)\cdot h = \frac{\partial f}{\partial x_1}(a)\, h_1 + \cdots + \frac{\partial f}{\partial x_p}(a)\, h_p$$

généralise terme $g'(a)h$, mais il prend en compte, du fait de la présence de plusieurs variables, les accroissements de f dans toutes les directions.
La forme linéaire $\mathrm{d}f(a)$ est l'application qui, à un « déplacement » h par rapport au point a, fait correspondre la modification (au premier ordre) sur la valeur de f qui en résulte.



Pour une fonction g d'une seule variable, de classe $\mathcal{C}^1$ sur I, pour tout $a \in I$, $\mathrm{d}g(a)$ est l'application

$$h \in \mathbb{R} \mapsto g'(a)h$$

de multiplication par $g'(a)$. En revanche, si $p \geqslant 2$, la notation $\cdot$ dans $\mathrm{d}f(a)\cdot h$ **ne désigne pas un produit**.

Propriété

Si f est de classe $\mathcal{C}^1$ sur U, elle est continue sur U.

Démonstration – La fonction f est de classe $\mathcal{C}^1$ sur U, donc pour tout point $a \in U$,

$$f(a+h) \underset{\|h\|\to 0}{=} f(a) + \mathrm{d}f(a)\cdot h + o(\|h\|) \underset{h\to(0,\ldots,0)}{\longrightarrow} f(a)$$

car $\mathrm{d}f(a)$ est continue. D'où le résultat. □

Propriété

- Toute fonction polynomiale définie sur un ouvert est de classe $\mathcal{C}^1$.
- En particulier, toute application linéaire de $\mathbb{R}^p$ dans $\mathbb{R}$ est de classe $\mathcal{C}^1$.
- Toute fraction rationnelle dont le dénominateur ne s'annule pas est de classe $\mathcal{C}^1$.

Démonstration – On considère les applications partielles et on applique les résultats analogues pour les fonctions d'une variable, d'où l'existence des dérivées partielles ; elles sont elles-mêmes soit polynomiales soit des fractions rationnelles dont le dénominateur ne s'annule pas, donc continues. $\square$

3. Opérations sur les fonctions de classe $\mathcal{C}^1$

Propriété – Combinaison linéaire

Soient $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ deux fonctions de classe $\mathcal{C}^1$, et $\lambda \in \mathbb{R}$.
Alors $\lambda f + g$ est de classe $\mathcal{C}^1$ sur U et pour tout $a \in U$,

$$\mathrm{d}(\lambda f + g)(a) = \lambda\,\mathrm{d}f(a) + \mathrm{d}g(a),$$
$$\text{et :}\quad \forall\, i \in [\![1, p]\!],\quad \frac{\partial(\lambda f + g)}{\partial x_i}(a) = \lambda\frac{\partial f}{\partial x_i}(a) + \frac{\partial g}{\partial x_i}(a).$$

Corollaire

L'ensemble $\mathcal{C}^1(U,\mathbb{R})$ des fonctions de classe $\mathcal{C}^1$ sur U à valeurs dans $\mathbb{R}$ est un $\mathbb{R}$-espace vectoriel.

Propriété – Produit

Soient $f : U \to \mathbb{R}$ et $g : U \to \mathbb{R}$ deux fonctions de classe $\mathcal{C}^1$.
Alors fg est de classe $\mathcal{C}^1$ sur U et pour tout $a \in U$,

$$\mathrm{d}(fg)(a) = (\mathrm{d}f(a))g(a) + f(a)(\mathrm{d}g(a)),$$
$$\text{et :}\quad \forall\, i \in [\![1, p]\!],\quad \frac{\partial(fg)}{\partial x_i}(a) = \frac{\partial f}{\partial x_i}(a)\,g(a) + f(a)\frac{\partial g}{\partial x_i}(a).$$

Propriété – Inverse

Soit $f : U \to \mathbb{R}^*$ une fonction de classe $\mathcal{C}^1$.
Alors $1/f$ est de classe $\mathcal{C}^1$ sur U et pour tout $a \in U$,

$$\mathrm{d}(1/f)(a) = -\frac{1}{f^2(a)}\,\mathrm{d}f(a),$$
$$\text{et :}\quad \forall\, i \in [\![1, p]\!],\quad \frac{\partial(1/f)}{\partial x_i}(a) = -\frac{1}{f^2(a)}\,\frac{\partial f}{\partial x_i}(a).$$

Démonstration des trois propriétés – C'est immédiat en considérant les applications partielles : les résultats sur les fonctions de la variable réelle prouvent l'existence des dérivées partielles ; elles sont continues par opérations sur les fonctions continues. $\square$

4. Composition : règle de la chaîne

Théorème – Règle de la chaîne

Soient I un intervalle de $\mathbb{R}$ et $x_1,\dots,x_p$ des fonctions de classe $\mathcal{C}^1$ sur I, à valeurs dans $\mathbb{R}$. Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$. On suppose que pour tout $t \in I$,

$$(x_1(t),\dots,x_p(t)) \in U.$$

Alors la fonction

$$g : t \mapsto f(x_1(t),\dots,x_p(t))$$

est définie et de classe $\mathcal{C}^1$ sur I, avec, pour tout $t \in I$,

$$g'(t) = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(x_1(t),\dots,x_p(t))\, x_i'(t).$$

Démonstration – La fonction g est bien définie car $(x_1,\dots,x_p)$ est à valeurs dans U. Soit $t \in I$; pour tout $i \in [\![1,p]\!]$, $x_i : I \to \mathbb{R}$ est de classe $\mathcal{C}^1$, donc il existe une fonction η_i qui a pour limite 0 en 0, telle que

$$x_i(t+h) = x_i(t) + x_i'(t)h + h\,\eta_i(h)$$

lorsque $t+h \in I$. De plus, f est de classe $\mathcal{C}^1$ sur U, donc en notant $a = (x_1(t),\dots,x_p(t))$, il existe une fonction ε qui a pour limite 0 en $(0,\dots,0)$, telle que

$$f(a+k) = f(a) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, k_i + \|k\|\,\varepsilon(k)$$

pour $k = (k_1,\dots,k_p)$ tel que $a+k \in U$. On écrit cette égalité avec

$$k = \big(x_1'(t)h + h\,\eta_1(h),\dots,x_p'(t)h + h\,\eta_p(h)\big)$$

lorsque $h \to 0$ avec $t+h \in I$; on a alors $k \to (0,\dots,0)$, donc $a+k \in U$ pour h assez proche de 0, d'où :

$$\begin{aligned}
g(t+h) &= f(x_1(t+h),\dots,x_p(t+h)) \\
&= f(x_1(t) + x_1'(t)h + h\,\eta_1(h),\dots,x_p(t) + x_p'(t)h + h\,\eta_p(h)) \\
&= f(x_1(t),\dots,x_p(t)) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\,\big(x_i'(t)h + h\eta_i(h)\big) + \|k\|\,\varepsilon(k) \\
&= g(t) + \left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, x_i'(t)\right) h + \left(\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, h\,\eta_i(h) + \|k\|\,\varepsilon(k)\right)
\end{aligned}$$

Pour conclure, il suffit donc de prouver que le terme dans la dernière parenthèse est un $o(h)$ lorsque $h \to 0$. Or, en choisissant la norme 1 ($\|\cdot\| = \|\cdot\|_1$), on a, pour $h \neq 0$,

$$\begin{aligned}
&\frac{1}{|h|}\left|\sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a)\, h\,\eta_i(h) + \|k\|\,\varepsilon(k)\right| \\
&\leqslant \sum_{i=1}^{p} \left(\left|\frac{\partial f}{\partial x_i}(a)\,\eta_i(h)\right| + |x_i'(t) + \eta_i(h)|\,|\varepsilon(k)|\right) \xrightarrow[h\to 0]{} 0.
\end{aligned}$$

On en déduit que g est dérivable sur I avec la formule annoncée pour g' ; de plus, cette formule montre que g' est continue sur I, car f et tous les x_i sont de classe $\mathcal{C}^1$. D'où le résultat. □

Remarques

• Si I est semi-ouvert ou fermé, la formule précédente doit être interprétée, aux extrémités de I, en termes de dérivées de g à gauche ou à droite.

• Dans la formule donnant $g'(t)$, x_i apparaît avec deux sens différents qu'il ne faut pas confondre : $\dfrac{\partial f}{\partial x_i}$ est une notation qui désigne la dérivée partielle de f par rapport à sa i-ième variable ; x_i apparaissant dans $x_i(t)$ ou $x_i'(t)$ désigne la fonction x_i. Il n'y a pas de confusion possible si l'on écrit, de façon équivalente,

$$g'(t) = \sum_{i=1}^{p} \partial_i f(x_1(t), \ldots, x_p(t))\, x_i'(t).$$

• La formule précédente s'écrit aussi, par définition de la différentielle,

$$\forall\, t \in I, \quad g'(t) = \mathrm{d}f(\gamma(t)) \cdot \gamma'(t),$$

où $\gamma = (x_1, \ldots, x_p)$.

• Avec les notations précédentes, (I,γ) est un arc paramétré de classe $\mathcal{C}^1$, et g' représente la dérivée de f le long de cet arc. $\square$

Dans la propriété suivante, on s'intéresse au cas de la composition

$$g : \left\{ \begin{array}{ccccc} V \subset \mathbb{R}^2 & \overset{(x,y)}{\rightarrow} & U \subset \mathbb{R}^2 & \overset{f}{\rightarrow} & \mathbb{R} \\ (u,v) & \mapsto & (x(u,v),y(u,v)) & \mapsto & f(x(u,v),y(u,v)) \end{array} \right.$$

Propriété – Application aux fonctions de deux variables

Soient V un ouvert de $\mathbb{R}^2$, x et y deux fonctions définies sur V, à valeurs dans $\mathbb{R}$, de classe $\mathcal{C}^1$ sur V. Soient U un ouvert de $\mathbb{R}^2$ et $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$. On suppose que pour tout $(u,v) \in V$,

$$(x(u,v), y(u,v)) \in U.$$

Alors la fonction

$$g : (u,v) \mapsto f(x(u,v), y(u,v))$$

est définie et de classe $\mathcal{C}^1$ sur V, avec, pour tout $(u,v) \in V$,

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

Démonstration – Il suffit d'appliquer le théorème précédent en faisant jouer à t le rôle de u à v fixé, puis celui de v à u fixé. La variable t décrit alors un ouvert de $\mathbb{R}$ (pas nécessairement un intervalle) comme on l'a montré au début de ce chapitre. On peut appliquer le théorème au voisinage de chaque point de cet ouvert. $\square$

Exemple – Passage en coordonnées polaires

Pour tout $(x,y) \in \mathbb{R}^2$, il existe $r \in \mathbb{R}_+$ et $\theta \in \mathbb{R}$ tels que $(x,y) = (r\cos(\theta), r\sin(\theta))$. On dit que r, θ sont des **coordonnées polaires** de (x,y). Il n'y a pas unicité de telles coordonnées : par exemple si $(x,y) = (0,0)$, $r = 0$ et tout $\theta \in \mathbb{R}$ conviennent. De même, si r, θ sont des coordonnées polaires de (x,y), alors pour tout $k \in \mathbb{Z}$, r et $\theta + 2k\pi$ en sont également.

