

Chapitre 3 : bases du calcul différentiel

Léo Bernus

15 février 2020

Introduction

Le calcul différentiel est la base absolue de toutes les techniques de calculs utilisées en physique. La grande idée du calcul différentiel est de ramener l'étude du courbe à l'étude du droit, c'est-à-dire de ramener les variations d'objets complexes à l'étude d'applications linéaires au voisinage d'un point. La grande idée du calcul différentiel peut se ramener à la formule suivante. Si f est différentiable en a , alors il existe un ouvert centré sur 0 tel que pour tout vecteur \mathbf{h} dans cet ouvert, on a :

$$f(a + \mathbf{h}) = f(a) + df(a)\mathbf{h} + o(\|\mathbf{h}\|), \quad (1)$$

soit :

$$\text{variations au voisinage d'un point} = \text{algèbre linéaire} + \text{majoration de normes}. \quad (2)$$

Autrement dit, il va s'agir de montrer que sous certaines hypothèses de régularité qui seront précisées dans le cours, sur un voisinage suffisamment petit, toute fonction peut être approchée par une application linéaire, notée ici $df(a)$, sous réserve que l'on soit capable de montrer que le reste, noté $o(\|\mathbf{h}\|)$, soit suffisamment majoré en norme pour que l'approximation linéaire soit vraie.

Un point délicat invoqué ici est la notion de voisinage. Pour rappel :

Définition 0.1. Soit E un espace topologique et $a \in E$. On appelle *voisinage* de a et on note $\mathcal{V}(a)$, tout ensemble qui contient un ouvert qui contient a .

Cela implique que pour parler de différentiabilité, il faut qu'il existe un voisinage de a , donc au moins un ouvert contenant a , dans lequel l'équation (1) est valable. Nous verrons plus tard que c'est une hypothèse forte. Cependant, nous pouvons montrer qu'il est toujours possible de parler de variation au voisinage d'un point en y ajoutant un vecteur (ici noté \mathbf{h}). En effet, si $\mathcal{V}(a)$ est un voisinage de a , alors il existe r réel strictement positif tel que la boule ouverte de centre a et de rayon r notée $B(a, r)$ soit incluse dans $\mathcal{V}(a)$. Dans ce cas, pour tout vecteur non nul \mathbf{h} , il est facile de vérifier que $\forall t \in]-\frac{r}{\|\mathbf{h}\|}, \frac{r}{\|\mathbf{h}\|}[$, $a + t\mathbf{h} \in B(a, r) \subset \mathcal{V}(a)$. Nous avons représenté cette opération sur la figure 1.

Tout ce chapitre va consister à tourner autour de l'équation (1). La définition autour de laquelle va tourner tout le cours sera la suivante :

Définition 0.2. Soit une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n$, soit $a \in \Omega$. On dit que f est différentiable en a s'il existe une application linéaire L_a , \mathcal{V}_a un voisinage de a , tels que pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $a + \mathbf{h} \in \mathcal{V}_a$,

$$f(a + \mathbf{h}) = f(a) + L_a(\mathbf{h}) + o(\|\mathbf{h}\|) \quad (3)$$

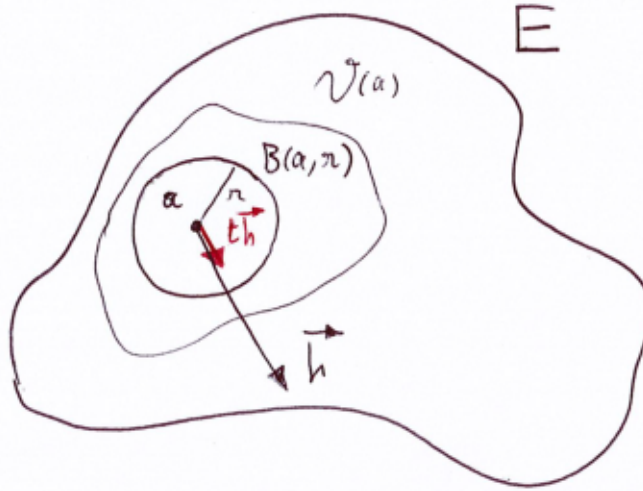


FIGURE 1 – Représentation d'un ouvert E contenant un point a , un voisinage de a noté $\mathcal{V}(a)$, d'une boule ouverte de rayon r centrée sur a notée $B(a, r)$, et comment faire rentrer $a + t\mathbf{v}$ dans $B(a, r)$ et donc dans $\mathcal{V}(a)$.

À titre de rappel :

Définition 0.3. Soit $f : \mathbb{R}^n \rightarrow \mathbb{R}^p$.

$$f(\mathbf{h}) = o(\|\mathbf{h}\|) \iff \exists \varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}^p / f(\mathbf{h}) = \|\mathbf{h}\|\varepsilon(\mathbf{h}), \quad \|\varepsilon(\mathbf{h})\| \xrightarrow{\|\mathbf{h} \rightarrow 0\|} 0. \quad (4)$$

Dans ce cas on dit que f est négligeable devant \mathbf{h} .

$$f(\mathbf{h}) = \mathcal{O}(\|\mathbf{h}\|) \iff \exists C > 0, \eta > 0 / \|\mathbf{h}\| < \eta \Rightarrow \|f(\mathbf{h})\| < C\|\mathbf{h}\| \quad (5)$$

Dans ce cas on dit que f est dominée par \mathbf{h} .

Si vous êtes perdus en calcul différentiel, rappelez-vous qu'il s'agit toujours de la même question fondamentale : comment se comporte cet objet (souvent une fonction) au voisinage d'un point ? Autrement dit : comment trouver une approximation linéaire de $f(a + \mathbf{h})$ du moment que la norme de \mathbf{h} est suffisamment petite ? Les questions plus subtiles et difficiles pour démontrer les théorèmes importants consisteront à déterminer la petitesse nécessaire de la norme de \mathbf{h} pour que les conditions de réalisation du théorème en question soient valides.

1 Fonctions de plusieurs variables à valeur dans \mathbb{R}

Dans cette section nous allons construire pas à pas la notion de dérivée partielle, sans nous soucier pour le moment de la différentiabilité des fonctions. Nous allons justement montrer quelques contre-exemples où les dérivées partielles peuvent être définies sans que la fonction ne soit différentiable.

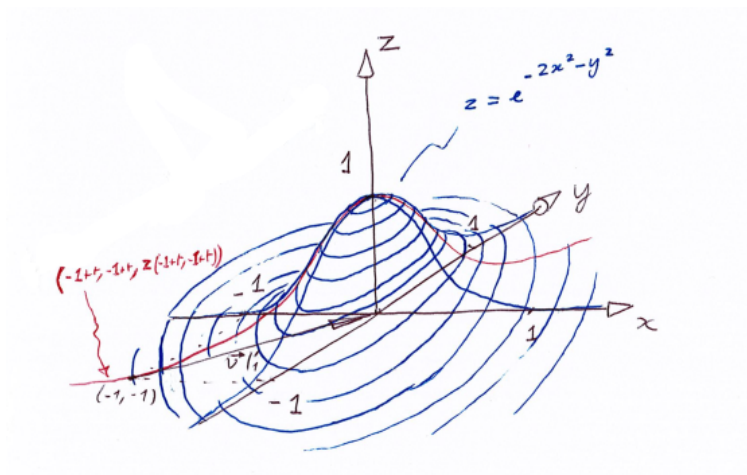


FIGURE 2 – Représentation de la nape d'équation $z = e^{-2x^2 - y^2}$ et d'une ligne passant par cette surface, d'équation paramétrique $t \mapsto (-1 + t, -1 + t, z((-1 + t, -1 + t)))$.

1.1 Dérivée directionnelle, ou dérivée le long d'un vecteur

Définition 1.1. Soit $f : E \rightarrow \mathbb{R}$, où $E \subset \mathbb{R}^n$ est un ouvert. Soit $a \in E$ et $\mathcal{V}(a)$ un voisinage de a . Soit $\mathbf{v} \in \mathbb{R}^n$. Soit I un intervalle ouvert de \mathbb{R} contenant 0 tel que $\forall t \in I, a + t\mathbf{v} \in \mathcal{V}(a)$.

On dit que f admet en a une dérivée le long du vecteur \mathbf{v} si la fonction d'une variable $\phi : I \rightarrow \mathbb{R}, t \mapsto f(a + t\mathbf{v})$, est dérivable en 0.

Si elle existe, on note

$$D_{\mathbf{v}}f(a) = \phi'(0) = \lim_{t \rightarrow 0} \frac{f(a + t\mathbf{v}) - f(a)}{t} \quad (6)$$

la dérivée de f le long de \mathbf{v} en a .

Remarque Si f , définie sur un ouvert $E \subset \mathbb{R}^n$, admet une dérivée le long de \mathbf{v} en $a \in E$, alors

$$\exists \mathcal{V}(a), \exists r > 0 / \forall t \in]-r, r[, a + t\mathbf{v} \in \mathcal{V}(a), f(a + t\mathbf{v}) = f(a) + t D_{\mathbf{v}}f(a) + o(t) \quad (7)$$

Cela veut dire qu'on peut écrire un développement limité de f le long de \mathbf{v} , mais cela n'implique pas nécessairement que f admette un développement limité sur un voisinage de a .

Exercices

1. Soit une montagne d'équation $z = e^{-2x^2 - y^2}$ (fig. 2). Un promeneur se situe en $(-1, -1, e^{-3})$ et marche à une vitesse dont la projection sur le plan (xOy) a pour coordonnées $(1, 1)$. Calculer la composante verticale de sa vitesse. On admet que la fonction $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto e^{-2x^2 - y^2}$ est différentiable en $(-1, -1)$ pour le moment.

Solution. Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto e^{-2x^2 - y^2}$. Dire que le promeneur se situe en $(-1, -1)$ et a une vitesse horizontale de composantes $(1, 1)$ revient à dire que, dans un voisinage

de $(-1, -1)$, si on note $(x(t), y(t))$ ses coordonnées horizontales à la date t et qu'on choisit par convention $(x(0), y(0)) = (-1, -1)$, on peut trouver un intervalle I ouvert contenant 0 et une fonction $\varepsilon : I \rightarrow \mathbb{R}^2$ qui vérifie $\|\varepsilon(t)\| \xrightarrow[t \rightarrow 0]{} 0$ telle que l'on ait

$$\forall t \in I, (x(t), y(t)) = (-1, -1) + t(1, 1) + t\varepsilon(t) \quad (8)$$

L'altitude du promeneur à la date t vaut

$$z(t) = f(x(t), y(t)) = f(-1+t+t\varepsilon_x(t), -1+t+t\varepsilon_y(t)) = \exp(-2(-1+t+t\varepsilon_x(t))^2 - (-1+t+t\varepsilon_y(t))^2) \quad (9)$$

Comme on a admis que f était différentiable en $(-1, -1)$, on peut admettre l'existence d'une application linéaire L_a telle que

$$f((-1, -1) + t(1, 1) + t\varepsilon(t)) = f(-1, -1) + L_a((t, t) + t\varepsilon(t)) + o((t, t) + t\varepsilon(t)) \quad (10)$$

mais comme L_a est linéaire et qu'en dimension finie toutes les applications linéaires sont continues, on est assuré que¹ $L_a(t\varepsilon(t)) = o(t)$. Par ailleurs il est évident que $o(t\varepsilon(t)) = o(t)$. Ainsi on peut écrire

$$z(t) = f(-1, -1) + L_a(t, t) + o(t). \quad (11)$$

Par unicité du développement limité, on est assuré que $L_a(t, t)$ sera bien égal à $D_{(1,1)}f(-1, -1)$. Ainsi il suffit de calculer $\phi'(0)$ où $\phi(t) = f(-1+t, -1+t)$. Par unicité du développement limité, on aura évidemment $z'(0) = \phi'(0)$. On calcule aisément que

$$\phi(t) = \exp(-2t^2 + 6t - 3) \quad \phi'(t) = (-4t + 6)\exp(-2t^2 + 6t - 3), \quad (12)$$

et donc $\phi'(0) = 6e^{-3} = D_{(1,1)}f(-1, -1) = z'(0)$.

Bien entendu il ne sera pas nécessaire d'être aussi immondément laborieux à chaque fois, cette proposition de correction était surtout pour montrer comment la majoration de norme permet d'identifier la fonction et son approximation linéaire lorsque l'on dérive.

2. Le but de cet exercice est de montrer que certaines fonctions peuvent n'être dérivables directionnellement presque nulle part.

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto |2x - y|$. Soit $\mathbf{w} = (u, v) \in \mathbb{R}^2$ avec $\|\mathbf{w}\| = 1$. À quelles conditions sur \mathbf{w} f est-elle dérivable le long de \mathbf{w} en $(0,0)$?