On définit, pour $(r,\theta) \in \mathbb{R}^2$,

$$x(r,\theta) = r\cos(\theta) \quad \text{et} \quad y(r,\theta) = r\sin(\theta).$$

Si $f : \mathbb{R}^2 \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^1$, on pose

$$g(r,\theta) = f(x(r,\theta), y(r,\theta)) = f(r\cos(\theta), r\sin(\theta)).$$

Par exemple, $g(\sqrt{2}, -\pi/4) = f(1, -1)$. D'après la propriété précédente, g est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et pour tout $(r,\theta) \in \mathbb{R}^2$,

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

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

Propriété – Caractérisation des fonctions constantes

Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur un ouvert U **convexe**.

Pour que f soit constante, il faut et il suffit que pour tout $i \in [\![1,p]\!]$, $\dfrac{\partial f}{\partial x_i} = 0$.

Démonstration

$\boxed{\Rightarrow}$ Évident, car toutes les applications partielles de f, qui sont des fonctions d'une variable, sont constantes et de classe $\mathcal{C}^1$.

$\boxed{\Leftarrow}$ Soient $a = (a_1, \dots, a_p) \in U$, $b = (b_1, \dots, b_p) \in U$ et

$$g : \left\{ \begin{array}{rcl} [0,1] & \to & \mathbb{R} \\ t & \mapsto & f(a + t\,(b-a)) = f(a_1 + t\,(b_1 - a_1), \dots, a_p + t\,(b_p - a_p)) \end{array} \right.$$

Cette fonction est bien définie, car lorsque t parcourt $[0,1]$, $a + t(b-a)$ parcourt le segment $[a,b]$, qui est contenu dans U car U est convexe. D'après la règle de la chaîne, g est de classe $\mathcal{C}^1$ sur l'intervalle $[0,1]$ et pour tout $t \in [0,1]$,

$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(a + t(b-a))\,(b_i - a_i) = 0.$$

La fonction d'une variable g est donc constante, et en particulier $g(0) = g(1)$, *i.e.* $f(a) = f(b)$. Ceci étant vrai pour tout $(a,b) \in U^2$, f est constante. □

Plus généralement, on a le résultat suivant :

Propriété

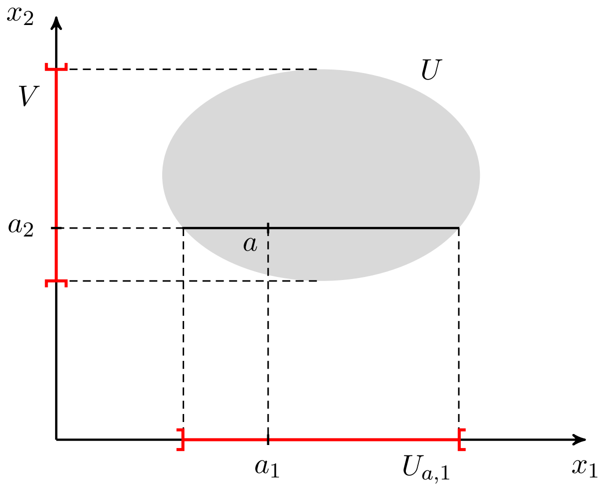
Soit $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur un ouvert U **convexe**, avec $p \geqslant 2$.
On suppose que

$$\frac{\partial f}{\partial x_1} = 0.$$

Alors f ne dépend pas de sa première variable : il existe un ouvert convexe V de $\mathbb{R}^{p-1}$ et une fonction $g : V \to \mathbb{R}$ de classe $\mathcal{C}^1$ telle que pour tout $(x_1, \dots, x_p) \in U$,

$$(x_2, \dots, x_p) \in V \quad \text{et} \quad f(x_1, \dots, x_p) = g(x_2, \dots, x_p).$$

Bien sûr, on peut généraliser ce résultat aux autres variables.



Démonstration

• Fixons $a = (a_1, \dots, a_p) \in U$ et considérons la première application partielle $f_{a,1}$ de f en a, définie sur l'ouvert non vide

$$U_{a,1} = \{t \in \mathbb{R};\ (t, a_2, \dots, a_p) \in U\}.$$

On sait déjà que $U_{a,1}$ est ouvert, en fait c'est un intervalle : si t_0 et t_1 sont deux éléments de $U_{a,1}$ avec $t_0 \leqslant t_1$ et si $t \in [t_0, t_1]$, alors le point $(t, a_2, \dots, a_p)$ appartient au segment joignant $(t_0, a_2, \dots, a_p)$ et $(t_1, a_2, \dots, a_p)$. Comme ces deux points appartiennent à U qui est convexe, on a $(t, a_2, \dots, a_p) \in U$, d'où : $t \in U_{a,1}$. Ceci prouve que $U_{a,1}$ est un intervalle ouvert. De plus, f étant de classe $\mathcal{C}^1$ sur U, $f_{a,1}$ est de classe $\mathcal{C}^1$ sur $U_{a,1}$, avec, pour tout $t \in U_{a,1}$,

$$f'_{a,1}(t) = \frac{\partial f}{\partial x_1}(t, a_2, \dots, a_p) = 0.$$

On en déduit que $f_{a,1}$ est constante sur $U_{a,1}$. Notons $g(a_2, \dots, a_p)$ l'unique valeur prise par $f_{a,1}$ sur $U_{a,1}$. On a donc

$$f(x_1, x_2, \dots, x_p) = g(x_2, \dots, x_p),$$

et ce, pour tout $(x_2, \dots, x_p)$ tel qu'il existe au moins une valeur x_1 telle que $(x_1, x_2, \dots, x_p) \in U$. Notons V l'ensemble de ces $(p-1)$-uplets $(x_2, \dots, x_p)$.

• V est un ouvert de $\mathbb{R}^{p-1}$: soient $(x_2, \dots, x_p) \in V$ et $x_1 \in \mathbb{R}$ tel que $(x_1, x_2, \dots, x_p) \in U$. Comme U est ouvert, il existe $r > 0$ tel que pour tout $y = (y_1, \dots, y_p) \in \mathbb{R}^p$ vérifiant $|y_i - x_i| < r$ pour tout i, on ait $y \in U$. Alors, pour tout $(y_2, \dots, y_p) \in \mathbb{R}^{p-1}$ vérifiant $|y_i - x_i| < r$ pour tout i, on a $(x_1, y_2, \dots, y_p) \in U$ et donc $(y_2, \dots, y_p) \in V$, d'où le résultat.

• V est convexe : soient $(x_2, \dots, x_p)$ et $(y_2, \dots, y_p)$ dans V et $\lambda \in [0,1]$. Il existe $x_1 \in \mathbb{R}$ et $y_1 \in \mathbb{R}$ tels que $(x_1, \dots, x_p) \in U$ et $(y_1, \dots, y_p) \in U$. Alors, par convexité de U,

$$(\lambda x_1 + (1-\lambda)y_1, \lambda x_2 + (1-\lambda)y_2, \dots, \lambda x_p + (1-\lambda)y_p) = \lambda(x_1, \dots, x_p) + (1-\lambda)(y_1, \dots, y_p) \in U,$$

et donc

$$\lambda(x_2, \dots, x_p) + (1-\lambda)(y_2, \dots, y_p) = (\lambda x_2 + (1-\lambda)y_2, \dots, \lambda x_p + (1-\lambda)y_p) \in V.$$

• Enfin, comme f est de classe $\mathcal{C}^1$ sur U, la formule définissant g montre que g est de classe $\mathcal{C}^1$ sur V, ce qui conclut la démonstration. □

5. Gradient

Si $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^1$, alors pour tout $a \in U$, on a le développement limité

$$f(a+h) \underset{\|h\| \to 0}{=} f(a) + \mathrm{d}f(a) \cdot h + o(\|h\|),$$

avec, pour tout $h = (h_1, \dots, h_p) \in \mathbb{R}^p$,

$$\mathrm{d}f(a) \cdot h = \sum_{i=1}^{p} \frac{\partial f}{\partial x_i}(a)\, h_i.$$

Dans $\mathbb{R}^p$ muni du produit scalaire canonique, ce terme se réinterprète comme un produit scalaire :

Propriété/Définition : Gradient

Soient $f : U \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ et $a \in U$. Le vecteur

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a) \right)$$

est appelé **gradient** de f en a. La fonction $\nabla f : U \to \mathbb{R}^p$ est appelée **gradient** de f.
Pour tout $h \in \mathbb{R}^p$, on a

$$\mathrm{d}f(a) \cdot h = (\nabla f(a) \,|\, h)$$

pour le produit scalaire canonique sur $\mathbb{R}^p$.

Remarque – La règle de la chaîne se réécrit, en notant $\gamma = (x_1, \dots, x_p)$:

$$\forall\, t \in I, \quad g'(t) = \big(\nabla f(\gamma(t)) \,|\, \gamma'(t)\big)\,.$$

Exemples

• D'après la loi de Fourier, la densité de flux de chaleur dans un matériau s'écrit $-\lambda\, \nabla T$, où T est la température et $\lambda > 0$ est la conductivité thermique.

• On note $\|\cdot\|$ la norme euclidienne canonique sur $\mathbb{R}^p$. L'application

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^p & \to & \mathbb{R} \\ x & \mapsto & \|x\|^2 \end{array} \right.$$

est de classe $\mathcal{C}^1$ sur $\mathbb{R}^p$. En effet, pour tout $x = (x_1, \dots, x_p) \in \mathbb{R}^p$,

$$f(x) = (x \,|\, x) = x_1^2 + \cdots + x_p^2;$$

l'application f est donc polynomiale. On a de plus, pour tout $a = (a_1, \dots, a_p) \in \mathbb{R}^p$,

$$\nabla f(a) = \left(\frac{\partial f}{\partial x_1}(a), \dots, \frac{\partial f}{\partial x_p}(a) \right) = (2a_1, \dots, 2a_p) = 2a.$$

• Revenons sur le calcul fait plus haut pour $g : (r, \theta) \mapsto f(r\cos(\theta), r\sin(\theta))$. Le calcul des dérivées partielles de g peut se mettre sous la forme

$$\begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta) \\ \dfrac{\partial g}{\partial \theta}(r,\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \sin(\theta) \\ -r\sin(\theta) & r\cos(\theta) \end{pmatrix} \begin{pmatrix} \dfrac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta)) \\ \dfrac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta)) \end{pmatrix}.$$

Or, pour tout $(r, \theta) \in \mathbb{R}^2$,

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

donc, si $r > 0$, la matrice précédente, notée $J(r, \theta)$, est inversible et on vérifie facilement que

$$J(r,\theta)^{-1} = \begin{pmatrix} \cos(\theta) & -\dfrac{1}{r}\sin(\theta) \\ \sin(\theta) & \dfrac{1}{r}\cos(\theta) \end{pmatrix}.$$

On a donc, pour tout $(x,y)=(r\cos(\theta),r\sin(\theta))$ tel que $r>0$,

$$\begin{pmatrix} \dfrac{\partial f}{\partial x}(r\cos(\theta),r\sin(\theta)) \\ \dfrac{\partial f}{\partial y}(r\cos(\theta),r\sin(\theta)) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\dfrac{1}{r}\sin(\theta) \\ \sin(\theta) & \dfrac{1}{r}\cos(\theta) \end{pmatrix} \begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta) \\ \dfrac{\partial g}{\partial \theta}(r,\theta) \end{pmatrix}$$
$$= \begin{pmatrix} \dfrac{\partial g}{\partial r}(r,\theta)\cos(\theta) - \dfrac{\partial g}{\partial \theta}(r,\theta)\dfrac{\sin(\theta)}{r} \\ \dfrac{\partial g}{\partial r}(r,\theta)\sin(\theta) + \dfrac{\partial g}{\partial \theta}(r,\theta)\dfrac{\cos(\theta)}{r} \end{pmatrix}.$$

Notons alors, pour tout $\theta\in\mathbb{R}$,

$$\vec{u}(\theta)=(\cos(\theta),\sin(\theta)),\quad \vec{v}(\theta)=(-\sin(\theta),\cos(\theta)).$$

La famille $(\vec{u}(\theta),\vec{v}(\theta))$ est une base orthonormée de $\mathbb{R}^2$ pour tout θ (la famille de fonctions $(\vec{u},\vec{v})$ est appelée repère polaire de $\mathbb{R}^2$). On remarque que pour tout $\theta\in\mathbb{R}$,

$$\vec{u}\,'(\theta)=\vec{v}(\theta),\quad \vec{v}\,'(\theta)=-\vec{u}(\theta).$$

Le calcul ci-dessus s'écrit alors : pour tout $(r,\theta)\in\mathbb{R}_+^*\times\mathbb{R}$ et $(x,y)=(r\cos(\theta),r\sin(\theta))$,

$$\nabla f(x,y)=\frac{\partial g}{\partial r}(r,\theta)\,\vec{u}(\theta)+\frac{1}{r}\frac{\partial g}{\partial \theta}(r,\theta)\,\vec{v}(\theta).$$

On parle de formule du gradient en coordonnées polaires.

III. Problèmes d'extrema

On recherche les extrema (c'est-à-dire la plus grande et/ou la plus petite valeur) d'une fonction $f:U\to\mathbb{R}$. Pour une fonction d'une variable réelle dérivable sur un intervalle I, on sait qu'en un point a où f atteint un extremum, si a est intérieur à I, on doit avoir $f'(a)=0$. Qu'en est-il pour les fonctions de plusieurs variables ?

Définition – Extremum

Soit $a\in U$.
- On dit que f admet un **minimum local** en a s'il existe $r>0$ tel que $B(a,r)\subset U$ et pour tout $x\in B(a,r)$, $f(x)\geqslant f(a)$.
- On dit que f admet un **minimum global** en a si pour tout $x\in U$, $f(x)\geqslant f(a)$.
- On dit que f admet un **maximum local** en a s'il existe $r>0$ tel que $B(a,r)\subset U$ et pour tout $x\in B(a,r)$, $f(x)\leqslant f(a)$.
- On dit que f admet un **maximum global** en a si pour tout $x\in U$, $f(x)\leqslant f(a)$.
- Enfin, un **extremum** est, par définition, un minimum ou un maximum.

Remarque – Évidemment, un extremum global est un extremum local du même type, et la réciproque est fausse. □

Supposons que f soit de classe $\mathcal{C}^1$ et qu'elle ait par exemple un minimum local en $a\in U$. Soit $(e_1,\dots,e_p)$ la base canonique de $\mathbb{R}^p$; pour tout $i\in[\![1,p]\!]$ et h assez petit, on a donc

$$f(a+he_i)\geqslant f(a),$$

et donc

$$\frac{f(a+he_i)-f(a)}{h}\geqslant 0\quad\text{si } h>0,\qquad \frac{f(a+he_i)-f(a)}{h}\leqslant 0\quad\text{si } h<0.$$

Lorsque h tend vers 0, on obtient respectivement $\dfrac{\partial f}{\partial x_i}(a) \geqslant 0$ et $\dfrac{\partial f}{\partial x_i}(a) \leqslant 0$. Finalement, pour tout $i \in [\![1,p]\!]$,

$$\frac{\partial f}{\partial x_i}(a) = 0.$$

On obtiendrait le même résultat avec un maximum local.

Définition – Point critique

Soient $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $a \in U$.
On dit que a est un **point critique** de f si $\nabla f(a) = (0,\dots,0)$. Ceci équivaut à

$$\forall\, i \in [\![1,p]\!], \quad \frac{\partial f}{\partial x_i}(a) = 0, \quad \text{ou encore à :} \quad \mathrm{d}f(a) = 0_{\mathscr{L}(\mathbb{R}^p,\mathbb{R})}.$$

Nous venons donc de montrer le résultat suivant :

Théorème – Condition nécessaire d'existence d'un extremum local

Si $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^1$ sur U et admet un extremum local en $a \in U$, alors a est un point critique de f : $\nabla f(a) = 0$.

Comme pour les fonctions d'une variable, cette condition n'est pas suffisante. Par exemple,

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^2 + y^2 - 4xy \end{array} \right.$$

est de classe $\mathcal{C}^1$ car polynomiale, et admet (0,0) comme point critique, car

$$\forall\,(x,y) \in \mathbb{R}^2, \quad \frac{\partial f}{\partial x}(x,y) = 2x - 4y \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = 2y - 4x.$$

Pourtant f n'a pas d'extremum local en (0,0). En effet, pour tout $x \in \mathbb{R}^*$,

$$f(x,0) = x^2 > 0 = f(0,0) \quad \text{tandis que} \quad f(x,x) = -2x^2 < 0.$$

Méthode – Si a est un point critique de f, pour vérifier si f a un extremum local en a, on peut considérer, pour h tel que $a + h \in U$,

$$f(a+h) - f(a).$$

- Si on prouve que cette expression a un signe constant pour h dans un voisinage de $(0,\dots,0)$, alors f a un extremum local en a.
- Si on trouve des points h arbitrairement proches de $(0,\dots,0)$ pour lesquels $f(a+h) - f(a) > 0$, et d'autres pour lesquels $f(a+h) - f(a) < 0$, alors f n'a pas d'extremum local en a. Pour cela, on procède souvent comme dans le contre-exemple ci-dessus, en cherchant des « directions » particulières.

Remarque – Si K est une partie fermée, bornée et non vide de $\mathbb{R}^p$, et si $f : K \to \mathbb{R}$ est continue, alors on sait que f est bornée et atteint ses bornes : en d'autres termes, f admet un minimum global et un maximum global sur K. Mais en général, le théorème des bornes atteintes ne permet pas de savoir *en quels points* de K ces bornes sont atteintes. La condition nécessaire ci-dessus permet de savoir, lorsque f est de classe $\mathcal{C}^1$ sur $\mathring{K}$ (intérieur de K, qui est un ouvert), en quels points de $\mathring{K}$ la fonction f est susceptible d'atteindre ses bornes. Mais il ne faut pas oublier qu'elles peuvent aussi être atteintes sur la frontière $Fr(K)$ de K. Il peut alors suffire de tester la valeur de f sur la frontière de K ainsi qu'aux éventuels points critiques de f dans $\mathring{K}$, pour déterminer les points en lesquels f atteint ses bornes.

On pourra garder en tête l'exemple très simple de la fonction $f : x \mapsto x$, continue sur le fermé borné non vide $[0,1]$ de $\mathbb{R}$. Elle atteint ses bornes en 0 et 1, qui ne sont pas des points critiques de f. Il n'y a pas de contradiction, car $[0,1]$ n'est pas ouvert, 0 et 1 sont sur sa frontière.

Exemple – On souhaite conditionner un produit en cartons d'une contenance de 1 litre. On se demande quelles sont les dimensions à donner à l'emballage afin d'utiliser le moins de carton possible (l'épaisseur du carton étant fixée).

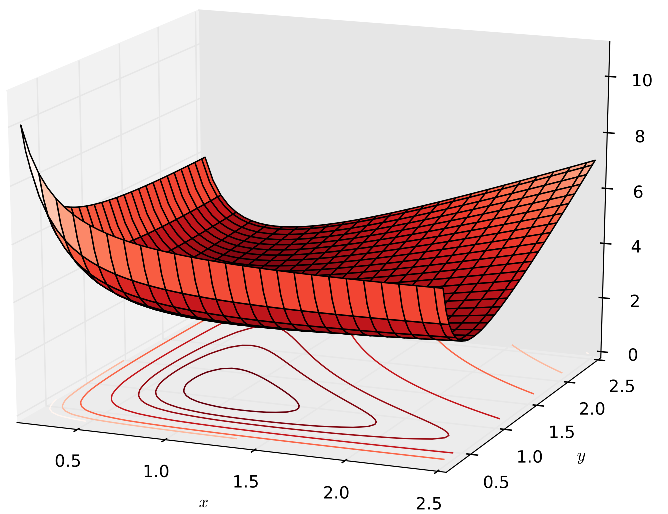
Soient x, y et z les longueurs (exprimées en décimètres) des trois côtés du carton, évidemment strictement positives. La contrainte sur le volume de l'emballage s'écrit

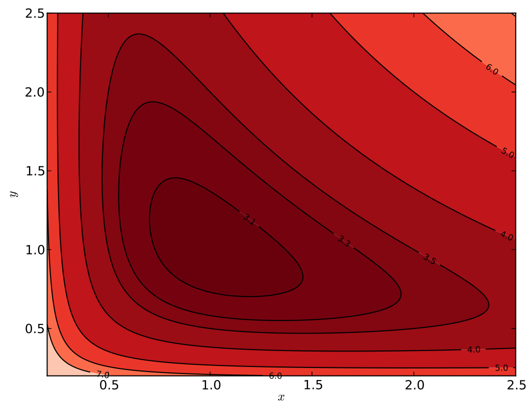
$$xyz = 1.$$

De plus, la surface utilisée est égale à

$$2(xy + yz + xz) = 2\left(xy + \frac{1}{x} + \frac{1}{y}\right) = 2S(x,y).$$

Le problème revient donc à déterminer l'éventuel minimum de S sur $U = (\mathbb{R}_+^*)^2$.
Voici la représentation graphique de S ainsi que certaines de ses lignes de niveaux, *i.e.*, les ensembles $\{(x,y) \in \mathbb{R}^2;\ S(x,y) = \lambda\}$ pour certaines valeurs de λ :





La fonction S est de classe $\mathcal{C}^1$ sur U et pour tout $(x,y) \in U$,

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

Il s'ensuit immédiatement que S possède un unique point critique sur U, égal à $(1,1)$. On va montrer que S possède un minimum global sur U en $(1,1)$.