Solution. Soit $t \neq 0$. La fonction f est continue sur \mathbb{R}^2 et pour tout t réel non nul, on a

$$\frac{f(t\mathbf{w}) - f(0,0)}{t} = \frac{|2tu - tv|}{t} = \frac{|t|}{t}|2u - v| = \text{signe}(t)|2u - v| \quad (13)$$

Si $2u \neq v$, cette expression n'admet pas de limite en 0 car elles diffèrent selon que t est positif ou négatif. Si $2u = v$, alors la dérivée le long de \mathbf{w} existe et vaut 0. Comme $\|\mathbf{w}\|^2 = 1 = u^2 + v^2$, cela ne laisse que deux choix possibles : $u = 1/\sqrt{5}, v = 2/\sqrt{5}$, ou $u = -1/\sqrt{5}$ et $v = -2/\sqrt{5}$.

1. En effet, si on pose, pour toute application linéaire u sur un espace vectoriel H $\|u\| = \sup_{x \in H} \frac{\|u(x)\|}{\|x\|}$, un résultat classique d'algèbre montre que ce nombre existe et est fini pour toute application linéaire sur H . Dans ce cas, on a $\|L_a(t\varepsilon(t))\| \leq \|L_a\| \|t\| \|\varepsilon(t)\| = o(t)$.

Remarques

1. De cet exercice on voit qu'en un point, une fonction peut admettre des dérivées directionnelles selon certains vecteurs et non d'autres.
2. Une fonction peut, en un point, être dérivable le long de tout vecteur, sans même être continue! Cela illustre que la continuité est une hypothèse forte. Pour rappel, une fonction f définie sur un ouvert E est continue en a si et seulement si

$$\forall \varepsilon > 0, \exists r > 0 / \|b - a\| < r \Rightarrow b \in E, \|f(b) - f(a)\| < \varepsilon \quad (14)$$

La continuité en un point demande d'être proche de ce point en norme quelque soit le chemin par lequel on s'approche de ce point. Nous allons voir que les droites qui convergent vers un point ne suffisent pas.

Exercice Soit

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \\ (x, y) \mapsto \begin{cases} \frac{xy^2}{x^2+y^4} & \text{si } (x, y) \neq (0, 0) \\ 0 & \text{sinon.} \end{cases} \quad (15)$$

Montrer que f est dérivable en 0 le long de tout vecteur mais qu'elle n'est pas continue en 0.

Solution. Soit $\mathbf{v} = (u, v) \in \mathbb{R}^2$. On calcule

$$\forall t \in \mathbb{R}^*, \frac{f(t\mathbf{v}) - f(0)}{t} = \frac{uv^2}{u^2 + t^2v^4} \xrightarrow{t \rightarrow 0} \begin{cases} \frac{v^2}{u} & \text{si } u \neq 0 \in E \\ 0 & \text{si } u = 0 \end{cases} \quad (16)$$

Donc $D_{\mathbf{v}}f(0, 0)$ existe pour tout \mathbf{v} . D'autre part, on constate que d'une part

$$f(t, t) = \frac{t^3}{t^2 + t^4} \xrightarrow{t \rightarrow 0} 0 \quad (17)$$

et d'autre part

$$f(t^2, t) = \frac{1}{2} \xrightarrow{t \rightarrow 0} \frac{1}{2} \neq 0 \quad (18)$$

ce qui montre que l'on atteint une autre valeur si on approche $(0, 0)$ par une parabole et non une droite. La fonction f n'est donc pas continue. Elle n'est donc pas différentiable non plus, puisque la différentiabilité implique la continuité.

1.2 Dérivée partielle

Définition 1.2. Soit $f : \mathcal{U} \rightarrow \mathbb{R}$ où $\mathcal{U} \subset \mathbb{R}^n$ est un ouvert. Soit $(\mathbf{e}_i)_{i \in [1, n]}$ la base canonique de \mathbb{R}^n . Soit $a \in \mathcal{U}$. On appelle dérivée partielle de f en a selon la i -ième coordonnée, ou i -ième dérivée partielle de f en a , si elle est définie, le nombre $D_{\mathbf{e}_i}f(a)$, que l'on note (selon les différents ouvrages) :

$$D_{\mathbf{e}_i}f(a) = \frac{\partial f}{\partial x_i}(a) = \partial_i f(a) = D_i f(a) = \partial_{x_i} f(a) = \dots \quad (19)$$

Dans ce cours nous utiliserons principalement les notations $\partial f / \partial x_i$ lorsque la spécification des variables sera nécessaire, et $\partial_i f$ lorsqu'aucune ambiguïté ne sera possible.

Remarque On a donc, lorsqu'elle existe

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a_1, \dots, a_i + t, \dots, a_n) - f(a_1, \dots, a_n)}{t} \quad (20)$$

Gradient Supposons que $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ soit différentiable en a : d'après la définition 0.2, il existe un voisinage de a , une application linéaire L_a tel que f admette le développement limité suivant

$$f(a + \mathbf{h}) = f(a) + L_a \mathbf{h} + o(\|\mathbf{h}\|) \quad (21)$$

Soit $\mathcal{V}(a)$ un voisinage de a dans l'ensemble de définition de f tel que le développement limité de la différentielle soit valide. Pour tout vecteur \mathbf{v} , on se donne $r > 0$ tel que $\forall t \in]-r, r[, a + t\mathbf{v} \in \mathcal{V}(a)$. Soit $(\mathbf{e}_i)_i$ la base canonique de \mathbb{R}^n . Si $\mathbf{v} = \mathbf{e}_i$, alors

$$f(a + t\mathbf{e}_i) = f(a) + t \frac{\partial f}{\partial x_i}(a) + o(t) \quad (22)$$

Si $\mathbf{v} = \sum_i v_i \mathbf{e}_i$, alors

$$f(a + t\mathbf{v}) = f(a) + tL_a \mathbf{v} + o(t) = f(a) + tL_a \left(\sum_i v_i \mathbf{e}_i \right) + o(t) = f(a) + t \sum_i v_i L_a(\mathbf{e}_i) + o(t) \quad (23)$$

Mais on sait que $L_a(\mathbf{e}_i) = \partial_i f(a)$, d'où

$$f(a + t\mathbf{v}) = f(a) + t \sum_i \frac{\partial f}{\partial x_i}(a) v_i + o(t) \quad (24)$$

Donc ici, $df(a)$ apparaît comme une forme linéaire définie de la façon suivante

$$\begin{aligned} df(a) : \quad \mathbb{R}^n &\rightarrow \mathbb{R} \\ \mathbf{v} = \sum_i v_i \mathbf{e}_i &\mapsto \sum_i \frac{\partial f}{\partial x_i} v_i \end{aligned} \quad (25)$$

Comme à toute forme linéaire de \mathbb{R}^n correspond un unique vecteur, on peut appeler gradient de f en a l'unique vecteur noté $\nabla f(a)$ tel que $df(a) = \langle \nabla f(a), \cdot \rangle$, et les coordonnées de $\nabla f(a)$ sont égales aux dérivées partielles de f en a .

Remarques

1. Matriciellement, si f est à variables dans \mathbb{R} , dans la base canonique on a $\nabla f = {}^t df$. En effet

$$\forall \mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n, \quad df(a)\mathbf{v} = (\partial_1 f(a), \dots, \partial_n f(a)) \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = {}^t \nabla f(a) \mathbf{v} \quad (26)$$

2. Certains ouvrages notent **grad** f au lieu de ∇f pour éviter la tentation de faire des raccourcis hâtifs en notant $\nabla = {}^t(\partial_1, \dots, \partial_n)$, égalité fautive dans un système de coordonnées quelconque. Cette relation n'est vraie que dans la base canonique d'un espace euclidien².

2. Nous verrons ultérieurement qu'il en est autrement dans les variétés différentiables.

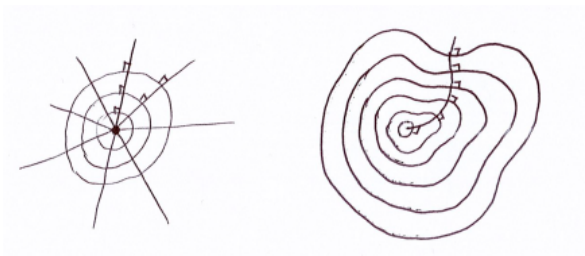


FIGURE 3 – Quelques exemples de lignes de niveaux de fonctions de deux variables à valeurs dans \mathbb{R} . Dans ce cas, les lignes de niveau sont également les surfaces de niveau. À gauche : les lignes de niveau du potentiel électrostatique engendré par une charge ponctuel sont des cercles concentriques, et les lignes de champ sont les droites qui concourent en la particule. À droite : lignes de niveau d’une fonction différentiable quelconque avec une ligne de champ.

Définition 1.3. Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, telle que $\forall x \in \Omega, \nabla f(x) \neq \mathbf{0}$.

- On appelle lignes de niveau de f l’ensemble des courbes régulières de \mathbb{R}^n telles que f ne varie pas le long de ces lignes.
- On appelle surfaces de niveau si $n = 3$ ou hypersurface de niveau si $n > 3$, l’ensemble des sous-variétés de \mathbb{R}^n de dimension $n - 1$ telles que f ne varie pas sur ces sous-variétés.
- On appelle lignes de champ l’ensemble des courbes qui sont tangentes à ∇f en tout point.

Nous représentons quelques exemples de lignes de niveau sur la figure 3. Un exemple connu d’utilisation des lignes de niveau sont les cartes de randonnées.

Exercices Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$ telle que $\forall x \in \Omega, \nabla f(x) \neq 0$.

1. Montrer qu’en tout point de Ω , le gradient de f est orthogonal aux lignes de niveau.
2. Montrer que le gradient de f correspond localement à la direction d’accroissement maximal de f .
3. On reprend la montagne d’équation $z = e^{-2x^2 - y^2}$ et son promeneur situé en $(-1, -1)$. Dans quelle direction devra aller le promeneur pour monter localement le plus vite possible ?

Solution.

1. Soit $a \in \Omega$, soit \mathcal{L} une ligne de niveau qui passe par a . Si \mathbf{v} est un vecteur directeur de \mathcal{L} en a , alors il existe un paramétrage de \mathcal{L} , c’est-à-dire un intervalle I ouvert connexe de \mathbb{R} , et une fonction $\gamma : I \rightarrow \Omega$ avec $\gamma(I) = \mathcal{L}$, on posera $\gamma(0) = a$, et il existe une fonction ϵ qui tend (en norme) vers 0 en 0 telle que

$$\gamma(t) = a + t\mathbf{v} + t\epsilon(t) \tag{27}$$

Sur la ligne de niveau, on a

$$f(a) = f(\gamma(t)) \tag{28}$$

$$= f(a + t\mathbf{v} + t\epsilon(t)) \tag{29}$$

$$= f(a) + t\langle \nabla f(a), t\mathbf{v} \rangle + t\langle \nabla f(a), \epsilon(t) \rangle + o(t) \tag{30}$$

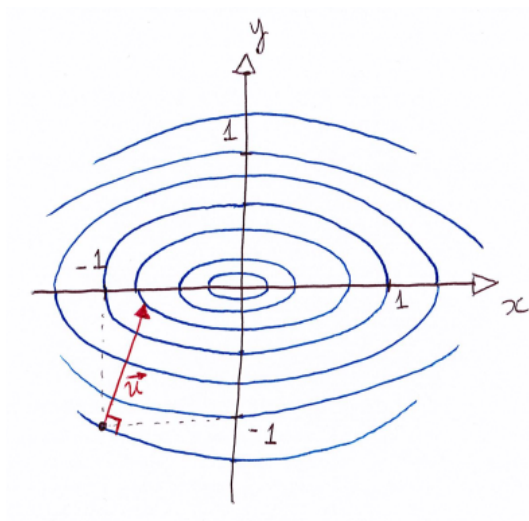


FIGURE 4 – Lignes de niveau de la montagne d'équation $z = e^{-2x^2 - y^2}$. On voit que le vecteur \mathbf{u} est bien orthogonal à la ligne de niveau.

Comme $\|\epsilon(t)\| \xrightarrow[t \rightarrow 0]{} 0$, $t\langle \nabla f(a), \epsilon(t) \rangle = o(t)$ d'où

$$t\langle \nabla f(a), \mathbf{v} \rangle + o(t) = 0 \quad (31)$$

Par unicité du développement limité, on en déduit $\langle \nabla f(a), \mathbf{v} \rangle = 0$, ce qui était à démontrer.

2. Soit une ligne passant par a et dirigée par un vecteur unitaire \mathbf{u} en a . Par le même raisonnement, on peut montrer que

$$f(\gamma(t)) = f(a) + t\langle \nabla f(a), \mathbf{u} \rangle + o(t) \quad (32)$$

Si $t > 0$, l'accroissement local de f dépendra du terme $\langle \nabla f(a), \mathbf{u} \rangle$. Ou bien \mathbf{u} est colinéaire à $\nabla f(a)$, auquel cas on a nécessairement $\mathbf{u} = \pm \nabla f(a) / \|\nabla f(a)\|$, et pour maximiser l'accroissement il faut choisir $+$. Dans ce cas, on a

$$\langle \nabla f(a), \mathbf{u} \rangle = \|\nabla f(a)\| \quad (33)$$

Ou bien \mathbf{u} n'est pas colinéaire à $\nabla f(a)$. Dans ce cas, on complète $\nabla f(a) / \|\nabla f(a)\|$ par $n-1$ vecteurs pour faire une base orthonormée de \mathbb{R}^n . Dans ce cas, on a

$$\mathbf{u} = u_{\nabla} \nabla f(a) / \|\nabla f(a)\| + \sum_{i=1}^{n-1} u_i \mathbf{e}_i \quad (34)$$

Comme $\|\mathbf{u}\| = 1$, nécessairement, $|u_{\nabla}| < 1$. Dans ce cas, on a

$$\langle \nabla f(a), \mathbf{u} \rangle = u_{\nabla} \|\nabla f(a)\| < \|\nabla f(a)\| \quad (35)$$

Cela prouve que $\nabla f(a)$ est la direction qui maximise l'accroissement de f .