On remarque que $S(1,1) = 3$ et que l'on a $S(x,y) > 3$ si $x < 1/3$ ou $y < 1/3$ ou $xy < 3$. Définissons

$$K = \{(x,y) \in (\mathbb{R}_+^*)^2;\ x \geqslant 1/3,\ y \geqslant 1/3,\ xy \leqslant 3\},$$

de sorte que $S(x,y) > 3$ si $(x,y) \notin K$. De plus, K est non vide, fermé (intersection de trois fermés, par continuité des fonctions $(x,y) \mapsto x - 1/3$, $(x,y) \mapsto y - 1/3$ et $(x,y) \mapsto 3 - xy$) et K est borné : si $(x,y) \in K$,

$$\frac{1}{3} \leqslant x \leqslant \frac{3}{y} \leqslant 9$$

et de même pour y. La fonction S a donc un minimum global sur K ; de plus, si $(x,y) \notin K$, $S(x,y) > 3 = S(1,1)$, donc S admet un minimum global sur U, qui doit être un point critique de S, c'est-à-dire $(1,1)$.

Finalement, la fonction S a un minimum global sur U en $(1,1)$, *i.e.* pour $x = y = z = 1$. L'emballage le plus économique répondant aux contraintes données est le cube de côté 10 cm. La surface utilisée correspond à $2S(1,1) = 6$ (elle vaut donc $6\,dm^2$).

IV. Dérivées partielles d'ordre 2

Définition

• On dit que f est de classe $\mathcal{C}^2$ sur U si f est de classe $\mathcal{C}^1$ sur U et si les dérivées partielles $\dfrac{\partial f}{\partial x_i}$, pour $i \in [\![1,p]\!]$, sont de classe $\mathcal{C}^1$ sur U.

• Si f est de classe $\mathcal{C}^2$ sur U, alors pour tout $(i,j) \in [\![1,p]\!]^2$,

$$\frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_j}\right) \quad \text{est notée} \quad \frac{\partial^2 f}{\partial x_i \partial x_j} \quad \text{ou} \quad \partial^2_{i,j} f.$$

Lorsque $i = j$, on écrit simplement

$$\frac{\partial^2 f}{\partial {x_i}^2} \quad \text{au lieu de} \quad \frac{\partial^2 f}{\partial x_i \, \partial x_i}.$$

Ces fonctions sont appelées **dérivées partielles** d'ordre 2 de f.

Remarque – On généralisera sans difficulté les résultats concernant les opérations sur les fonctions de classe $\mathcal{C}^1$ (combinaison linéaire, produit, quotient) aux fonctions de classe $\mathcal{C}^2$. De plus, les applications linéaires, les fonctions polynomiales, et les fractions rationnelles dont le dénominateur ne s'annule pas, sont de classe $\mathcal{C}^2$.

Théorème de Schwarz (admis : démonstration hors programme)

Si f est de classe $\mathcal{C}^2$ sur U, alors pour tout $(i,j) \in [\![1,p]\!]^2$,

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

Exemple – Soit

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R} \\ (x,y) & \mapsto & x^4 + y^3 - 5x^2 y \end{array} \right.$$

La fonction f est de classe $\mathcal{C}^\infty$ sur $\mathbb{R}^2$ car elle est polynomiale. Pour tout $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x,y) = 4x^3 - 10xy, \qquad \frac{\partial f}{\partial y}(x,y) = 3y^2 - 5x^2,$$

et en ce qui concerne les dérivées d'ordre 2 :

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 12x^2 - 10y, \quad \frac{\partial^2 f}{\partial y^2}(x,y) = 6y \quad \text{et} \quad \frac{\partial^2 f}{\partial x \, \partial y}(x,y) = \frac{\partial^2 f}{\partial y \, \partial x}(x,y) = -10x.$$

V. Résolution d'équations aux dérivées partielles

De très nombreux phénomènes physiques, chimiques, biologiques, économiques sont modélisables par des équations aux dérivées partielles, c'est-à-dire, par une relation entre les différentes dérivées partielles (d'ordre 1 ou 2 très souvent) d'une certaine quantité. Elles sont souvent associées à une condition initiale et/ou une condition « au bord », c'est-à-dire sur la frontière du domaine d'espace.

Exemples

• L'équation de Poisson

$$\Delta f(x) = g(x) \quad \text{pour } x \in U \subset \mathbb{R}^p$$

où $\Delta f = \sum_{i=1}^{p} \frac{\partial^2 f}{\partial {x_i}^2}$ est le laplacien de f, intervient par exemple en électrostatique ; g correspond à la distribution de charges, et f est le potentiel associé.

• L'équation des ondes

$$\frac{\partial^2 f}{\partial t^2}(t,x) = c^2\, \Delta f(t,x) \quad \text{pour } (t,x) \in U \subset \mathbb{R} \times \mathbb{R}^p,$$

modélise la propagation d'une onde (par exemple, sonore ou électromagnétique) dans $\mathbb{R}^p$ ($c > 0$ est la vitesse ou célérité). Ici, Δf désigne le laplacien de f par rapport aux variables d'espace représentées par x. Lorsque $p = 1$, on obtient l'équation

$$\frac{\partial^2 f}{\partial t^2} = c^2\, \frac{\partial^2 f}{\partial x^2},$$

qui modélise par exemple la vibration unidirectionnelle d'une corde infinie, $f(t,x)$ représentant le déplacement au temps t du point de la corde d'abscisse x.

• L'équation du transfert thermique, ou équation de la chaleur,

$$\frac{\partial f}{\partial t}(t,x) = D\, \Delta f(t,x) \quad \text{pour } (t,x) \in U \subset \mathbb{R} \times \mathbb{R}^p,$$

modélise l'évolution de la température f dans un milieu ($D > 0$ est le coefficient de diffusivité thermique). Ici aussi, Δf désigne le laplacien de f par rapport aux variables d'espace.

• L'équation du transport

$$\frac{\partial f}{\partial t}(t,x) + c(t,x)\, \frac{\partial f}{\partial x}(t,x) = 0 \quad \text{pour } (t,x) \in U \subset \mathbb{R} \times \mathbb{R}^p,$$

où $c : U \to \mathbb{R}$ est continue.

Un principe important de résolution d'équations aux dérivées partielles est d'effectuer un changement de variable ; si $f : U \subset \mathbb{R}^p \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^1$ (ou $\mathcal{C}^2$ pour une équation d'ordre 2) solution d'une équation aux dérivées partielles, on écrit, pour $x = (x_1, \ldots, x_p) \in U$,

$$f(x) = g(u_1(x), \ldots, u_p(x))$$

où $u_1, \ldots, u_p$ sont des fonctions définies sur U, à valeurs dans un ouvert V de $\mathbb{R}^p$, et $g : V \to \mathbb{R}$.

Pour que cela définisse correctement la fonction g, on choisit les fonctions $u_1, \ldots, u_p$ et l'ouvert V de sorte que

$$\phi : \left\{ \begin{array}{ccc} U & \to & V \\ x & \mapsto & (u_1(x), \ldots, u_p(x)) \end{array} \right.$$

soit bijective. Ainsi, la relation $f = g \circ \phi$ que l'on veut utiliser équivaut à $g = f \circ \phi^{-1}$. On souhaite également que g et ϕ soient de classe $\mathcal{C}^1$ (resp. $\mathcal{C}^2$ pour l'ordre 2), ce qui est le cas si toutes les fonctions coordonnées de ϕ et ϕ^{-1} sont de classe $\mathcal{C}^1$ (resp. $\mathcal{C}^2$), par application de la règle de la chaîne.

Le changement de variable est choisi pour que g vérifie une équation aux dérivées partielles la plus simple possible. Lorsque $p = 2$, on se ramène par exemple à l'une des équations suivantes (les variables de la fonction g sont notées y_1 et y_2) :

• $\frac{\partial g}{\partial y_1} = 0$ sur V.

Si V est convexe, on sait que cela entraîne que g ne dépend pas de sa première variable, et qu'il existe F de classe $\mathcal{C}^1$ sur un intervalle ouvert de $\mathbb{R}$ tel que, pour tout $(y_1,y_2) \in V$, $g(y_1,y_2) = F(y_2)$. La réciproque est évidente. On peut bien sûr adapter avec la deuxième variable.

- $\dfrac{\partial^2 g}{\partial y_1 \, \partial y_2} = 0$ sur V.

De même, si V est convexe, ceci entraîne l'existence de $\tilde{G}$ de classe $\mathcal{C}^1$ sur un intervalle ouvert de $\mathbb{R}$ tel que, pour tout $(y_1,y_2) \in V$,

$$\frac{\partial g}{\partial y_2}(y_1,y_2) = \tilde{G}(y_2).$$

En notant G une primitive de $\tilde{G}$ sur cet intervalle, la fonction

$$\tilde{g} : (y_1,y_2) \mapsto g(y_1,y_2) - G(y_2)$$

vérifie $\dfrac{\partial \tilde{g}}{\partial y_2} = 0$ sur V. D'après le premier point, la fonction $\tilde{g}$ ne dépend pas de sa deuxième variable, et finalement g se met sous la forme

$$g : (y_1,y_2) \mapsto F(y_1) + G(y_2)$$

où F et G sont de classe $\mathcal{C}^2$ sur des intervalles ouverts de $\mathbb{R}$. La réciproque est évidente.

Dans chaque cas, on en déduit f par la relation $f = g \circ \phi$, puis on vérifie la réciproque. On peut aussi, dans certains cas, raisonner entièrement par équivalences.

Voici deux exemples fondamentaux de changements de variables qu'il faut savoir utiliser, avec, dans chaque cas, un exemple détaillé :

▶ **Transformation affine**

Soit ψ un isomorphisme de $\mathbb{R}^p$ sur $\mathbb{R}^p$. Une transformation affine consiste à choisir

$$\phi : x \in U \mapsto \psi(x) + a,$$

où $a \in \mathbb{R}^p$. Dans ce cas, ϕ est une bijection, et $\phi^{-1} : y \mapsto \psi^{-1}(y - a)$, dont chaque fonction coordonnée est de classe $\mathcal{C}^1$ (et même $\mathcal{C}^2$) sur $V = \phi(U)$, car linéaire. Par exemple, dans $\mathbb{R}^2$, une transformation affine convenable pour effectuer un changement de variable est une application de la forme

$$(x_1,x_2) \mapsto (\alpha x_1 + \beta x_2 + a_1, \gamma x_1 + \delta x_2 + a_2)$$

avec $\alpha\delta - \beta\gamma \neq 0$.

Exemple – On cherche à déterminer toutes les fonctions f de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ telles que pour tout $(x, y) \in \mathbb{R}^2$,

$$\frac{\partial^2 f}{\partial x^2}(x, y) - 3\frac{\partial^2 f}{\partial x \, \partial y}(x, y) + 2\frac{\partial^2 f}{\partial y^2}(x, y) = 0. \qquad (E)$$

Le changement de variable utilisé ici est donné par la fonction

$$\phi : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \to & \mathbb{R}^2 \\ (x, y) & \mapsto & (x + y, 2x + y) \end{array} \right.$$

La fonction ϕ est linéaire, c'est une bijection de $\mathbb{R}^2$ sur $\mathbb{R}^2$ car, pour tout $(x, y) \in \mathbb{R}^2$ et $(u, v) \in \mathbb{R}^2$,

$$\phi(x, y) = (u, v) \quad \Leftrightarrow \quad \begin{cases} x + y = u \\ 2x + y = v \end{cases} \quad \Leftrightarrow \quad \begin{cases} x = v - u \\ y = u - x = 2u - v \end{cases}$$

Ainsi, ϕ est un isomorphisme de $\mathbb{R}^2$ sur $\mathbb{R}^2$, et pour tout $(u, v) \in \mathbb{R}^2$, $\phi^{-1}(u, v) = (-u+v, 2u-v)$.

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction et

$$g : (u, v) \mapsto (f \circ \phi^{-1})(u, v) = f(-u + v, 2u - v),$$

de sorte que pour tout $(x, y) \in \mathbb{R}^2$,

$$f(x, y) = (g \circ \phi)(x, y) = g(x + y, 2x + y).$$

D'après la règle de la chaîne pour les fonctions de deux variables, g est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ si et seulement si f est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$, et dans ce cas, pour tout $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial g}{\partial u}(x+y,2x+y)+2\frac{\partial g}{\partial v}(x+y,2x+y), \quad \frac{\partial f}{\partial y}(x,y) = \frac{\partial g}{\partial u}(x+y,2x+y)+\frac{\partial g}{\partial v}(x+y,2x+y),$$

et, en tenant compte du théorème de Schwarz,

$$\begin{aligned}\frac{\partial^2 f}{\partial x^2}(x,y) &= \frac{\partial^2 g}{\partial u^2}(x+y,2x+y) + 2\frac{\partial^2 g}{\partial v\,\partial u}(x+y,2x+y) + 2\frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) + 4\frac{\partial^2 g}{\partial v^2}(x+y,2x+y)\\ &= \frac{\partial^2 g}{\partial u^2}(x+y,2x+y) + 4\frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) + 4\frac{\partial^2 g}{\partial v^2}(x+y,2x+y),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 f}{\partial x\,\partial y}(x,y) &= \frac{\partial^2 g}{\partial u^2}(x+y,2x+y) + 2\frac{\partial^2 g}{\partial v\,\partial u}(x+y,2x+y) + \frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) + 2\frac{\partial^2 g}{\partial v^2}(x+y,2x+y)\\ &= \frac{\partial^2 g}{\partial u^2}(x+y,2x+y) + 3\frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) + 2\frac{\partial^2 g}{\partial v^2}(x+y,2x+y),\end{aligned}$$

$$\begin{aligned}\frac{\partial^2 f}{\partial y^2}(x,y) &= \frac{\partial^2 g}{\partial u^2}(x+y,2x+y) + \frac{\partial^2 g}{\partial v\,\partial u}(x+y,2x+y) + \frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) + \frac{\partial^2 g}{\partial v^2}(x+y,2x+y)\\ &= \frac{\partial^2 g}{\partial u^2}(x+y,2x+y) + 2\frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) + \frac{\partial^2 g}{\partial v^2}(x+y,2x+y).\end{aligned}$$

Après simplifications, on obtient que f est solution de (E) sur $\mathbb{R}^2$ si et seulement si g est de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ avec, pour tout $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial^2 g}{\partial u\,\partial v}(x+y,2x+y) = 0.$$

L'image de ϕ étant $\mathbb{R}^2$, ceci équivaut au fait que g soit de classe $\mathcal{C}^2$ sur $\mathbb{R}^2$ avec

$$\frac{\partial^2 g}{\partial u\,\partial v} = 0.$$

En refaisant le raisonnement du deuxième point ci-dessus ($\mathbb{R}^2$ est convexe), on obtient que ceci équivaut à l'existence de deux fonctions F et G de $\mathbb{R}$ dans $\mathbb{R}$ de classe $\mathcal{C}^2$ telles que pour tout $(u,v) \in \mathbb{R}^2$,

$$g(u,v) = F(u) + G(v).$$

Les solutions de (E) sur $\mathbb{R}^2$ sont donc exactement les fonctions de la forme

$$(x,y) \mapsto F(x+y) + G(2x+y)$$

pour F et G deux fonctions quelconques de classe $\mathcal{C}^2$ sur $\mathbb{R}$ à valeurs dans $\mathbb{R}$.

▶ **Coordonnées polaires**

Soit $\theta_0 \in [-\pi,\pi[$. Notons $U = \mathbb{R}^2 \setminus D$, où D est la demi-droite

$$\{(x,y) \in \mathbb{R}^2;\ \arg(x+iy) = \theta_0\}$$

(en considérant que $0 \in D$). Si $(x,y) \in U$, il existe un unique $r > 0$ et un unique $\theta \in]\theta_0,\theta_0+2\pi[$ tels que

$$(x,y) = (r\cos(\theta), r\sin(\theta)).$$

La fonction

$$\psi : \left\{ \begin{array}{ccc}]0,+\infty[\times]\theta_0,\theta_0+2\pi[& \to & U \\ (r,\theta) & \mapsto & (r\cos(\theta), r\sin(\theta)) \end{array} \right.$$

est bijective, de classe $\mathcal{C}^1$ sur l'ouvert $V =]0,+\infty[\times]\theta_0,\theta_0+2\pi[$.

Pour tout $(r,\theta) \in V$, l'égalité $(x,y) = \psi(r,\theta)$ entraîne que $r = \sqrt{x^2+y^2}$ et donc

$$\cos(\theta) = \frac{x}{r} = \frac{x}{\sqrt{x^2+y^2}}, \qquad \sin(\theta) = \frac{y}{r} = \frac{y}{\sqrt{x^2+y^2}}.$$

Par exemple, lorsque $\theta_0 = -\pi$, soit $(x,y) = \psi(r,\theta) \in U$ avec $(r,\theta) \in V$. En se restreignant à $x > 0$, on a

$$\tan(\theta) = \frac{y}{x} \quad \text{avec} \quad \theta \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[,$$

donc

$$\psi^{-1}(x,y) = \left(\sqrt{x^2+y^2}, \arctan\left(\frac{y}{x}\right)\right),$$

ce qui montre que chaque fonction-coordonnée de $\phi = \psi^{-1}$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}_+^* \times \mathbb{R}$. Selon les situations, on pourra considérer une autre valeur de θ_0 et/ou adapter les formules précédentes.

Le passage en coordonnées polaires, c'est-à-dire le changement de variable défini par la fonction ϕ précédente (ou une forme analogue selon l'ouvert sur lequel on travaille), permet de résoudre un certain nombre d'équations aux dérivées partielles.

Remarque – La formule donnant $\phi = \psi^{-1}$ dépend de l'ouvert sur lequel on travaille. Il est parfois plus simple de travailler avec la fonction ψ, c'est-à-dire, à partir de la relation

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

Exemple – Soit $U = \mathbb{R}_+^* \times \mathbb{R}$. On recherche toutes les fonctions $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ telles que, pour tout $(x,y) \in U$, $\nabla f(x,y)$ soit colinéaire à (x,y). Cette condition équivaut au fait que pour tout $(x,y) \in U$,

$$\begin{vmatrix} \dfrac{\partial f}{\partial x}(x,y) & x \\ \dfrac{\partial f}{\partial y}(x,y) & y \end{vmatrix} = 0, \quad i.e., \quad y\frac{\partial f}{\partial x}(x,y) - x\frac{\partial f}{\partial y}(x,y) = 0.$$

On résout cette équation aux dérivées partielles en passant en coordonnées polaires : avec les notations ci-dessus, on pose, pour f solution du problème,

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

pour tout $(r,\theta) \in V =]0,+\infty[\, \times \,]-\frac{\pi}{2}, \frac{\pi}{2}[$. D'après la règle de la chaîne, la fonction g est de classe $\mathcal{C}^1$ sur V et pour tout $(r,\theta) \in V$,

$$\frac{\partial g}{\partial \theta}(r,\theta) = \frac{\partial f}{\partial x}(r\cos(\theta), r\sin(\theta))\,(-r\sin(\theta)) + \frac{\partial f}{\partial y}(r\cos(\theta), r\sin(\theta))\,(r\cos(\theta)) = 0;$$

V étant convexe, la fonction g est donc indépendante de θ : il existe $\tilde{F} :]0,+\infty[\to \mathbb{R}$ de classe $\mathcal{C}^1$ telle que, pour tout $(r,\theta) \in V$, $g(r,\theta) = \tilde{F}(r)$. La fonction ψ étant une bijection de V sur U, pour tout $(x,y) \in U$, il existe $(r,\theta) \in V$ tel que $(x,y) = \psi(r,\theta)$, et alors

$$f(x,y) = (g \circ \phi)(x,y) = \tilde{F}(\sqrt{x^2+y^2}) = F(x^2+y^2),$$

où $F : r \mapsto \tilde{F}(\sqrt{r}\,)$ est de classe $\mathcal{C}^1$ sur $]0,+\infty[$ par composition. Réciproquement, soit f une fonction de la forme précédente. Alors f est de classe $\mathcal{C}^1$ sur U par composition et, pour tout $(x,y) \in U$,

$$y\frac{\partial f}{\partial x}(x,y) - x\frac{\partial f}{\partial y}(x,y) = y \times 2x\,F'(x^2+y^2) - x \times 2y\,F'(x^2+y^2) = 0.$$

Les solutions du problème sont donc exactement les fonctions de la forme

$$(x,y) \mapsto F(x^2+y^2)$$

avec $F :]0,+\infty[\to \mathbb{R}$ de classe $\mathcal{C}^1$. Ce sont des fonctions « isotropes », c'est-à-dire, dépendant de (x,y) uniquement via sa norme euclidienne usuelle.

VI. Courbes et surfaces

Dans le chapitre **Fonctions vectorielles – Arcs paramétrés**, nous avons étudié les courbes données par une représentation paramétrique, et nous avons notamment décrit l'allure locale d'une telle courbe, en lien avec la notion de tangente. Dans cette partie, nous allons étudier le cas d'une courbe ou surface définie par une équation cartésienne de la forme $f(x,y) = 0$ ou $f(x,y,z) = 0$.

On travaillera dans $\mathbb{R}^2$ ou $\mathbb{R}^3$ munis de leur structure euclidienne canonique.

1. Courbes du plan données par une équation cartésienne

Dans ce paragraphe, $p = 2$. Dans de nombreuses situations, une courbe $\mathcal{C}$ du plan n'est pas donnée par un paramétrage, mais par une équation cartésienne, c'est-à-dire que $\mathcal{C}$ est l'ensemble des points (x,y) de U tels que $f(x,y) = 0$ où $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^1$. Il se peut que ceci définisse la courbe de façon implicite, car il faut *a priori* résoudre une équation pour tracer cette courbe. On peut citer l'exemple des courbes équipotentielles, isoclines, de même altitude, etc ...

Exemples

• Si $\phi : I \to \mathbb{R}$ est de classe $\mathcal{C}^1$ sur un intervalle ouvert I de $\mathbb{R}$, le graphe de ϕ est la courbe $\mathcal{C}$ d'équation $y = \phi(x)$, c'est-à-dire $f(x,y) = 0$ avec $f : (x,y) \mapsto \phi(x) - y$, de classe $\mathcal{C}^1$ sur l'ouvert $U = I \times \mathbb{R}$. Dans ce cas, la représentation est explicite car y est fonction de x. La courbe $\mathcal{C}$ est en fait l'image de l'arc paramétré (I,γ) où, pour tout $x \in I$, $\gamma(x) = (x,\phi(x))$.