3. La direction que doit suivre le promeneur est bien entendu $\mathbf{u} = \nabla z(-1, -1) / \|\nabla z(-1, -1)\| = (2, 1) / \sqrt{5}$. Pour descendre le plus vite possible il faut aller vers $-\mathbf{u}$. Nous illustrons ce résultat dans la figure 4.

2 Différentiabilité

Cette fois-ci on considère des fonctions à n variables à valeurs dans \mathbb{R}^p .

2.1 Différentielle

Dans cette section nous voulons généraliser et formaliser ce qui a été abordé de façon heuristique dans la première section. Rappelons la définition de la différentiabilité :

Définition 2.1. Soit une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n$, soit $a \in \Omega$. On dit que f est différentiable en a s'il existe une application linéaire L_a , \mathcal{V}_a un voisinage de a , tels que pour tout $\mathbf{h} \in \mathbb{R}^n$ tel que $a + \mathbf{h} \in \mathcal{V}_a$,

$$f(a + \mathbf{h}) = f(a) + L_a(\mathbf{h}) + o(\|\mathbf{h}\|) \quad (36)$$

Théorème 2.1. (Définition). Si f est différentiable en a , alors avec les mêmes notations, L_a est unique. On appelle différentielle de f en a l'application linéaire ainsi définie : $df(a) = L_a$.

Démonstration. Par exemple, supposer l'existence de L_1 et L_2 distinctes et calculer la limite $\lim_{t \rightarrow 0} \frac{f(a+tv) - f(a)}{t}$ dans les deux cas et conclure. \square

Théorème 2.2. Soit une fonction f définie sur un ouvert $\Omega \subset \mathbb{R}^n$, soit $a \in \Omega$. Si f est différentiable en a , alors pour tout $\mathbf{h} \in \mathbb{R}^n / a + \mathbf{h} \in \mathcal{V}_a$ où \mathcal{V}_a est un voisinage de a tel que le développement limité à l'ordre 1 de f soit valable, on a

$$f(a + \mathbf{h}) = f(a) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(a) h_i + o(\|\mathbf{h}\|) \quad (37)$$

où

$$\frac{\partial f}{\partial x_i}(a) = \lim_{t \rightarrow 0} \frac{f(a + t\mathbf{e}_i) - f(a)}{t} \quad (38)$$

est la dérivée partielle de f en a selon la i -ième coordonnée.

Démonstration. C'est la même preuve que le gradient ! Il suffit de décomposer \mathbf{h} sur la base canonique. \square

Attention, ici $\partial_i f(a) \in \mathbb{R}^p$ et non \mathbb{R} , car f est à valeur dans \mathbb{R}^p . Dans ces conditions, la différentielle de f en a est une application linéaire de \mathbb{R}^n dans \mathbb{R}^p et vérifie

$$df(a) : \mathbb{R}^n \rightarrow \mathbb{R}^p, \mathbf{h} = \sum_i h_i \mathbf{e}_i \mapsto \sum_i \frac{\partial f}{\partial x_i}(a) h_i \quad (39)$$

Représentation matricielle Si f est à valeur dans \mathbb{R}^p et différentiable en a on a

$$f(a + \mathbf{h}) = \begin{pmatrix} f_1(a_1 + h_1, \dots, a_n + h_n) \\ \vdots \\ f_p(a_1 + h_1, \dots, a_n + h_n) \end{pmatrix} = \begin{pmatrix} f_1(a_1, \dots, a_n) \\ \vdots \\ f_p(a_1, \dots, a_n) \end{pmatrix} + \begin{pmatrix} \sum_{i=1}^n \partial_i f_1(a) h_i \\ \vdots \\ \sum_{i=1}^n \partial_i f_p(a) h_i \end{pmatrix} + o(\mathbf{h}) \quad (40)$$

Le deuxième terme est une application linéaire qu'on peut représenter matriciellement. La matrice représentative de $df(a)$ est appelée la matrice jacobienne de f en a relative à un système de coordonnées (ici la base canonique) et s'exprime ainsi

$$J_f(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_j}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial f_i}{\partial x_1}(a) & \dots & \frac{\partial f_i}{\partial x_j}(a) & \dots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_j}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \quad (41)$$

En effet on vérifie que

$$f(a + \mathbf{h}) = \begin{pmatrix} f_1(a_1, \dots, a_n) \\ \vdots \\ f_p(a_1, \dots, a_n) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \dots & \frac{\partial f_1}{\partial x_j}(a) & \dots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial f_i}{\partial x_1}(a) & \dots & \frac{\partial f_i}{\partial x_j}(a) & \dots & \frac{\partial f_i}{\partial x_n}(a) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \frac{\partial f_p}{\partial x_1}(a) & \dots & \frac{\partial f_p}{\partial x_j}(a) & \dots & \frac{\partial f_p}{\partial x_n}(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_j \\ \vdots \\ h_n \end{pmatrix} + o(\mathbf{h}) \quad (42)$$

Ici le vecteur colonne \mathbf{h} a n lignes, et la matrice $J_f(a)$ a p lignes et n colonnes. Retenons que la matrice jacobienne a pour composantes à la i -ième ligne et à la j -ième colonne

$$J_f(a)_{[i,j]} = \frac{\partial f_i}{\partial x_j}(a) \quad (43)$$

ou pas... en cas d'oubli il suffit de développer $f(a+\mathbf{h})$ et de représenter matriciellement l'application linéaire obtenue pour retrouver immédiatement l'expression de la jacobienne.

Exemple Soit f la fonction définie ainsi

$$f :]-\pi, \pi[\times \mathbb{R} \rightarrow \mathbb{R}^3 \\ \begin{pmatrix} \theta \\ u \end{pmatrix} \mapsto \begin{pmatrix} \cos \theta \operatorname{chu} \\ \sin \theta \operatorname{chu} \\ \operatorname{shu} \end{pmatrix} \quad (44)$$

où ch et sh sont les fonctions hyperboliques. On admet que f est différentiable en tout point de $]-\pi, \pi[\times \mathbb{R}$. On calcule tout simplement

$$J_f(\theta, u) = \begin{pmatrix} -\sin \theta \operatorname{chu} & \cos \theta \operatorname{shu} \\ \cos \theta \operatorname{chu} & \sin \theta \operatorname{shu} \\ 0 & \operatorname{chu} \end{pmatrix} \quad (45)$$

En fait, comme on le verra plus tard, f est une immersion qui permet de représenter une surface plongée dans \mathbb{R}^3 . Si on impose $\theta = cst$, on voit que dans le plan vertical qui fait un angle θ avec l'axe Ox , on trace une branche d'hyperbole. Une fois cela fait, lorsque l'on fait varier θ , cela fait une

surface de révolution autour de l'axe Oz . Il suffit donc de faire tourner une branche d'hyperbole d'équation cartésienne $x^2 - z^2 = 1$ autour de l'axe Oz pour dessiner cette figure. De plus, les deux colonnes de la matrice jacobienne, $\partial_\theta f$ et $\partial_u f$, sont des vecteurs tangents à cette surface : ils indiquent comment un point varie localement lorsqu'on fait varier localement les coordonnées θ et u . Toute cette heuristique sera développée formellement et rigoureusement dans le chapitre sur les surfaces régulières. En attendant on peut s'en faire une approche intuitive avec la figure 5

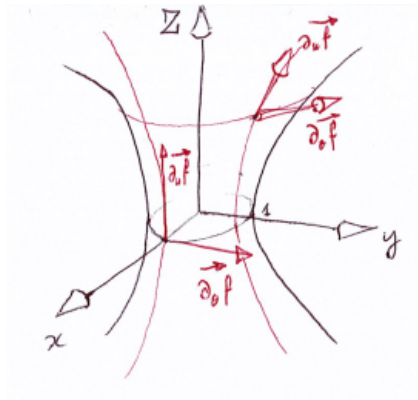


FIGURE 5 – On représente la surface plongée dont le paramétrage est décrit par l'équation (44). Nous avons représenté les vecteurs colonnes du jacobien (eq. (45)), $\partial_\theta f(\theta, u)$ et $\partial_u f(\theta, u)$ en deux points de la surface. Nous voyons qu'ils sont tangents aux lignes $u = cst$ et $\theta = cst$ respectivement.

2.2 Propriétés

- Si f est différentiable en a , alors f est continue en a . C'est immédiat car toutes les applications linéaires sont continues en dimension finie. En dimension infinie, il faudra ajouter à la définition de la différentiabilité que l'application linéaire en question soit continue.
- Corollaire : si f n'est pas continue en a , elle n'est pas différentiable en a .
- f peut admettre des dérivées partielles en a voire même des dérivées le long de tout vecteur en a , sans être différentiable ! On a déjà vu que cela était possible sans même que f ne soit continue avec l'exemple $f(x, y) = xy^2/(x^2 + y^4)$ si $(x, y) \neq 0$ et 0 sinon.
- Si f est différentiable en a , alors

$$\frac{f(a + \mathbf{h}) - f(a) - \mathrm{d}f(a)\mathbf{h}}{\|\mathbf{h}\|} \xrightarrow{\|\mathbf{h}\| \rightarrow 0} 0 \quad (46)$$

2.3 Condition suffisante

Nous allons (enfin) donner une condition suffisante pour vérifier que f est différentiable en a .

Théorème 2.3. *Soit $f : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$, $a \in \Omega$. S'il existe un voisinage ouvert de a tel que toutes les dérivées partielles de f existent et sont continues en a , alors f est différentiable en a .*

La preuve de ce théorème est technique mais pas insurmontable. Nous préférons la reporter en annexe de ce chapitre pour éviter d'alourdir la première lecture de l'exposé. L'idée de la preuve

consiste tout simplement à majorer le reste en cheminant d'abord de (a_1, a_2) à $(a_1 + h_1, a_2)$ puis $(a_1 + h_1, a_2 + h_2)$, en utilisant le théorème des accroissements finis pour exprimer le reste à chaque fois. L'idée est ensuite de majorer la norme de la différence avec la différentielle pour conclure que le développement limité peut s'écrire dans un voisinage de a . Nous recommandons aux lecteurs motivés d'essayer de faire et/ou de lire la preuve en deuxième lecture de ce cours s'ils découvrent le calcul différentiel en lisant ce texte, ou de le faire d'emblée s'ils ont déjà suivi des cours.

Définition 2.2. On appelle fonction de classe \mathcal{C}^1 sur Ω l'ensemble des fonctions différentiables en tout point de Ω . On note $\mathcal{C}^1(\Omega)$ cet ensemble.

On a $\mathcal{C}^1(\Omega) \subset \mathcal{C}^0(\Omega)$.

Exemples

- $(x, y, z) \mapsto (xyz - 42x^{42}, z^{2020})$ est \mathcal{C}^1 sur \mathbb{R}^3 .
- $(x, y) \mapsto \sqrt{x^2 + y^2}$ est \mathcal{C}^0 sur \mathbb{R}^2 et \mathcal{C}^1 sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.
- $(x, y) \mapsto 1/\sqrt{x^2 + y^2}$ est \mathcal{C}^0 et \mathcal{C}^1 sur $\mathbb{R}^2 \setminus \{(0, 0)\}$.

Une fois qu'on sait que la fonction est différentiable sur tout un ensemble on peut essayer d'en dégager des propriétés. En voici une.

Théorème 2.4. Soit $f : \mathcal{U} \subset \mathbb{R}^2 \rightarrow \mathbb{R}^p$. Si \mathcal{U} est un ouvert convexe alors les deux propositions suivantes sont équivalentes

1. $\forall (x, y) \in \mathcal{U}, \partial_x f(x, y) = 0$
2. Il existe une fonction d'une variable g, \mathcal{C}^1 , telle que $\forall (x, y) \in \mathcal{U}, f(x, y) = g(y)$

Démonstration. $2 \Rightarrow 1$ est évident, montrons $1 \Rightarrow 2$. Supposons les hypothèses vérifiées. Soient (x, y) et (x', y) dans \mathcal{U} . Pour tout t dans $[x, x']$, $(t, y) \in \mathcal{U}$ grâce à sa convexité. D'autre part on a

$$f(x', y) = f(x, y) + \int_x^{x'} \partial_t f(t, y) dt = f(x, y) \quad (47)$$

ce qui montre que la fonction dont le graphe est $g(x) = f(x, y)$ est bien une fonction d'une variable qui satisfait ce qui est à démontrer. \square

Ce genre de théorème est très utile en physique, d'ailleurs les physiciens ne se privent pas d'écrire des choses du genre " $\partial_x f(x, y) = 0 \Rightarrow f(y)$ " qui est faux d'un point de vue strictement mathématique (car les deux objets désignés par f ont une nature mathématique différente) mais correct du point de vue de la physique si f désigne une quantité et non une fonction mathématique.