• Le cercle unité $\mathcal{C}$ de $\mathbb{R}^2$ a pour équation $x^2 + y^2 = 1$. On peut choisir $f(x,y) = x^2 + y^2 - 1$ pour tout $(x,y) \in \mathbb{R}^2$. $\square$

Dans l'exemple précédent, on remarquera que l'on peut entièrement résoudre l'équation, mais ce n'est pas toujours possible. On souhaite trouver un moyen de décrire, malgré cela, la courbe $\mathcal{C}$. On sait notamment le faire pour les arcs paramétrés : si $\Gamma = (I,\gamma)$ est un arc de classe $\mathcal{C}^1$, Γ possède, en chaque point régulier $M(t_0)$, une tangente dirigée par $\gamma'(t_0)$ (on rappelle que $M(t_0)$ est un point régulier de Γ si et seulement si $\gamma'(t_0) \neq 0$). On souhaite se ramener à cette situation.

Définition – Point régulier

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $\mathcal{C}$ la partie de $\mathbb{R}^2$ d'équation $f(x,y) = 0$.
On appelle **point régulier** de $\mathcal{C}$ tout point $(x_0,y_0) \in \mathcal{C}$ tel que

$$\nabla f(x_0,y_0) \neq (0,0),$$

c'est-à-dire, tel que $\dfrac{\partial f}{\partial x}(x_0,y_0) \neq 0$ ou $\dfrac{\partial f}{\partial y}(x_0,y_0) \neq 0$.
En d'autres termes, il s'agit des points de $\mathcal{C}$ qui ne sont pas des points critiques de f.

Théorème (admis)

Avec les notations précédentes, soit (x_0,y_0) un point régulier de $\mathcal{C}$.
Alors il existe $r > 0$, $\eta > 0$ et $\gamma :]-\eta,\eta[\to U$ de classe $\mathcal{C}^1$, tels que :

- $(x_0,y_0) = \gamma(0)$;
- $(]-\eta,\eta[,\gamma)$ soit un arc paramétré simple et régulier ;
- $B((x_0,y_0),r) \subset U$ et pour tout $(x,y) \in B((x_0,y_0),r)$, on a l'équivalence :

$$f(x,y) = 0 \quad \Leftrightarrow \quad \exists\, t \in]-\eta,\eta[\,;\ (x,y) = \gamma(t).$$

On dit que $(]-\eta,\eta[,\gamma)$ est un paramétrage local de $\mathcal{C}$ au voisinage de (x_0,y_0).

Exemple – Soit $\mathcal{C}$ le cercle unité d'équation $f(x,y)=0$ avec $f:(x,y)\mapsto x^2+y^2-1$. La fonction f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et pour tout $(x,y)\in\mathbb{R}^2$,

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

qui est non nul sauf à l'origine, qui n'est pas un point de $\mathcal{C}$. Tous les points de $\mathcal{C}$ sont donc réguliers. En fait, dans ce cas, on peut construire explicitement un paramétrage de $\mathcal{C}$ au voisinage de chaque point. Par exemple, au voisinage d'un point (a,b) de $\mathcal{C}$ tel que $b>0$, on peut paramétrer $\mathcal{C}$ par

$$\gamma_1:\left\{\begin{array}{rcl}[-a-1,-a+1] & \to & \mathbb{R}^2\\ t & \mapsto & (a+t,\sqrt{1-(a+t)^2})\end{array}\right.$$

Au voisinage de $(1,0)$, on peut paramétrer $\mathcal{C}$ par

$$\gamma_2:\left\{\begin{array}{rcl}[-1,1] & \to & \mathbb{R}^2\\ t & \mapsto & (\sqrt{1-t^2},t)\end{array}\right.$$

On peut procéder de même dans les autres cas.

Il est important de bien comprendre les différentes notions de point régulier selon le type de courbe considéré, et les liens entre ces notions :

Bilan

• Si la courbe est donnée par un paramétrage (I,γ) où $\gamma:I\to\mathbb{R}^2$ est de classe $\mathcal{C}^1$, un point $M(t)$ est régulier si et seulement si $\gamma'(t)\neq 0$.

• Si la courbe est donnée par une équation cartésienne $f(x,y)=0$ où $f:U\to\mathbb{R}$ est de classe $\mathcal{C}^1$, un point (x_0,y_0) de $\mathcal{C}$ est régulier si et seulement si $\nabla f(x_0,y_0)\neq(0,0)$.

Le résultat admis ci-dessus montre que si l'on est dans la situation du deuxième point, on est également dans la situation du premier : au voisinage d'un point régulier au sens du deuxième point, une courbe donnée de façon implicite peut être « explicitée », et être vue comme l'image d'un arc paramétré régulier (on peut écrire x et y comme fonctions d'un paramètre t). □

Soit maintenant $\mathcal{C}$ une partie de $\mathbb{R}^2$ donnée par une équation cartésienne $f(x,y)=0$, et supposons que l'on soit dans le cadre d'application du théorème précédent en un point (x_0,y_0), c'est-à-dire que f est de classe $\mathcal{C}^1$ sur U avec $\nabla f(x_0,y_0)\neq(0,0)$. Avec les notations du théorème, et en notant $\gamma=(x,y)$, on a par définition même, pour tout $t\in\,]-\eta,\eta[$,

$$f(x(t),y(t))=0.$$

D'après la règle de la chaîne, ceci définit une fonction de classe $\mathcal{C}^1$ sur $]-\eta,\eta[$ et, pour tout $t\in\,]-\eta,\eta[$,

$$\frac{\partial f}{\partial x}(x(t),y(t))\,x'(t)+\frac{\partial f}{\partial y}(x(t),y(t))\,y'(t)=0,\quad i.e.\quad \big(\nabla f(\gamma(t))\,|\,\gamma'(t)\big)=0,$$

et donc, pour $t=0$,

$$\frac{\partial f}{\partial x}(x_0,y_0)\,x'(0)+\frac{\partial f}{\partial y}(x_0,y_0)\,y'(0)=0,\quad i.e.\quad \big(\nabla f(x_0,y_0)\,|\,\gamma'(0)\big)=0.$$

La tangente à la courbe $\mathcal{C}$ au point (x_0,y_0) (en tant que support d'un arc paramétré simple et régulier au voisinage de ce point) est la droite passant par (x_0,y_0) et dirigée par $\gamma'(0)$. Or, l'égalité précédente montre que le vecteur (non nul) $\nabla f(x_0,y_0)$ est orthogonal à $\gamma'(0)$: $\nabla f(x_0,y_0)$ est un vecteur normal à la tangente.

La tangente à la courbe $\mathcal{C}$ au point (x_0,y_0) est donc l'ensemble des points $(x,y)\in\mathbb{R}^2$ tels que

$$(\nabla f(x_0,y_0)\,|\,(x-x_0,y-y_0))=0.$$

On a ainsi démontré le résultat suivant :

Propriété/Définition : Tangente en un point régulier

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $\mathcal{C}$ la partie de $\mathbb{R}^2$ d'équation $f(x,y) = 0$.
Soit (x_0,y_0) un point régulier de $\mathcal{C}$.
• La courbe $\mathcal{C}$ possède une tangente en (x_0,y_0), d'équation

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

Si $\dfrac{\partial f}{\partial x}(x_0,y_0) = 0$, il s'agit d'une tangente horizontale, si $\dfrac{\partial f}{\partial y}(x_0,y_0) = 0$, il s'agit d'une tangente verticale.

• On appelle **normale** à $\mathcal{C}$ au point (x_0,y_0), la droite passant par (x_0,y_0) et dirigée par le vecteur

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

On dit que $\nabla f(x_0,y_0)$ est orthogonal à $\mathcal{C}$ au point (x_0,y_0).

Exemple – Soit $\mathcal{C}$ la partie de $\mathbb{R}^2$ d'équation

$$x^3 + 3y^2 + 6xy + 4 = 0.$$

Elle a pour équation cartésienne $f(x,y) = 0$ où $f : (x,y) \mapsto x^3 + 3y^2 + 6xy + 4$ est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. Pour tout $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x,y) = 3x^2 + 6y \quad \text{et} \quad \frac{\partial f}{\partial y}(x,y) = 6y + 6x.$$

On a les équivalences

$$\begin{cases} 3x^2 + 6y = 0 \\ 6y + 6x = 0 \end{cases} \Leftrightarrow \begin{cases} x^2 + 2y = 0 \\ y = -x \end{cases} \Leftrightarrow \begin{cases} x^2 - 2x = 0 \\ y = -x \end{cases}$$
$$\Leftrightarrow \quad (x,y) = (0,0) \quad \text{ou} \quad (x,y) = (2,-2).$$

Les points critiques de f sont donc $(0,0)$ et $(2,-2)$. De ces deux points, seul $(2,-2)$ appartient à $\mathcal{C}$. Tout autre point de $\mathcal{C}$ est donc régulier, et l'équation de la tangente à $\mathcal{C}$ en l'un de ses points réguliers (x_0,y_0) est

$$\frac{\partial f}{\partial x}(x_0,y_0)\,(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)\,(y - y_0) = 0$$
$$i.e. \quad (x_0^2 + 2y_0)(x - x_0) + 2(x_0 + y_0)(y - y_0) = 0.$$

Cas particulier – Si ϕ est une fonction de classe $\mathcal{C}^1$ sur un intervalle ouvert I, son graphe $\mathcal{C}$ est la courbe d'équation $f(x,y) = 0$ avec $f : (x,y) \mapsto \phi(x) - y$, de classe $\mathcal{C}^1$ sur $U = I \times \mathbb{R}$. Tout point de $\mathcal{C}$ est régulier car

$$\nabla f(x,y) = (\phi'(x), -1) \neq (0,0)$$

pour tout $(x,y) \in \mathcal{C}$. La tangente à $\mathcal{C}$ en un point (x_0,y_0) a pour équation

$$\frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0) = 0,$$
$$i.e. \quad \phi'(x_0)(x - x_0) - (y - y_0) = 0.$$

Sachant que $y_0 = \phi(x_0)$, on retrouve bien sûr l'équation

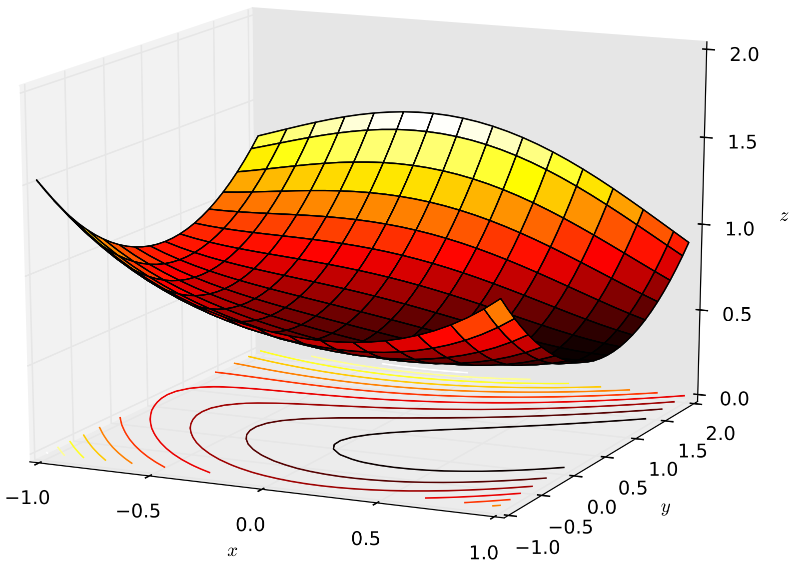
$$y = \phi'(x_0)(x - x_0) + \phi(x_0).$$

Définition – Ligne de niveau

Soit $f : U \to \mathbb{R}$ et $\lambda \in \mathbb{R}$.
On appelle **ligne de niveau** λ de f la partie de U d'équation $f(x,y) = \lambda$.