La convexité est une hypothèse très importante. Voici un contre-exemple :

$$f : \mathbb{R}^2 \setminus \{x, y | y = 0, x \geq 0\} \rightarrow \mathbb{R}$$

$$(x, y) \mapsto \begin{cases} x^2 & \text{si } x \geq 0, y > 0 \\ -x^2 & \text{si } x \geq 0, y < 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (48)$$

On peut vérifier que f est \mathcal{C}^1 sur son ensemble de définition et que partout, on a $\partial_y f = 0$. Pourtant, on a $f(1, 1) = 1$ et $f(1, -1) = -1$, on ne peut donc pas trouver une fonction ne dépendant que de x qui suffirait à décrire le comportement de f . Ici, c'est la non convexité de Ω qui nous empêche de passer d'un point à l'autre par une droite. Nous avons représenté ce contre-exemple sur la figure 6. Nous montrerons comment ce théorème peut servir à résoudre des équations aux dérivées partielles avec l'exemple de l'équation d'onde (voir sec. 3.5).

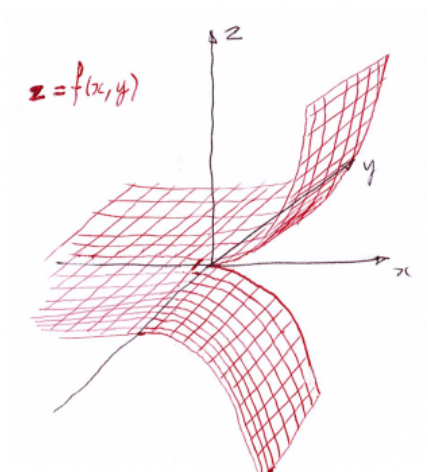


FIGURE 6 – Représentation du graphe de la fonction définie à l'équation (48). On voit clairement que toutes les dérivées le long de la direction y sont nulles mais que ça n'empêche aucunement la fonction de dépendre de y à cause de la non convexité de l'ensemble de définition.

3 Composition et opération sur les différentielles

3.1 Opérations élémentaires

Proposition 3.1. Soient $\alpha : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$, f et $g : \Omega \rightarrow \mathbb{R}^p$, avec Ω ouvert. Soit $a \in \Omega$ tel que α, f, g soient différentiables en a . Alors $\alpha f + g$ est différentiable en a et on a

$$d(\alpha f + g)(a) = f(a)d\alpha(a) + \alpha(a)df(a) + dg(a) \quad (49)$$

Matriciellement, on a

$$f(a) {}^t \nabla \alpha(a) + \alpha(a) J_f(a) + J_g(a) = \begin{pmatrix} f_1(a) \\ \vdots \\ f_p(a) \end{pmatrix} (\partial_1 \alpha(a), \dots, \partial_n \alpha(a)) + \alpha(a) \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ & \ddots & \\ \vdots & \partial_j f_i(a) & \vdots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix} + \begin{pmatrix} \partial_1 g_1(a) & \dots & \partial_n g_1(a) \\ & \ddots & \\ \vdots & \partial_j g_i(a) & \vdots \\ \partial_1 g_p(a) & \dots & \partial_n g_p(a) \end{pmatrix} \quad (50)$$

Le premier terme de cette équation est une matrice de taille $p \times n$ dont le coefficient i, j vaut

$f_i(a)\partial_j\alpha(a)$. Appliqué à un vecteur h , on a

$$\begin{pmatrix} f_1(a) \\ \vdots \\ f_p(a) \end{pmatrix} (\partial_1\alpha(a), \dots, \partial_n\alpha(a)) \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^n f_1(a)\partial_i\alpha(a)h_i \\ \vdots \\ \sum_{i=1}^n f_p(a)\partial_i\alpha(a)h_i \end{pmatrix} \quad (51)$$

Démonstration. Comme α, f, g sont différentiables en a , on peut trouver un voisinage de a sur lequel les trois fonctions admettent un développement limité à l'ordre 1 (il suffit de prendre l'intersection des trois voisinages). Soit \mathcal{V}_a un tel voisinage, soit $h \in \mathbb{R}^n$ tel que $a + h \in \mathcal{V}_a$. On a

$$\alpha(a+h)f(ah) + g(a+h) = (\alpha(a) + d\alpha(a)h + o(h))(f(a) + df(a)h + o(h)) + g(a) + dg(a)h + o(h) \quad (52)$$

$$\begin{aligned} &= \alpha(a)f(a) + g(a) + (d\alpha(a)h)f(a) + \alpha(a)df(a)h + dg(a)h \\ &\quad + f(a)o(h) + df(a)h o(h) + \alpha(a)o(h) + d\alpha(a)h o(h) + o(h) \end{aligned} \quad (53)$$

Comme les normes de $f(a)$ et $\alpha(a)$ sont finies, $f(a)o(h) + \alpha(a)o(h) = o(h)$. Comme les applications linéaires en dimension finie ont des normes finies on a $\|df(a)h\| \leq \|df(a)\| \|h\|$ et $|d\alpha(a)h| \leq \|\nabla\alpha(a)\| \|h\|$, donc $d\alpha(a)ho(h) = o(h)$ et $df(a)ho(h) = o(h)$. De telle sorte que

$$\alpha(a+h)f(ah) + g(a+h) = \alpha(a)f(a) + g(a) + f(a)d\alpha(a)h + \alpha(a)df(a)h + dg(a)h + o(h) \quad (54)$$

$h \mapsto f(a)(d\alpha(a)h) + \alpha(a)df(a)h + dg(a)h$ est une application linéaire, donc $\alpha f + g$ est différentiable en a et on a l'égalité demandée. \square

3.2 Composition des différentielles

Théorème 3.1. Soient $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathcal{V} \subset \mathbb{R}^p$, $g : \mathcal{V} \rightarrow \mathbb{R}^q$ avec \mathcal{U}, \mathcal{V} , ouverts. Soit $a \in \mathcal{U}$. Si f est différentiable en a et g est différentiable en $f(a)$, alors $g \circ f$ est différentiable en a et on a

$$d(g \circ f)(a) = dg(f(a)) \circ df(a) \quad (55)$$

Démonstration. g est différentiable en $f(a)$ donc il existe une boule ouverte de rayon ϵ dans laquelle pour tout vecteur ℓ tel que $f(a) + \ell$ est dans ce voisinage,

$$g(f(a) + \ell) = g(f(a)) + dg(f(a))\ell + o(\ell) \quad (56)$$

D'autre part, comme f est différentiable en a on a l'existence d'un voisinage ouvert de a tel que tant que $a + h$ est dans ce voisinage,

$$f(a+h) = f(a) + df(a)h + o(h) \quad (57)$$

Comme $df(a)$ est continue, on sait que $df(a)h + o(h) \xrightarrow{\|h\| \rightarrow 0} 0$ donc il existe une boule ouverte de rayon r telle que tant que $\|h\| < r$, $\|df(a)h + o(h)\| < \epsilon$. Il existe donc un voisinage de a tel que tant que $a + h$ est dans ce voisinage,

$$g(f(a+h)) = g(f(a) + df(a)h + o(h)) = g(f(a)) + dg(f(a))(df(a)h + o(h)) + o(df(a)h + o(h)) \quad (58)$$

Comme $\forall h \in \mathbb{R}^n \setminus \{\mathbf{0}\}$, $\|df(a)h\|/\|h\| \leq \|df(a)\|$, on a

$$o(\|df(a)h + o(h)\|) = o\left(\|h\| \frac{\|df(a)h\|}{\|h\|} + o(h)\right) = o(\|h\| \|df(a)\| + o(h)) = o(h) \quad (59)$$

De plus,

$$dg(f(a))o(h) \leq \|dg(f(a))\|o(h) = o(h) \quad (60)$$

D'où l'égalité demandée. \square

Les deux démonstrations des deux théorèmes précédents illustrent bien ce qu'annonçait l'introduction de ce cours. Le calcul différentiel consiste à majorer suffisamment des normes pour que l'algèbre linéaire puisse s'appliquer localement.

Matrices La représentation matricielle de $dg(f(a)) \circ df(a)$ est tout simplement le produit des jacobiniennes. On a donc

$$J_g(f(a))J_f(a)h = \begin{pmatrix} \partial_1 g_1(f(a)) & \dots & \partial_p g_1(f(a)) \\ & \ddots & \\ \vdots & \partial_j g_i(f(a)) & \vdots \\ & & \ddots \\ \partial_1 g_q(f(a)) & \dots & \partial_p g_q(f(a)) \end{pmatrix} \begin{pmatrix} \partial_1 f_1(a) & \dots & \partial_n f_1(a) \\ & \ddots & \\ \vdots & \partial_k f_\ell(a) & \vdots \\ & & \ddots \\ \partial_1 f_p(a) & \dots & \partial_n f_p(a) \end{pmatrix} \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} \quad (61)$$

La première matrice a q lignes et p colonnes, la deuxième matrice a p lignes et n colonnes, et la dernière a évidemment n lignes. Le produit de ces deux matrices est une matrice à q lignes et n colonnes et a pour composante i, j

$$(J_g(f(a))J_f(a))_{[i,j]} = \sum_{k=1}^p \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a) \quad (62)$$

Appliqué au vecteur h , le résultat d'une telle opération est un vecteur de p colonnes donc la i -ième ligne vaut

$$(J_g(f(a))J_f(a)\mathbf{h})_i = \sum_{j=1}^n \sum_{k=1}^p \frac{\partial g_i}{\partial y_k}(f(a)) \frac{\partial f_k}{\partial x_j}(a) h_j \quad (63)$$

Ici nous avons posé $\partial g_i/\partial y_j = \partial_j g_i$ avec la lettre y et non x pour éviter la confusion entre les deux espaces vectoriels distincts sur lesquels portent la dérivation. La différentielle de f en a agit sur un vecteur de \mathbb{R}^n et le transporte linéairement dans \mathbb{R}^p , alors que la différentielle de g en $b = f(a)$ agit sur un vecteur de \mathbb{R}^p et le transporte linéairement dans \mathbb{R}^q . Voir la figure 7.

3.3 Difféomorphisme et changement de variables

Il arrive souvent aux physiciens de noter abusivement $y = y(x)$ puis $f(x) = f(y)$ puis $df/dx = df/dy \times dy/dx$ pour noter un changement de variables. Par exemple, considérons la trajectoire d'une pierre en chute libre dans le champ de pesanteur au voisinage de la surface de la Terre lancée à 45 degrés vers le ciel. Il est possible de paramétrer l'altitude z de la pierre avec le temps, ou

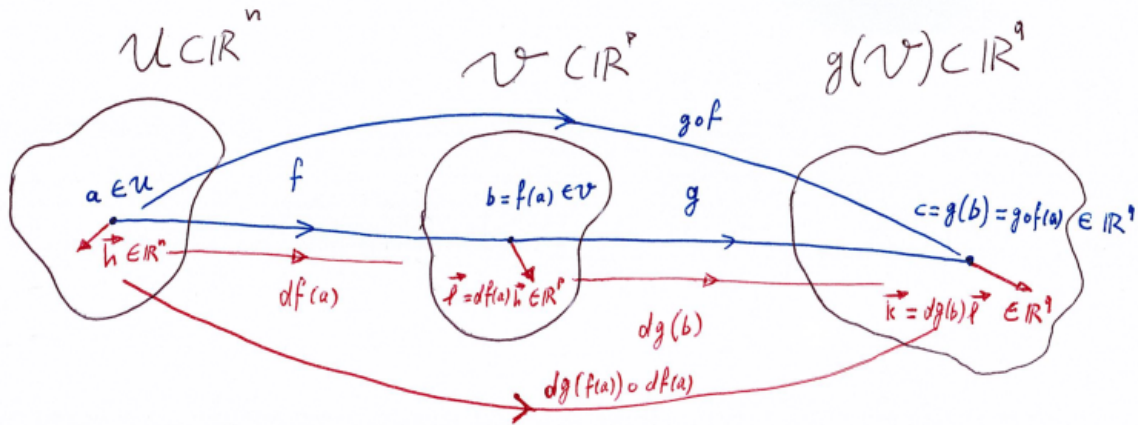


FIGURE 7 – Représentation symbolique de la composition des différentielles. En bleu : composition de f et g . En rouge : composition de $df(a)$ et $dg(f(a))$

avec l'espace. Les physiciens noteront $z = z(t) = z(x)$. Pourtant, la fonction $t \mapsto z$ est différente de la fonction $x \mapsto z$ car la dépendance en x et en t est différente, l'équation $z(x) = z(t)$ est donc nécessairement fautive. Il serait plus rigoureux de noter \tilde{z} l'altitude exprimée comme fonction du temps et z l'altitude exprimée comme fonction de l'espace, et relier les deux ainsi : $\tilde{z}(t) = z(x(t))$ par exemple. Cependant, les physiciens s'intéressent plus aux quantités mesurables qu'aux mathématiques (ce qui est parfaitement leur droit) et à ce titre ils ont raison de nommer les mêmes quantités par les mêmes lettres. L'inconvénient de cette approche est qu'elle peut semer la confusion en mathématiques. Il est plus prudent d'avoir à l'esprit les subtilités mathématiques que nous allons voir lorsque l'on fait de la physique, même si on utilise les mêmes lettres pour les mêmes quantités mais qui font références à des fonctions différentes.

Le changement de variable consiste à utiliser un point de vue différent plus simple mais équivalent au précédent. Il faut donc considérer une transformation d'un jeu de variables défini dans un ouvert, à un autre jeu de variables défini dans ce même ouvert. La définition la plus adaptée au changement de variables est la suivante.

3.3.1 Règle de la chaîne

Définition 3.1. Soient \mathcal{U}, \mathcal{V} , deux ouverts de \mathbb{R}^n . On dit que f est un \mathcal{C}^1 difféomorphisme, ou difféomorphisme de classe \mathcal{C}^1 , de \mathcal{U} dans \mathcal{V} , si f est une bijection de \mathcal{U} dans \mathcal{V} et que f et f^{-1} sont \mathcal{C}^1 respectivement sur \mathcal{U} et \mathcal{V} .

Proposition 3.2. Si f est un \mathcal{C}^1 difféomorphisme de \mathcal{U} dans \mathcal{V} , alors

$$\forall a \in \mathcal{U}, df^{-1}(a) = (df(a))^{-1} \quad (64)$$

Démonstration. $f^{-1} \circ f(a) = a = Id(a)$ donc $d(f^{-1} \circ f)(a) = Id = df^{-1}(f(a)) \circ df(a)$ d'où le résultat. \square

L'idée du changement de variables est que si l'on dispose d'un difféomorphisme dont le graphe est $y = \phi(x)$, on peut utiliser indifféremment y ou x comme jeu de variables, y compris s'il y a besoin de dériver les fonctions qui en dépendent. Si l'on note une fonction f de \mathcal{U} dans \mathbb{R}^p , on peut très bien poser la fonction g définie de \mathcal{V} dans \mathbb{R}^p telle que $g(\phi(x)) = g(y) = f(x)$. Pour réconcilier les physiciens et les mathématiciens, on peut noter que la "quantité" z peut s'exprimer ou bien comme une fonction de x ou bien comme une fonction de y : $z = f(x) = g(y)$. Une question très importante en physique est de savoir comment varie la quantité z lorsque x varie, ou lorsque y varie. Lorsqu'on considère z comme une fonction de x , il suffit d'écrire

$$\frac{\partial z}{\partial x_i} = \frac{\partial f}{\partial x_i} \quad (65)$$

et de même en substituant x_i par y_i et f par g . Comment obtenir l'une des expressions en fonction de l'autre ? Il suffit pour cela d'utiliser les résultats sur la composition des différentielles. On peut calculer très facilement que

$$\frac{\partial z}{\partial x_i} = \frac{\partial f}{\partial x_i} = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial g}{\partial y_k} \quad (66)$$

ou, noté autrement

$$\frac{\partial z}{\partial x_i} = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial z}{\partial y_k} \quad (67)$$

Matriciellement, on a multiplié la matrice jacobienne de g par la matrice jacobienne de ϕ . Comme cette équation est valable pour toute quantité z , on peut exprimer cette équation dans l'algèbre des opérateurs et écrire simplement

$$\frac{\partial}{\partial x_i} = \sum_{k=1}^n \frac{\partial y_k}{\partial x_i} \frac{\partial}{\partial y_k} \quad (68)$$

Cette équation est ce qu'on appelle la "règle de la chaîne". On peut évidemment l'écrire dans l'autre sens

$$\frac{\partial}{\partial y_i} = \sum_{k=1}^n \frac{\partial x_k}{\partial y_i} \frac{\partial}{\partial x_k} \quad (69)$$

et là il s'agit de multiplier la jacobienne de la fonction potentiellement considérée par la jacobienne de ϕ^{-1} .

3.3.2 Exemple simple

Soit $f : \mathbb{R}^2 \rightarrow \mathbb{R}^n, \mathcal{C}^1$. On pose $u = x + y$ et $v = 2x - y$. On pose $z = f(x, y)$. Exprimer $\partial z / \partial x$, puis $\partial z / \partial y$ en fonction de $\partial z / \partial u$ et $\partial z / \partial v$.

Solution : d'abord, on remarque que $\phi : (x, y) \mapsto (x + y, 2x - y)$ est bien un \mathcal{C}^1 difféomorphisme : ϕ est différentiable partout et on peut calculer son inverse, $\phi^{-1}(u, v) = ((u + v)/3, (2u - v)/3)$ qui est aussi différentiable partout. Pour le reste il suffit d'utiliser la règle de la chaîne

$$\frac{\partial}{\partial x} = \frac{\partial u}{\partial x} \frac{\partial}{\partial u} + \frac{\partial v}{\partial x} \frac{\partial}{\partial v} = \frac{\partial}{\partial u} + 2 \frac{\partial}{\partial v} \quad (70)$$

de même, on obtient

$$\frac{\partial}{\partial y} = \frac{\partial}{\partial u} - \frac{\partial}{\partial v} \quad (71)$$

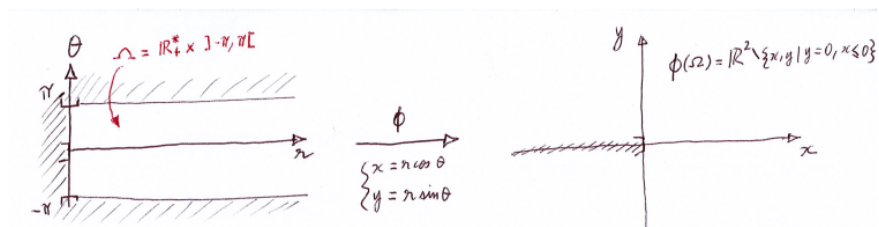


FIGURE 8 – Représentation du changement de variables en coordonnées polaires.

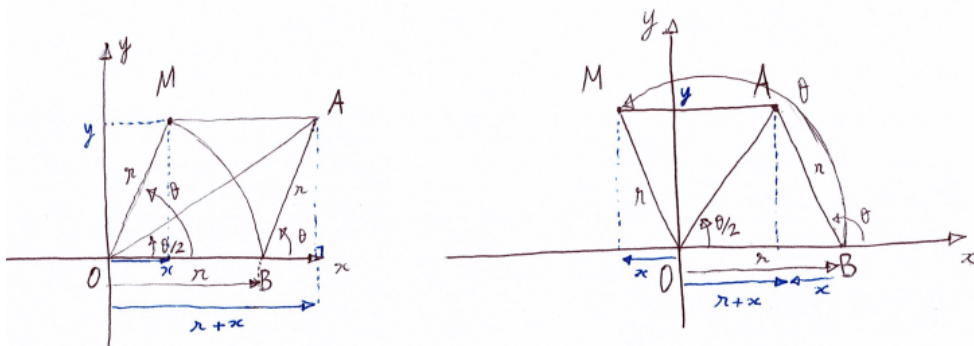


FIGURE 9 – À gauche on a représenté un cas où $x > 0$, à droite un cas où $x < 0$.

3.3.3 Coordonnées polaires

On pose $x = r \cos \theta$ et $y = r \sin \theta$ et $\phi : (r, \theta) \mapsto (x, y)$. Déterminer les ensembles de départ et d'arrivée tels que ϕ soit un \mathcal{C}^1 difféomorphisme. Écrire la bijection réciproque. Exprimer les dérivées partielles d'un jeu de variables en fonction de l'autre.

Solution. D'abord on constate que ϕ est différentiable sur \mathbb{R}^2 . Son jacobien vaut $\det J_\phi = r$ et s'annule en $r = 0$. Pour que ϕ soit inversible il faut donc enlever la droite $r = 0$ de \mathbb{R}^2 . Ensuite il faut que ϕ soit une bijection. Pour l'instant nous admettons le théorème d'inversion globale 4.2 qui affirme que si une fonction définie sur un ouvert est différentiable et de jacobien non nul, et injective, sur un ouvert, alors c'est un \mathcal{C}^1 difféomorphisme. Il faut donc trouver l'ensemble de départ tel que ϕ soit injective. On constate que $\phi(r, \theta + 2\pi) = \phi(r, \theta)$. On décide donc de garder un intervalle de largeur 2π pour θ , par exemple $]-\pi, \pi[$. Il reste encore un problème, car on a $\phi(-r, \theta + \pi) = \phi(r, \theta)$. On décide donc de ne garder que les r strictement positifs. L'ensemble de départ sera donc le suivant : $\Omega = \mathbb{R}_+^* \times]-\pi, \pi[$. L'ensemble d'arrivée sera $\mathbb{R}^2 \setminus \{x, y \mid y = 0, x \leq 0\}$ qui est la demi-droite $(-xO)$. Voir une représentation figure 8.

Écrivons la bijection réciproque. Tout d'abord on a évidemment $r = \sqrt{x^2 + y^2}$. On voit d'ailleurs ici la non inversibilité de $d\phi$ en $r = 0$ puisque r n'est pas différentiable en $x = y = 0$.

On serait tenté d'écrire $\theta = \text{atan}(y/x)$ mais cette formule présente un inconvénient : elle n'est pas valable quand $x = 0$. Une façon de procéder est de considérer le losange de la figure 9. Nous y considérons le point $M(x, y)$. Le point B est obtenu en reportant la longueur r sur l'axe (Ox) .

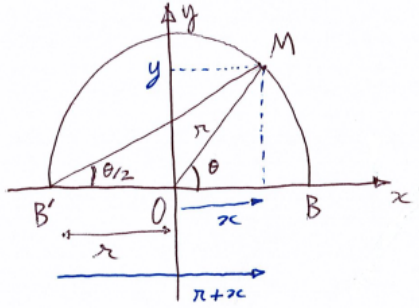


FIGURE 10 – Cercle polaire.

Le point A est obtenu de telle sorte que $\overrightarrow{OM} = \overrightarrow{BA}$. Dans ce cas, le quadrilatère $MOBA$ est un losange. De là, il suit que $(\overrightarrow{OB}, \overrightarrow{OA}) = (\overrightarrow{OB}, \overrightarrow{OM})/2 = \theta/2$. D'autre part, on voit sur la figure 9 que

$$\tan(\overrightarrow{OB}, \overrightarrow{OA}) = \frac{\theta}{2} = \frac{y}{r+x} = \frac{y}{\sqrt{x^2+y^2}+x} \quad (72)$$

d'où

$$\theta = 2 \operatorname{atan} \frac{y}{\sqrt{x^2+y^2}+x} \quad (73)$$

Cette équation est valable sur tout $\phi(\Omega)$. Une autre façon de procéder est la suivante. On constate sur la figure 10 que l'angle θ est un angle au centre et que $(\overrightarrow{B'B}, \overrightarrow{B'M})$ est un angle inscrit qui coupe la même corde de cercle que θ , d'où $(\overrightarrow{B'B}, \overrightarrow{B'M}) = \theta/2$. Par un raisonnement géométrique simple on en déduit la même formule.

À présent calculons les dérivées partielles. On a

$$\partial_r = \frac{\partial x}{\partial r} \partial_x + \frac{\partial y}{\partial r} \partial_y = \cos \theta \partial_x + \sin \theta \partial_y = \frac{x}{\sqrt{x^2+y^2}} \partial_x + \frac{y}{\sqrt{x^2+y^2}} \partial_y \quad (74)$$

ou de façon plus compacte et sans perdre en généralité

$$r \partial_r = x \partial_x + y \partial_y \quad (75)$$

D'autre part

$$\partial_\theta = \frac{\partial x}{\partial \theta} \partial_x + \frac{\partial y}{\partial \theta} \partial_y = -r \sin \theta \partial_x + r \cos \theta \partial_y = -y \partial_x + x \partial_y \quad (76)$$

De là on serait tenté de prendre l'expression de ϕ^{-1} et de différentier. Le calcul devient assez vite immonde. Le plus simple est d'inverser les équations (75) et (76). C'est un système linéaire à deux inconnues (∂_x et ∂_y) dont l'inversion s'apprend en classe secondaire, on trouve après quelques calculs élémentaires

$$\partial_x = \sin \theta \partial_r + \frac{\cos \theta}{r} \partial_\theta, \quad \partial_y = \cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \quad (77)$$

3.4 Dérivées d'ordre supérieur

Si une fonction est \mathcal{C}^1 sur un ouvert et que ses dérivées partielles sont non seulement continues mais aussi elles-mêmes \mathcal{C}^1 , on peut se demander ce qu'il advient des dérivées partielles des dérivées partielles, et pourquoi pas aller plus loin dans l'ordre des dérivées. Nous verrons ensuite comment les fonctions évoluent au voisinage d'un point aux ordres supérieurs.