Exemple – Voici le graphe et certaines lignes de niveau de la fonction

$$f : (x,y) \mapsto \frac{1}{2} + \frac{1}{10}\left((1-x)^2 + \left(\frac{3}{2}\,y - x^2\right)^2\right)$$



Remarque – Bien sûr, l'étude des lignes de niveau de f entre dans le cadre précédent, *via* l'étude de l'équation $f(x,y) - \lambda = 0$. Les points réguliers de cette ligne de niveau sont ses points en lesquels ∇f ne s'annule pas, puisque $\nabla(f - \lambda) = \nabla f$.

Propriété – Gradient et lignes de niveau

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $\lambda \in \mathbb{R}$. Soit (x_0,y_0) un point régulier de la ligne de niveau λ de f, c'est-à-dire que $f(x_0,y_0) = \lambda$ et $\nabla f(x_0,y_0) \neq (0,0)$.
Alors $\nabla f(x_0,y_0)$ est orthogonal à la ligne de niveau λ de f, et orienté dans le sens des valeurs croissantes de f, c'est-à-dire qu'il existe $\eta > 0$ tel que la fonction

$$t \mapsto f((x_0,y_0) + t\,\nabla f(x_0,y_0)) = f\left(x_0 + t\,\frac{\partial f}{\partial x}(x_0,y_0), y_0 + t\,\frac{\partial f}{\partial y}(x_0,y_0)\right)$$

soit strictement croissante sur $]-\eta,\eta[$.

Démonstration – La première conclusion est déjà connue, d'après la propriété et la remarque précédentes. Pour la seconde conclusion, la fonction f est de classe $\mathcal{C}^1$ sur U et les fonctions

$$x : t \mapsto x_0 + t\,\frac{\partial f}{\partial x}(x_0,y_0) \quad \text{et} \quad y : t \mapsto y_0 + t\,\frac{\partial f}{\partial y}(x_0,y_0)$$

sont de classe $\mathcal{C}^1$ sur $\mathbb{R}$ et ont pour limites respectives x_0 et y_0 en 0, le point (x_0,y_0) appartenant à l'ouvert U. La fonction

$$g : t \mapsto f((x_0,y_0) + t\,\nabla f(x_0,y_0)) = f(x(t),y(t))$$

est donc bien définie au voisinage de 0, et d'après la règle de la chaîne, elle est de classe $\mathcal{C}^1$ au voisinage de 0 avec, pour tout $t \in \mathbb{R}$ assez proche de 0,

$$\begin{aligned} g'(t) &= \frac{\partial f}{\partial x}(x(t),y(t))\, x'(t) + \frac{\partial f}{\partial y}(x(t),y(t))\, y'(t) \\ &= \frac{\partial f}{\partial x}(x(t),y(t))\, \frac{\partial f}{\partial x}(x_0,y_0) + \frac{\partial f}{\partial y}(x(t),y(t))\, \frac{\partial f}{\partial y}(x_0,y_0), \end{aligned}$$

et en particulier

$$g'(0) = \left(\frac{\partial f}{\partial x}(x_0,y_0)\right)^2 + \left(\frac{\partial f}{\partial y}(x_0,y_0)\right)^2 = \|\nabla f(x_0,y_0)\|^2.$$

Sachant que $\nabla f(x_0,y_0) \neq (0,0)$, on a donc $g'(0) > 0$. La fonction g étant de classe $\mathcal{C}^1$ au voisinage de 0, il existe $\eta > 0$ tel que $g'(t) > 0$ pour tout $t \in]-\eta,\eta[$, d'où le résultat. $\square$

Exemple – Soit $f : (x,y) \mapsto x^2 + y^2$. Les lignes de niveau de f sont les sous-ensembles de $\mathbb{R}^2$ d'équation $x^2 + y^2 = \lambda$ où $\lambda \in \mathbb{R}$. Si $\lambda < 0$, cet ensemble est vide, si $\lambda = 0$, il est réduit au point $(0,0)$, et si $\lambda > 0$, il s'agit du cercle de centre $(0,0)$ et de rayon $\sqrt{\lambda}$.

La fonction f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ et pour tout $(x,y) \in \mathbb{R}^2$,

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

il est donc colinéaire à (x,y) (ce qui est cohérent avec le deuxième exemple d'équation aux dérivées partielles que nous avons traité). Pour tout $(x_0,y_0) \neq (0,0)$, $\nabla f(x_0,y_0)$ est non nul et orthogonal à la ligne de niveau $\lambda = x_0^2 + y_0^2$ de f, orienté dans le sens des valeurs croissantes de f, c'est-à-dire, « s'éloignant » de l'origine.

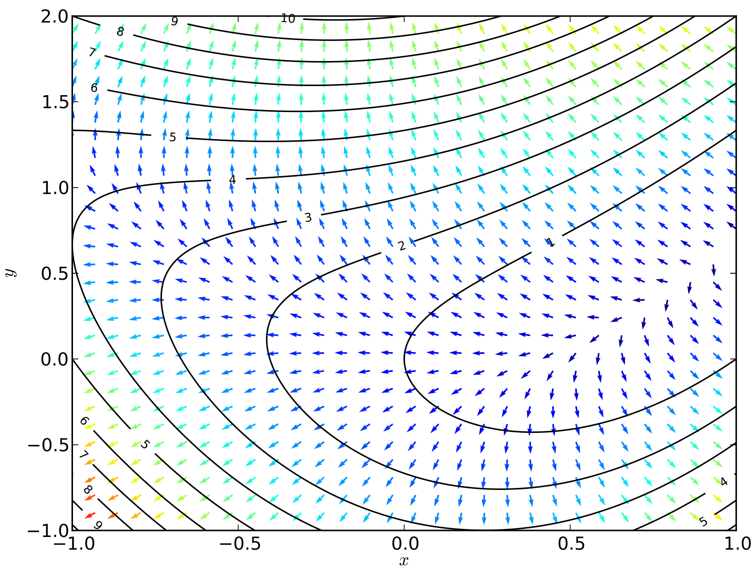
Remarque – En électrostatique par exemple :

• Si une fonction V représente un potentiel électrique V, les lignes de niveau de V sont appelées lignes équipotentielles.

• Si le champ électrostatique $\overrightarrow{E}$ dérive de V, c'est-à-dire vérifie $\overrightarrow{E} = -\nabla V$, on appelle ligne de champ de $\overrightarrow{E}$ toute courbe $\mathcal{C}$ régulière telle que pour tout $(x,y) \in \mathcal{C}$, $\overrightarrow{E}(x,y)$ soit un vecteur tangent à $\mathcal{C}$ en (x,y).

D'après ce qui précède, les lignes de champ de $\overrightarrow{E}$ sont orthogonales aux lignes équipotentielles de V. De plus, $\overrightarrow{E}$ est dirigé dans le sens des potentiels décroissants.

Exemple – Ci-dessous, on visualise certaines lignes de niveau de la fonction f représentée précédemment, ainsi que les valeurs de ∇f sur un maillage du plan, à partir desquelles on devine les lignes de champ de ∇f. On constate graphiquement sur cet exemple la propriété précédente :



2. Surfaces données par une équation cartésienne

Dans ce paragraphe, $p = 3$. On souhaite étudier les surfaces données par une équation cartésienne de la forme $f(x,y,z) = 0$, où $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^1$.

Exemples

• La sphère unité de $\mathbb{R}^3$ a pour équation cartésienne $x^2 + y^2 + z^2 = 1$.

• Si $g : V \to \mathbb{R}$ est une fonction de classe $\mathcal{C}^1$ sur un ouvert V de $\mathbb{R}^2$, la surface représentative de g a pour équation cartésienne $z = g(x,y)$, ce qui entre dans le cadre précédent, en posant $f(x,y,z) = g(x,y) - z$ pour tout $(x,y,z) \in \mathbb{R}^3$ tel que $(x,y) \in V$. Dans ce cas, il s'agit d'une représentation explicite car z est directement donné en fonction de x et y.

Définition – Point régulier, plan tangent

Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $\mathcal{S}$ la partie de $\mathbb{R}^3$ d'équation $f(x,y,z) = 0$.

• On appelle **point régulier** de $\mathcal{S}$ tout point $(x_0,y_0,z_0) \in \mathcal{S}$ tel que

$$\nabla f(x_0,y_0,z_0) \neq (0,0,0)$$

c'est-à-dire, tel que $\dfrac{\partial f}{\partial x}(x_0,y_0,z_0) \neq 0$ ou $\dfrac{\partial f}{\partial y}(x_0,y_0,z_0) \neq 0$ ou $\dfrac{\partial f}{\partial z}(x_0,y_0,z_0) \neq 0$.

• Si (x_0,y_0,z_0) est un point régulier de $\mathcal{S}$, on appelle **plan tangent** à $\mathcal{S}$ en (x_0,y_0,z_0) le plan orthogonal à $\nabla f(x_0,y_0,z_0)$ et passant par (x_0,y_0,z_0), c'est-à-dire, le plan de $\mathbb{R}^3$ d'équation

$$\begin{aligned} &\left(\nabla f(x_0,y_0,z_0) \,\middle|\, (x-x_0, y-y_0, z-z_0)\right) = 0, \\ \textit{i.e.} \quad &\frac{\partial f}{\partial x}(x_0,y_0,z_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0,z_0)(y-y_0) + \frac{\partial f}{\partial z}(x_0,y_0,z_0)(z-z_0) = 0. \end{aligned}$$

Exemple – Soit $g : V \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ sur un ouvert V de $\mathbb{R}^2$ et soit $\mathcal{S}$ la surface représentative de g, c'est-à-dire, la surface d'équation $z = g(x,y)$.