Théorème de Schwarz

Théorème 3.2. Soit $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$ où \mathcal{U} est un ouvert, de classe \mathcal{C}^2 (elle est \mathcal{C}^1 et ses dérivées partielles sont toutes \mathcal{C}^1 sur \mathcal{U}). Alors

$$\forall a \in \mathcal{U}, \quad \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right) (a) = \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (a) = \frac{\partial^2 f}{\partial x_i \partial x_j} (a) = \frac{\partial^2 f}{\partial x_j \partial x_i} (a) \quad (78)$$

Démonstration. Soit $a = (a_1, a_2, \dots, a_n) \in \mathcal{U}$. Soit $\eta > 0$ tel que

$$\forall (x, y) \in \Omega = [a_i - \eta, a_i + \eta] \times [a_j - \eta, a_j + \eta], \quad a + x e_i + y e_j \in \mathcal{U} \quad (79)$$

Pour $(x, y) \in \Omega$, on pose

$$F(x, y) = \int_{x'=a_i}^x \left(\int_{y'=a_j}^y \partial_j(\partial_i f(a_1, \dots, x', \dots, y', \dots, a_n)) \, dy' \right) dx' \quad (80)$$

où x' est à la i -ième position et y' à la j -ième. En intégrant deux fois on calcule immédiatement que

$$F(x, y) = f(\dots, x, \dots, y, \dots) - f(\dots, a_i, \dots, y, \dots) + f(a) - f(\dots, x, \dots, a_j, \dots) \quad (81)$$

Mais d'après le théorème de Fubini, on a aussi

$$F(x, y) = \int_{y'=a_j}^y \left(\int_{x'=a_i}^x \partial_j(\partial_i f(\dots, x', \dots, y', \dots)) \, dx' \right) dy' \quad (82)$$

D'après (81) on a

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right) (a_i, a_j) = \partial_i \partial_j f(a) \quad (83)$$

mais d'après (82) on calcule aussi

$$\frac{\partial}{\partial x} \left(\frac{\partial F}{\partial y} \right) (a_i, a_j) = \partial_j \partial_i f(a) \quad (84)$$

d'où le résultat. □

Il y a une arnaque : dans cette preuve, nous avons admis le théorème de Fubini. Le théorème de Fubini est difficile à démontrer et s'inscrit dans la théorie de l'intégration. Cela montre que la notion d'intégration est plus puissante que la notion de dérivation.

Remarque : en terme d'opérateurs, cela veut dire que $\partial_i \partial_j = \partial_j \partial_i$.

Développement de Taylor

Théorème 3.3. (Formule de Taylor-Lagrange)

Soit $f : [a, b] \rightarrow \mathbb{R}$, de classe \mathcal{C}^n , telle que $f^{(n+1)}$ existe sur $]a, b[$. Alors

$$\exists c \in]a, b[\quad f(b) = f(a) + \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \frac{(b-a)^{n+1}}{(n+1)^{n+1}} f^{(n+1)}(c) \quad (85)$$

Démonstration. Considérons la fonction

$$\varphi : [a, b] \rightarrow \mathbb{R}; x \mapsto f(b) - f(x) - \sum_{k=1}^n \frac{(b-x)^k}{k!} f^{(k)}(x) - A \frac{(b-x)^{n+1}}{(n+1)!} \quad (86)$$

On choisit A telle que $\varphi(a) = 0$. φ est continue sur $[a, b]$ et dérivable sur $]a, b[$ et un calcul par télescopage de la somme montre que

$$\forall x \in]a, b[, \quad \varphi'(x) = -\frac{(b-x)^n}{n!} f^{(n+1)}(x) + A \frac{(b-x)^n}{n!} \quad (87)$$

On a $\varphi(a) = \varphi(b) = 0$ donc d'après le théorème de Rolle, il existe c dans $]a, b[$ tel que $\varphi'(c) = 0$, ce qui implique $A = f^{(n+1)}(c)$, ce qui achève la preuve. \square

Corollaire 3.1. Sous les mêmes hypothèses, on a

$$f(b) = f(a) + \sum_{k=1}^n \frac{(b-a)^k}{k!} f^{(k)}(a) + \mathcal{O}((b-a)^{n+1}) \quad (88)$$

Nous avons besoin de ces préliminaires sur les développements limités à une variable pour passer à plusieurs variables comme nous allons le voir.

Théorème 3.4. (Taylor-Lagrange à plusieurs variables)

Soit $f : \mathcal{U} \subset \mathbb{R}^p \rightarrow \mathbb{R}^q$ où \mathcal{U} est ouvert, f est \mathcal{C}^n , $x \in \mathcal{U}$, $\mathbf{h} \in \mathbb{R}^p$ tel que $\forall t \in [-1, 1], x + t\mathbf{h} \in \mathcal{U}$. Alors il existe $\theta \in]0, 1[$ tel que

$$f(x + \mathbf{h}) = f(x) + \sum_{k=1}^{n-1} \frac{1}{k!} \partial_{[k]} f(x) h^{[k]} + \frac{1}{n!} \partial_{[n]} f^{[n]}(x + \theta \mathbf{h}) \quad (89)$$

où l'on a posé la notation suivante

$$\partial_{[k]} f^{[k]}(x) h^{[k]} = \sum_{i_1 + i_2 + \dots + i_p = k, i_j \geq 0} \frac{k!}{i_1! i_2! \dots i_p!} \frac{\partial^n f}{\partial x_1^{i_1} \dots \partial x_p^{i_p}}(x) h_1^{i_1} \dots h_p^{i_p} \quad (90)$$

Démonstration. Il suffit d'appliquer la formule de Taylor-Lagrange à une variable sur la fonction $F : [0, 1] \rightarrow \mathbb{R}^q, t \mapsto f(x + t\mathbf{h})$. \square

Corollaire 3.2. Sous les mêmes hypothèses on a

$$f(x + \mathbf{h}) = \sum_{k=0}^{n-1} \frac{1}{k!} \partial_{[k]} f(x) h^{[k]} + \mathcal{O}(\|\mathbf{h}\|^n) \quad (91)$$

3.5 Exercices

Équation d'onde Trouver l'ensemble des fonctions \mathcal{C}^2 sur \mathbb{R}^2 qui vérifient l'équation d'onde

$$\frac{\partial^2 y}{\partial x^2} - \frac{\partial^2 y}{\partial t^2} = 0 \quad (92)$$

Indication : utiliser le changement de variables $u = x + t$, $v = x - t$.

Solution On remarque que le changement de variables s'inverse aisément $x = (u + v)/2$, $t = (v - u)/2$ et que tout se dérive bien dans les deux sens. On a donc bien un \mathcal{C}^1 difféomorphisme de \mathbb{R}^2 dans \mathbb{R}^2 , et même un \mathcal{C}^2 difféomorphisme. Calculons comment se transforment les dérivées partielles.

$$\partial_x = (\partial_x u)\partial_u + (\partial_x v)\partial_v = \partial_u + \partial_v \quad (93)$$

$$\partial_x^2 = \partial_x(\partial_x) = (\partial_u + \partial_v)(\partial_u + \partial_v) = \partial_u^2 + \partial_v^2 + \partial_u\partial_v + \partial_v\partial_u = \partial_u^2 + \partial_v^2 + 2\partial_u\partial_v \quad (94)$$

en vertu du théorème de Schwarz. De même, on calcule que

$$\partial_t^2 = \partial_u^2 + \partial_v^2 - 2\partial_u\partial_v \quad (95)$$

De telle sorte que si l'on pose $y(x, t) = z(u, v)$, l'équation d'onde s'écrit

$$\partial_x^2 y - \partial_t^2 y = 4\partial_u\partial_v z = 0 \quad (96)$$

La dernière égalité nous indique que la fonction $\partial_u z$ ne dépend que de u puisque sa dérivée partielle par rapport à v est nulle sur \mathbb{R}^2 . \mathbb{R}^2 est convexe, donc d'après le théorème 2.4, il existe une fonction f d'une variable \mathcal{C}^1 telle que

$$\partial_u z = f(u) \quad (97)$$

f est intégrable sur \mathbb{R} . Si l'on note F une primitive (qui est au moins \mathcal{C}^2), alors on peut récrire l'équation précédente ainsi

$$\partial_u(z - F(u)) = 0 \quad (98)$$

donc il existe une fonction d'une variable G , \mathcal{C}^2 sur \mathbb{R} telle que $z - F(u) = G(v)$. Réexprimé dans les anciennes variables, on résoud l'équation d'onde ainsi :

$$\exists F, G \in \mathcal{C}^2(\mathbb{R}) / \forall (x, t) \in \mathbb{R}^2, y(x, t) = F(x + t) + G(x - t) \quad (99)$$

Comme on l'enseigne en physique ondulatoire, le premier terme correspond à un signal se propageant dans la direction $-x$ et l'autre à un signal se propageant dans la direction $+x$. Ici nous avons pris des unités naturelles en posant que la célérité des ondes était égale à 1. Pour retrouver le terme physique il suffirait de poser $\tau = ct$, ce qui ne correspond à rien d'autre qu'à un changement d'unité de mesure du temps.

Laplacien

Définition 3.2. Soit une fonction de classe \mathcal{C}^2 sur un ouvert \mathcal{U} de \mathbb{R}^2 . Alors on définit le laplacien de f ainsi :

$$\forall (x, y) \in \mathcal{U}, \Delta f(x, y) = \partial_x^2 f + \partial_y^2 f. \quad (100)$$

Autrement dit, on a $\Delta = \partial_x^2 + \partial_y^2$.

1. On reprend le changement en coordonnées polaires évoqué plus haut. Montrer que

$$\Delta = \partial_r^2 + \frac{1}{r}\partial_r + \frac{1}{r^2}\partial_\theta^2 \quad (101)$$

2. Trouver les fonctions harmoniques, c'est-à-dire les fonctions qui vérifient $\Delta f = 0$, qui ne dépendent que de $r = \sqrt{x^2 + y^2}$.
3. Généralisation : pour n variables, on pose $\Delta = \sum_{i=1}^n \partial_i^2$. Trouver les fonctions harmoniques ne dépendant que de $r = \sqrt{x_1^2 + \dots + x_n^2}$.

Solution.

1. On a $\partial_x = \cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta$. On calcule

$$\partial_x^2 = \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) \quad (102)$$

$$= \cos \theta \partial_r \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) - \frac{\sin \theta}{r} \left(\cos \theta \partial_r - \frac{\sin \theta}{r} \partial_\theta \right) \quad (103)$$

$$= \cos^2 \theta \partial_r^2 + 2 \frac{\cos \theta \sin \theta}{r^2} \partial_\theta + \frac{\sin^2 \theta}{r} \partial_r + \frac{\sin^2 \theta}{r^2} \partial_\theta^2 - 2 \frac{\cos \theta \sin \theta}{r} \partial_\theta \partial_r \quad (104)$$

Avec la même technique on calcule

$$\partial_y^2 = \sin^2 \theta \partial_r^2 - 2 \frac{\sin \theta \cos \theta}{r^2} \partial_\theta + \frac{\cos^2 \theta}{r} \partial_r + \frac{\cos^2 \theta}{r^2} \partial_\theta^2 + 2 \frac{\sin \theta \cos \theta}{r} \partial_\theta \partial_r \quad (105)$$

En utilisant le théorème de Schwarz et en faisant la somme, on trouve l'égalité annoncée.

2. Si f ne dépend que de r alors on a $\partial_\theta f = 0$, d'où

$$\partial_r^2 f + \frac{1}{r} \partial_r f = 0 \quad (106)$$

Comme $\mathbb{R}_+^* \times]-\pi, \pi[$ est convexe, on peut considérer f comme une fonction d'une variable, de telle sorte que $f(x, y) = g(r)$. La fonction g est solution de l'équation différentielle suivante

$$g'' + \frac{1}{r} g' = 0 \quad (107)$$

sur $]0, +\infty[$. On sait que la solution générale est $g'(r) = a \exp(-\ln r) = a/r$ où $a \in \mathbb{R}$, soit $g(r) = a \ln r + b$ où $a, b \in \mathbb{R}$.

3. Si f ne dépend que de r , alors elle ne dépend que de $R = r^2$. Cela revient à étudier F avec $f(r) = F(R)$. On a

$$\Delta F = \sum_i \partial_i^2 \left(F \left(\sum_j x_j^2 \right) \right) = \sum_i \partial_i (2x_i F'(R)) = 2(2RF''(R) + nF'(R)) \quad (108)$$

Les fonctions harmoniques sont donc solutions de

$$F'' + \frac{n}{2R} F' = 0 \quad (109)$$

La solution générale est de la forme $F'(R) = \alpha/R^{n/2}$ où $\alpha \in \mathbb{R}$. Examinons dimension par dimension ce que l'on trouve.

- $n = 1$: $F(R) = a\sqrt{R} + b$ ou encore $f(r) = ar + b$ avec $a, b \in \mathbb{R}$.
- $n = 2$: $F(R) = \tilde{a} \ln R + b$ ou encore $f(r) = a \ln r + b$ où $a, b \in \mathbb{R}$ (en fait on a $a = 2\tilde{a}$).
On retrouve bien le cas précédent.
- $n \geq 3$: $F(R) = a/R^{\frac{n}{2}-1} + b$ soit $f(r) = a/r^{n-2} + b$ avec $a, b \in \mathbb{R}$. Dans le cas particulier où $n = 3$, on retrouve $f(r) = a/r + b$ qui est le potentiel de Coulomb. En effet, pour un champ électrique à symétrie sphérique, pour tout point en un lieu différent de la charge, l'équation du champ est bien $\Delta\phi = 0$.

4 Théorèmes d'inversion

En première approche, on se contente d'énoncer les théorèmes, d'en donner l'heuristique et de les appliquer. Pour ne pas alourdir la première lecture, les démonstrations sont reportées en annexe. Ces théorèmes vont révéler la grande puissance du calcul différentiel, notamment tout ce qui concerne les problèmes inverses, mais permettra aussi de définir les variétés différentielles plongées qui permettent de définir les fonctions implicites, les courbes, les surfaces, les systèmes de coordonnées généralisées. Ici nul besoin d'insister sur la richesse et la multitude des applications physiques qu'on peut en tirer. Nous montrerons quelques exemples.

4.1 Inversion locale

C'est le théorème le plus important de cette section. De ce théorème va découler tout le reste.

Théorème 4.1. (*inversion locale*).

Soit $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, avec \mathcal{U} ouvert, et f de classe \mathcal{C}^1 . Soit $a \in \mathcal{U}$. Si $df(a)$ est une application linéaire bijective, alors

- Il existe $\mathcal{V}_a \subset \mathcal{U}$ un voisinage ouvert de a ,
- Il existe $\mathcal{W}_b \subset \mathbb{R}^n$ un voisinage ouvert de $b = f(a)$,

tels que f soit un \mathcal{C}^1 difféomorphisme de \mathcal{V}_a dans $\mathcal{W}_b = f(\mathcal{V}_a)$. De plus, en notant $f^{-1} : \mathcal{W}_b \rightarrow \mathcal{V}_a$ la bijection réciproque de f restreinte à \mathcal{W}_b , on a

$$\forall (x, y) \in \mathcal{V}_a \times \mathcal{W}_b / y = f(x), \quad df^{-1}(y) = (df(x))^{-1}. \quad (110)$$

En pratique, il faudra vérifier que le jacobien de f en a est non nul. Le théorème semble intuitivement évident : si f est différentiable en a et que sa différentielle est bijective, alors f est localement bijective puisqu'elle se comporte localement comme sa différentielle. Pourtant la preuve est très technique car les majorations de normes nécessaires sont assez subtiles.

Comme on l'a vu avec les difféomorphismes, on peut interpréter ce théorème comme servant de base aux changements de coordonnées. Pour ce faire on a besoin du corollaire suivant.

Corollaire 4.1. Soit $f : \mathcal{U} \rightarrow \mathbb{R}^n$ où \mathcal{U} est un ouvert de \mathbb{R}^n , telle que $\forall x \in \mathcal{U}$, $df(x)$ est inversible. Alors $f(\mathcal{U})$ est ouvert.

Démonstration. Soit $x \in \mathcal{U}$. D'après le théorème d'inversion locale, il existe deux ouverts $\mathcal{V}_x, \mathcal{W}_y$, ouverts de \mathcal{U} et \mathbb{R}^n , voisinages de x et $y = f(x)$, tels que $f(\mathcal{V}_x) = \mathcal{W}_y$. On en déduit, puisque $\mathcal{U} = \bigcup_{x \in \mathcal{U}} \mathcal{V}_x$,

$$f(\mathcal{U}) = f\left(\bigcup_{x \in \mathcal{U}} \mathcal{V}_x\right) = \bigcup_{x \in \mathcal{U}} \mathcal{W}_y \quad (111)$$

qui est ouvert. □

De là il suit le théorème suivant.

Théorème 4.2. Soient f_1, \dots, f_n des fonctions réelles \mathcal{C}^1 au voisinage de $a \in \mathbb{R}^n$. Les relations $u_i = f(x_1, \dots, x_n)$ définissent un changement de coordonnées sur un voisinage de a si et seulement si $df_i(a)$ sont des formes linéaires indépendants.

C'est une autre façon d'exprimer le théorème d'inversion locale.

4.2 Inversion globale

Corollaire 4.2. (Théorème d'inversion globale).

Soit $f : \mathcal{U} \subset \mathbb{R}^n \rightarrow \mathcal{V} = f(\mathcal{U})$ où \mathcal{U} est ouvert et $f \in \mathcal{C}^1$. Si f est injective alors on a l'équivalence entre les deux propositions suivantes.

1. $\forall x \in \mathcal{U}, df(x)$ est inversible.
2. $\mathcal{V} = f(\mathcal{U})$ est ouvert, f est globalement inversible et f^{-1} est \mathcal{C}^1 .

Démonstration. $2 \Rightarrow 1$. f et f^{-1} sont toutes deux \mathcal{C}^1 et en tout point on a $f^{-1} \circ f = Id$, soit $\det(df^{-1}(f(x)))\det(df(x)) = 1$, ce qui prouve que $df(x)$ est inversible pour tout x .

$1 \Rightarrow 2$. Il est évident que f est surjective par construction, par hypothèse elle est donc bijective. Pour prouver que son inverse est \mathcal{C}^1 il suffit d'appliquer le théorème d'inversion locale en chaque point pour obtenir l'expression de df^{-1} qui est donc bien définie. \square

Exemple important : coordonnées polaires Nous invitons les lecteurs à relire l'exercice 3.3.3 à la lumière des nouveaux théorèmes que nous avons admis à ce moment-là. Remarquons que nous pourrions prendre tout intervalle ouvert de longueur 2π comme ensemble de définition de l'angle pour assurer l'injectivité de la transformation. Alors la demi-droite que l'on excluerait serait l'origine arbitrairement choisie du repère polaire.

Courbe paramétrée Soit $\gamma : I \subset \mathbb{R} \rightarrow \Gamma \subset \mathbb{R}^3, t \mapsto \gamma(t) = (x(t), y(t), z(t))$. Si I est ouvert, $\gamma \in \mathcal{C}^1$ sur I (x, y, z le sont), et injective, alors γ est un \mathcal{C}^1 difféomorphisme de I dans Γ . Remarquons que dire que γ est injective revient à dire que la courbe Γ ne se recoupe pas elle-même (voir figure 11).

Cet exemple devrait vous rappeler quelque chose

$$\gamma(t) = \left(-\frac{1}{2}gt^2 + v_{0z}t\right) e_z + v_{0x}te_x, \quad t \in \left[0, \frac{2v_{0z}}{g}\right], \quad v_{0x}, v_{0z}, g > 0 \quad (112)$$

Il est aisé de montrer que cette courbe paramétrée réalise une bijection entre $[0, 2v_{0z}/g]$ et la courbe d'équation cartésienne (voir figure 12)

$$z = -\frac{g}{2v_{0x}^2}x^2 + \frac{v_{0z}}{v_{0x}}x, \quad y = 0, \quad x \in [0, 2v_{0x}v_{0z}/g]. \quad (113)$$

Un exemple que nous traiterons abondamment ultérieurement constitue les surfaces régulières.

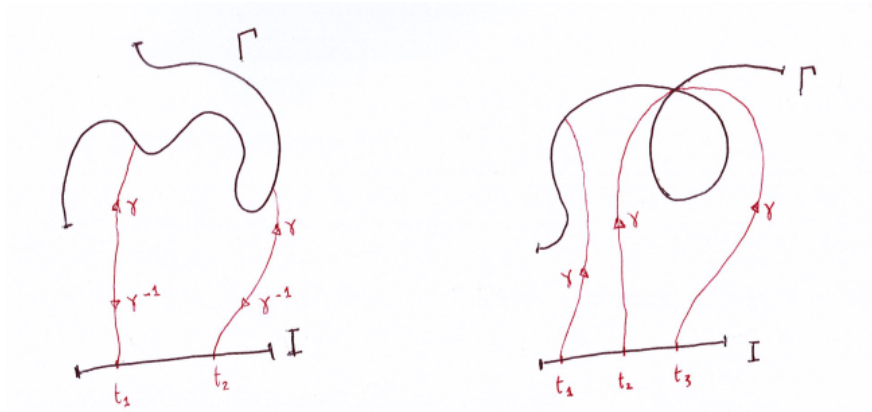


FIGURE 11 – À gauche, un arc paramétré injectif et régulier, donc bijectif. À droite, un arc paramétré régulier, mais non injectif : on voit que $\gamma(t_2) = \gamma(t_3)$ avec $t_2 \neq t_3$.

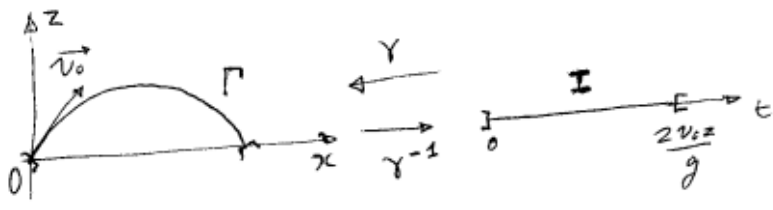


FIGURE 12 – Graphe de la trajectoire d'une chute libre.

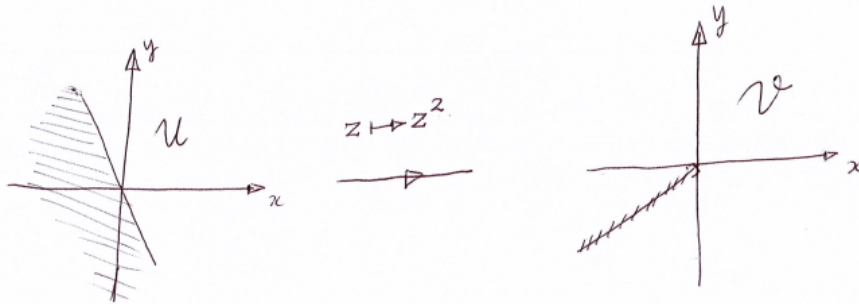


FIGURE 13 – Exemple d'ensemble de départ et d'ensemble d'arrivée pour que la fonction carré soit un \mathcal{C}^1 difféomorphisme.

Exercice Soit $\mathcal{U} = \mathbb{R}^2 \setminus \{(0, 0)\}$. On pose $f : \mathcal{U} \rightarrow \mathbb{R}^2, (x, y) \mapsto (x^2 - y^2, 2xy)$.

1. Montrer que f est localement inversible sur tout \mathcal{U} mais pas globalement inversible.
2. Déterminer deux ouverts \mathcal{V} et \mathcal{W} tels que f soit un \mathcal{C}^1 difféomorphisme de \mathcal{V} dans \mathcal{W} .

Solution.

1. f est différentiable sur \mathcal{U} et on calcule aisément $J_f = 4(x^2 + y^2) \neq 0$ sur tout \mathcal{U} . D'après le théorème d'inversion locale, en tout point de \mathcal{U} , il existe un voisinage et un voisinage de l'image de ce point tels que f soit un \mathcal{C}^1 difféomorphisme reliant ces deux ensembles. Autrement dit, f est localement inversible sur tout \mathcal{U} . Cependant, on constate qu'en tout point de \mathcal{U} , on a $f(x, y) = f(-x, -y)$. f n'est donc pas injective, donc n'est pas globalement inversible.
2. En posant $z = x + iy$, on constate que dans \mathbb{C} , f est représentée par la fonction carré. Le résultat de la question précédente montre qu'il est possible de trouver en tout point une détermination univoque locale de la racine carrée d'un nombre complexe. Pour trouver le plus "gros" \mathcal{C}^1 difféomorphisme global, il faut trouver un ensemble qui ne contient pas à la fois z et $-z$ afin de rendre la fonction carré injective. Une solution est de prendre tout demi-plan dont la coupure contient l'origine pour \mathcal{U} , et l'image de \mathcal{U} donnera \mathbb{R}^2 privé de la demi-droite dont l'origine est $(0, 0)$ et dont le vecteur directeur sera orthogonal à la coupure du demi-plan antécédent et dirigé vers le complémentaire du demi-plan. Voir figure 13.

4.3 Fonctions implicites

4.3.1 Approche heuristique

Imaginons qu'on cherche à résoudre une équation du type $f(x, y) = 0$ au voisinage de (a, b) sachant que $f(a, b) = 0$. Autrement dit, on cherche à déterminer la ligne de niveau de f au voisinage de (a, b) et passant par (a, b) , représentée comme le graphe d'une fonction de x . Sans souci de rigueur pour le moment, on peut écrire

$$f(x, y) \approx f(a, b) + \partial_x f(a, b)(x - a) + \partial_y f(a, b)(y - b) \quad (114)$$

et en déduire que, si $\partial_y f(a, b) \neq 0$,

$$y = \phi(x) \approx b + \frac{\partial_x f(a, b)}{\partial_y f(a, b)}(x - a) \quad (115)$$

et on peut même calculer $\phi'(x) = \partial_x f(a, b) / \partial_y f(a, b)$. Le théorème d'inversion locale nous assurera l'existence et l'unicité de ϕ .