Comme on l'a expliqué ci-dessus, c'est un cas particulier de surface donnée par une équation cartésienne $f(x,y,z) = 0$ avec $f : (x,y,z) \mapsto g(x,y) - z$ définie sur l'ouvert de $\mathbb{R}^3$

$$U = \{(x,y,z) \in \mathbb{R}^3;\ (x,y) \in V\}.$$

La fonction f est de classe $\mathcal{C}^1$ sur U de même que g sur V, et pour tout $(x,y,z) \in U$,

$$\nabla f(x,y,z) = \left(\frac{\partial g}{\partial x}(x,y), \frac{\partial g}{\partial y}(x,y), -1\right) \neq (0,0,0).$$

En particulier, chaque point de $\mathcal{S}$ est régulier. Si $(x_0,y_0,z_0) \in \mathcal{S}$, le plan tangent à $\mathcal{S}$ en (x_0,y_0,z_0) a pour équation

$$\begin{aligned} &\frac{\partial g}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial g}{\partial y}(x_0,y_0)(y-y_0) - (z-z_0) = 0, \\ \textit{i.e.} \quad z = &\frac{\partial g}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial g}{\partial y}(x_0,y_0)(y-y_0) + g(x_0,y_0). \end{aligned}$$

3. Courbes tracées sur une surface

Définition

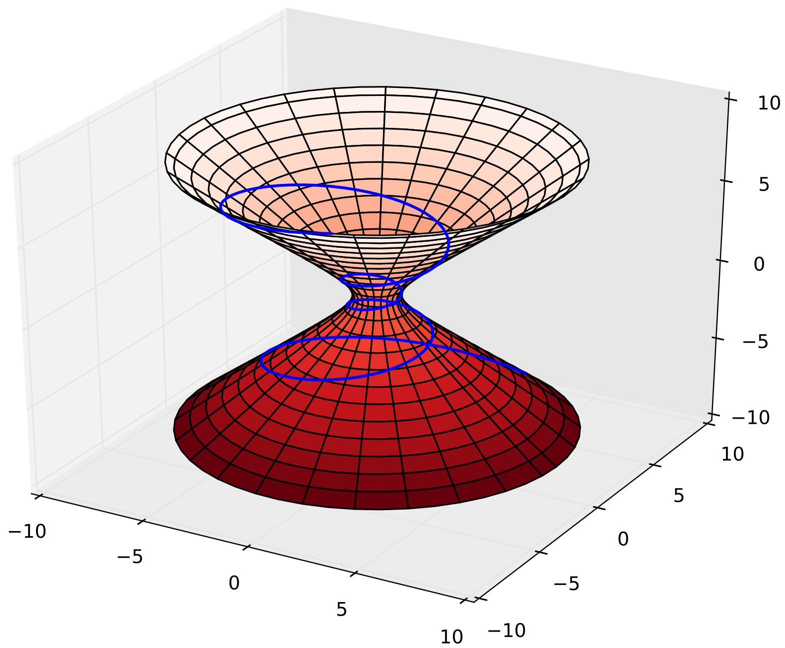
Soit $f : U \to \mathbb{R}$ de classe $\mathcal{C}^1$ et $\mathcal{S}$ la partie de $\mathbb{R}^3$ d'équation $f(x,y,z) = 0$.

On appelle **courbe tracée sur la surface** $\mathcal{S}$ tout arc paramétré (I,γ) où I est un intervalle de $\mathbb{R}$ et $\gamma = (x,y,z) : I \to \mathbb{R}^3$ vérifie, pour tout $t \in I$, $(x(t),y(t),z(t)) \in \mathcal{S}$.

Par exemple, ci-dessous, on visualise une partie de la courbe paramétrée par

$$\begin{cases} x(t) = \cos(t)\,\mathrm{ch}(t/4) \\ y(t) = \sin(t)\,\mathrm{ch}(t/4) \\ z(t) = \mathrm{sh}(t/4) \end{cases}$$

tracée sur la surface d'équation cartésienne $x^2 + y^2 - z^2 = 1$.



Par définition même, on a, avec les notations précédentes : pour tout $t \in I$,

$$f(x(t), y(t), z(t)) = 0.$$

Si γ est de classe $\mathcal{C}^1$, alors d'après la règle de la chaîne, $t \mapsto f(x(t), y(t), z(t))$ est de classe $\mathcal{C}^1$ sur I. Comme cette fonction est nulle, on a, pour tout $t \in I$,

$$\frac{\partial f}{\partial x}(x(t), y(t), z(t))\, x'(t) + \frac{\partial f}{\partial y}(x(t), y(t), z(t))\, y'(t) + \frac{\partial f}{\partial z}(x(t), y(t), z(t))\, z'(t) = 0$$

$$\textit{i.e.} \quad \big(\nabla f(\gamma(t)) \,|\, \gamma'(t)\big) = 0,$$

et donc $\nabla f(\gamma(t))$ est orthogonal à $\gamma'(t)$, qui dirige la tangente à la courbe en chacun de ses points réguliers.

On en déduit le résultat suivant :

Propriété – Tangente à une courbe tracée sur une surface

Soit $\Gamma = (I,\gamma)$ une courbe tracée sur la surface $\mathcal{S}$ d'équation $f(x, y, z) = 0$ où $f : U \to \mathbb{R}$ est de classe $\mathcal{C}^1$. On note $\mathcal{C}$ le support de Γ.

Soit $(x_0,y_0,z_0) = M(t_0) \in \mathcal{S}$ un point régulier en tant qu'élément de $\mathcal{S}$ et en tant que point de Γ.

Alors la tangente à Γ en $M(t_0)$ est contenue dans le plan tangent à $\mathcal{S}$ en (x_0,y_0,z_0).

Cas particulier – Soit $\mathcal{S}$ la surface représentative d'une fonction $g : V \to \mathbb{R}$ de classe $\mathcal{C}^1$, c'est-à-dire, la surface d'équation $z = g(x, y)$.

Fixons l'une des coordonnées x ou y, ce qui revient à considérer l'intersection de $\mathcal{S}$ avec des plans parallèles aux plans de coordonnées (yOz) ou (xOz). Par exemple, fixons $y = y_0$ et considérons le sous-ensemble

$$\{(x, y_0, g(x,y_0));\ (x,y_0) \in V\}.$$

C'est le support d'une courbe tracée sur $\mathcal{S}$, que l'on peut paramétrer par

$$x \mapsto (x, y_0, g(x,y_0));$$

elle est régulière. La situation est analogue si l'on fixe $x = x_0$. Les courbes de cette forme sont appelées **courbes coordonnées** de $\mathcal{S}$.

Enfin, si l'on fixe $z = z_0$, on obtient le sous-ensemble

$$\{(x,y,z_0) \in U;\ g(x, y) = z_0\},$$

qui s'identifie à la ligne de niveau z_0 de g. Si $\nabla g(x, y) \neq (0,0)$ pour tout $(x, y) \in V$ tel que $g(x, y) = z_0$, l'ensemble considéré est une courbe régulière et le théorème admis au paragraphe précédent permet de la paramétrer localement par des fonctions de la forme

$$t \mapsto (x(t), y(t), z_0);$$

à nouveau, on obtient une courbe tracée sur $\mathcal{S}$.

Remarque – Contrairement aux courbes coordonnées, qui sont toujours régulières, le sous-ensemble

$$\{(x,y,z_0) \in U;\ z_0 = g(x, y)\}$$

peut ne pas être une courbe régulière ; il peut même ne pas être une courbe, si par exemple g prend la valeur z_0 sur une partie de surface non nulle, comme un disque, une couronne... Pourtant, $\mathcal{S}$ est toujours une surface dont tous les points sont réguliers, elle possède bien un plan tangent en chacun de ses points !

Exemple – Soit $g : (x, y) \mapsto x^2 - y^2$. La fonction g est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$. À $y = y_0$ fixé, on obtient le sous-ensemble

$$\{(x, y_0, x^2 - y_0^2);\ x \in \mathbb{R}\},$$

qui est une parabole dont les branches sont tournées « vers le haut ». À $x = x_0$ fixé, on obtient le sous-ensemble

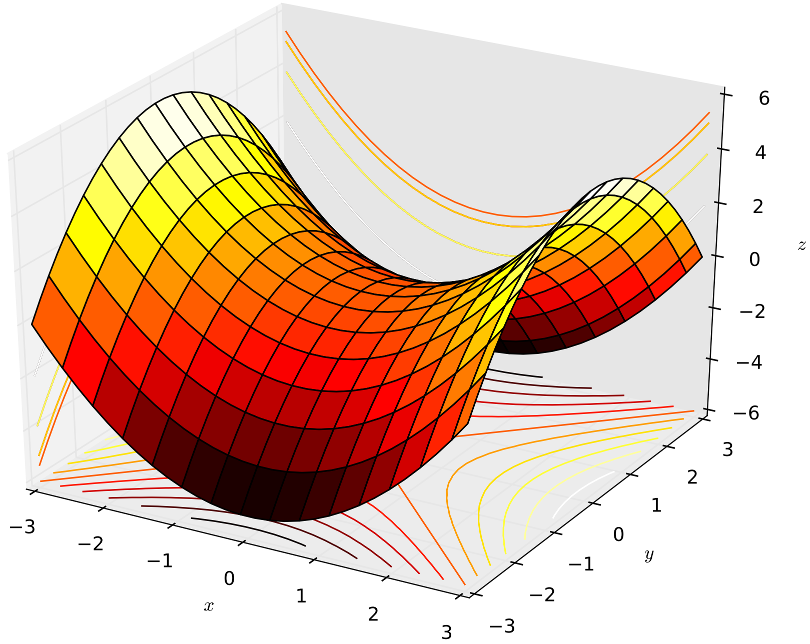
$$\{(x_0, y, x_0^2 - y^2);\ y \in \mathbb{R}\}$$

qui est une parabole dont les branches sont tournées « vers le bas ». À $z = z_0$ fixé, on obtient le sous-ensemble

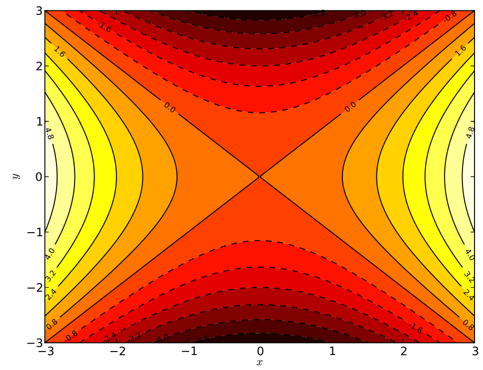
$$\{(x, y, z) \in \mathbb{R}^3;\ x^2 - y^2 = z_0\}.$$

Si $z_0 = 0$, il s'agit de la réunion de deux droites sécantes, sinon, il s'agit d'une hyperbole.

Voici la représentation graphique de la fonction g, sur laquelle on peut visualiser des courbes coordonnées de la surface (directement sur la surface), certaines de leurs projections sur des plans « verticaux », et certaines lignes de niveau de g :



Sur la figure suivante, on visualise plus particulièrement certaines lignes de niveaux de g :



On constate que toutes les lignes de niveau de g sont régulières sauf la ligne de niveau 0 (réunion des deux droites sécantes), dont le point $(0,0)$ n'est pas un point régulier.

Cela se démontre très facilement : la fonction g est de classe $\mathcal{C}^1$ sur $\mathbb{R}^2$ car polynomiale, et pour tout $(x,y) \in \mathbb{R}^2$,

$$\nabla g(x,y) = (2x, -2y);$$

ainsi ∇g s'annule uniquement en $(0,0)$, qui est un point de la ligne 0 de g.