4.3.2 Théorème des fonctions implicites

Avant d'énoncer le théorème nous aurons besoin d'une notation.

Définition 4.1. Soient E_1, \dots, E_n des espaces vectoriels normés et $f : \Omega \subset E_1 \times \dots \times E_n \rightarrow F$, où Ω est ouvert. Soit $a = (a_1, \dots, a_n) \in \Omega$. Pour $i \in \llbracket 1, n \rrbracket$, la fonction $f_i : x \mapsto f(a_1, \dots, x, \dots, a_n)$ où x est à la i -ième position, est définie sur un voisinage de a_i dans E_i . Si f_i est différentiable en a_i , alors on dit que f admet une différentielle partielle d'indice i , et on note

$$\partial_i f(a) = df_i(a_i) \quad (116)$$

Évidemment, si f est différentiable en a , on a

$$df(a)\mathbf{h} = \sum_{i=1}^n \partial_i f(a)\mathbf{h}_i \quad (117)$$

Si $E_i = \mathbb{R}$, $\partial_i f(a) = \partial f(a) / \partial x_i$.

Théorème 4.3. (fonctions implicites)

Soient E, F deux espaces vectoriels normés, $\mathcal{U} \subset E$, $\mathcal{V} \subset F$; ouverts, et $f : \mathcal{U} \times \mathcal{V} \rightarrow F$ de classe \mathcal{C}^1 . Soit $(a, b) \in \mathcal{U} \times \mathcal{V}$. Si $\partial_2 f(a, b)$ est une application linéaire bijective, alors

- il existe $\mathcal{U}_a \subset \mathcal{U}$ un voisinage ouvert de a ;
- il existe $\mathcal{V}_b \subset \mathcal{V}$ un voisinage ouvert de b ;
- il existe $\mathcal{W}_c \subset F$ un voisinage ouvert de $c = f(a, b)$;
- il existe une fonction φ , unique, de classe \mathcal{C}^1 , $\varphi : \mathcal{U}_a \times \mathcal{W}_c \rightarrow \mathcal{V}_b$

telles que

$$\forall z \in \mathcal{W}_c, \forall x \in \mathcal{U}_a, f(x, \varphi(x, z)) = z \quad (118)$$

L'idée de la preuve est d'appliquer le théorème d'inversion locale à la fonction $(x, y) \mapsto (x, f(x, y))$. Nous reportons cette preuve en annexe.

Souvent, lorsqu'on cherche des fonctions implicites, on fixe la valeur de z et on ne cherche qu'une ligne de niveau de f . Mais la version générale de ce théorème permet aussi de caractériser comment les lignes de niveau varient quand l'image de f varie.

4.3.3 Interprétation graphique

Essayons d'interpréter graphiquement ce théorème dans le cas où $E = F = \mathbb{R}$. Nous avons représenté un cas fictif sur la figure 14. Pour que le théorème puisse s'appliquer, il faut que f varie le long de e_y . Tant que les lignes de niveau ne sont pas verticales, alors le long de l'axe y , puisqu'on traverse des lignes de niveau, la valeur de $f(x, y)$ varie, donc $\partial_y f \neq 0$. Pour avoir un graphe local de y comme fonction de x , il faut donc que ses lignes de niveau ne soient pas tangentes à e_y , autrement dit il ne faut pas qu'elles soient verticales si on les représente sur un repère cartésien classique. Si tel est le cas, il est toujours possible d'appliquer ce théorème à la variable x et non y , et d'exprimer localement x comme le graphe d'une fonction de y .

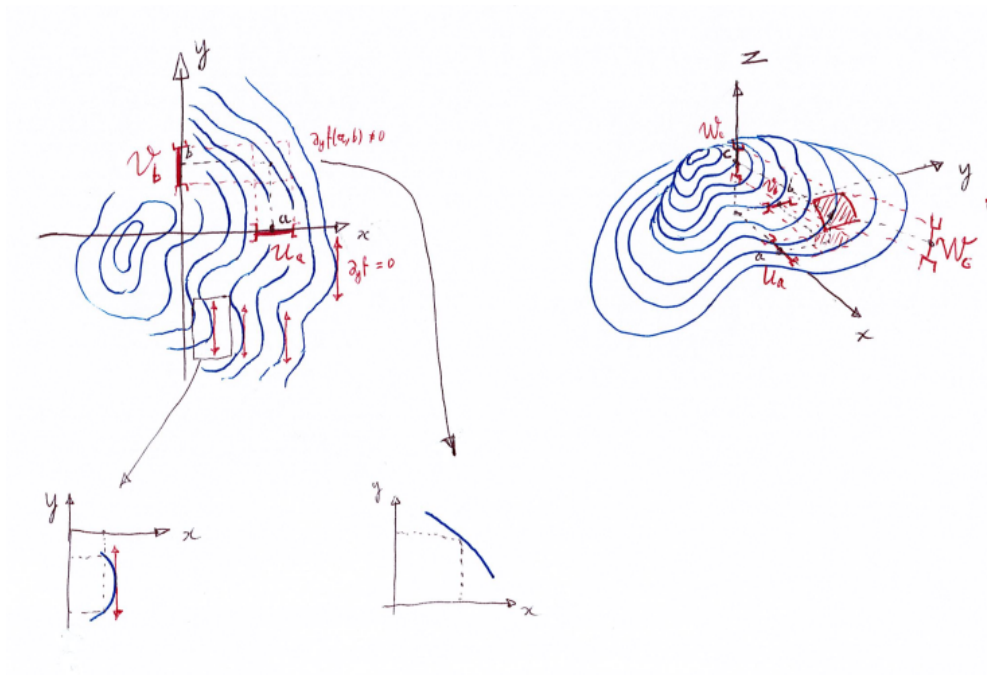


FIGURE 14 – Nous représentons en haut à gauche les lignes de niveau d’une fonction de deux variables à valeurs dans \mathbb{R} . Nous avons représenté les voisinages correspondant aux notations du théorème en un point où $\partial_2 f(a, b) \neq 0$. Nous avons indiqué quelques cas où $\partial_2 f(a, b) \neq 0$ donc avec une tangente verticale. En dessous, nous avons représenté deux zooms : l’un où $\partial_2 f = 0$, auquel cas la ligne n’est localement pas le graphe d’une fonction (un antécédent peut donner plusieurs images), et l’autre où $\partial_2 f \neq 0$, dans ce cas la ligne est localement le graphe d’une fonction de x . À droite nous avons tenté de faire une représentation en perspective pour montrer à quoi correspond $\mathcal{W}_c = f(\mathcal{U}_a \times \mathcal{V}_b)$.

4.3.4 Différentielle de la fonction implicite

Supposons les conditions du théorème des fonctions implicites réalisées. Localement, on a $f(x, \varphi(x, z)) = z$ pour tous x et z dans le bon voisinage. Soient $\mathbf{h}_1, \mathbf{h}_3 \in E \times F$. On a

$$d(f(x, \varphi(x, z))) (\mathbf{h}_1, \mathbf{h}_3) = \partial_1 f(x, \varphi(x, z)) \mathbf{h}_1 + \partial_2 f(x, \varphi(x, z)) \circ d\varphi(x, z) (\mathbf{h}_1, \mathbf{h}_3) \quad (119)$$

mais aussi

$$d(f(x, \varphi(x, z))) (\mathbf{h}_1, \mathbf{h}_3) = \mathbf{h}_3 \quad (120)$$

Par hypothèse, $\partial_2 f(x, \varphi(x, z))$ est inversible, on peut donc calculer $d\varphi(x, z)$:

$$d\varphi(x, z) (\mathbf{h}_1, \mathbf{h}_3) = (\partial_2 f(x, \varphi(x, z)))^{-1} (\mathbf{h}_3 - \partial_1 f(x, \varphi(x, z)) \mathbf{h}_1) \quad (121)$$

soit

$$d\varphi(x, z) = (\partial_2 f(x, \varphi(x, z)))^{-1} \circ (Id_F - \partial_1 f(x, \varphi(x, z))) \quad (122)$$

Remarquons que si l'on considère que $z = c$ n'est pas une variable, la formule devient simplement

$$d\varphi_c(x) = -(\partial_2 f(x, \varphi_c(x)))^{-1} \circ \partial_1 f(x, \varphi_c(x)) \quad (123)$$

en ayant posé $\varphi_c(x) = \varphi(x, c)$.

En règle générale, il est absolument inutile d'apprendre par coeur cette formule. En effet, on peut toujours retomber sur ses pieds en différenciant l'équation $f(x, y) = 0$ comme nous allons le voir dans les exemples qui suivent.

4.3.5 Exercices sur les fonctions implicites

Échauffement : le cercle On souhaite résoudre l'équation $x^2 + y^2 = 1$ en y . On pose $f(x, y) = x^2 + y^2 - 1$ et on cherche les lignes de niveau de f lorsqu'elle prend la valeur 0. On calcule $\partial_y f(x, y) = 2y$, pour pouvoir appliquer le théorème des fonctions implicites il faut donc exclure les points où $y = 0$. Cette équation définit deux fonctions "implicites" (qu'on peut ici déterminer explicitement, en général ce n'est pas le cas) :

$$\varphi_+ :]-1, 1[\rightarrow]0, 1[, x \mapsto \sqrt{1 - x^2} \quad (124)$$

$$\varphi_- :]-1, 1[\rightarrow]0, 1[, x \mapsto -\sqrt{1 - x^2} \quad (125)$$

qui définissent les deux demi-cercles supérieurs et inférieurs. On ne peut pas définir de fonction de x au voisinage des points $(-1, 0)$ et $(1, 0)$, mais on peut le faire dans l'autre sens. Cela donnerait les deux fonctions suivantes

$$\psi_d :]-1, 1[\rightarrow]0, 1[, y \mapsto \sqrt{1 - y^2} \quad (126)$$

$$\psi_g :]-1, 1[\rightarrow]0, 1[, y \mapsto -\sqrt{1 - y^2} \quad (127)$$

qui correspondent aux demi-cercles à droite et à gauche de l'axe $(0y)$. Calculons les différentielles (ici, simplement les dérivées) des fonctions φ_{\pm} . On pourrait dériver brutalement les fonctions explicitement obtenues ou utiliser la formule obtenue plus haut, mais soyons plus malins. En différenciant l'équation $f(x, \varphi_{\pm}(x)) = 0$ on obtient $x dx + \varphi_{\pm}(x) d\varphi_{\pm} = 0$ soit

$$\varphi'_{\pm}(x) = -\frac{x}{\varphi_{\pm}(x)} = \mp \frac{x}{\sqrt{1 - x^2}} \quad (128)$$

Ce petit exemple montre la méthode pour obtenir la dérivée de la fonction implicite.

Théorie des perturbations Le théorème des fonctions implicites est à la base de la théorie des perturbations. En effet, la plupart des "vraies" équations de la physique sont insolubles, mais on peut tenter de les résoudre approximativement au voisinage d'un point connu. On la résoud d'abord "à l'ordre 0" en assimilant le voisinage du point au point connu, puis "à l'ordre 1" en utilisant les dérivées premières et en y injectant le résultat à l'ordre 0, et éventuellement aux ordres supérieurs pour repousser le reste à des ordres plus élevés. Une question mathématique intéressante est de savoir si la série converge et si le reste devient nul lorsque l'ordre tend vers l'infini. Ce genre de question occupe les mathématiques des systèmes dynamiques non intégrables qui s'écrivent sous la forme d'équations différentielles insolubles du type $y' = f(y) + \varepsilon g(y)$ dans le cas où l'on connaît la solution de $y' = f(y)$ et que le paramètre ε est petit. La recherche mathématique actuelle s'occupe de déterminer des conditions pour que la théorie des perturbation converge vers une solution exacte (dans la plupart des cas, les séries ne convergent pas). Ce type de problème est notamment étudié en M2 DSG. Pour ce cours, nous allons donner une illustration plus simple qui permet de toucher le problème du doigt.

Soit l'équation (E) : $x^7 + 0,99x - 2,03 = 0$.

1. Montrer que cette équation a une solution unique dans \mathbb{R} .
2. Trouver sans calculatrice une solution à 10^{-1} près, à 10^{-3} près, puis à 10^{-5} près. On pensera à poser $f : \mathbb{R}^3 \rightarrow \mathbb{R}, (x, p, q) \mapsto x^7 + px - q$ et à rechercher les bonnes fonctions implicites.
3. Calculer les dérivées partielles à l'ordre 1 et à l'ordre 2 de la fonction implicite φ qui vérifie $f(\varphi(p, q), p, q) = 0$ au voisinage de $(p, q) = (1, 2)$.

Solution.

1. Soit $g : x \mapsto x^7 + 0,99x - 2,03$. g est continue et dérivable sur \mathbb{R} et $\forall x \in \mathbb{R}, g'(x) = 7x^6 + 0,99 > 0$. De plus on a $\lim_{\pm\infty} g = \pm\infty$ donc g réalise une bijection de \mathbb{R} dans \mathbb{R} donc $g(x) = 0$ admet une solution unique.
2. On pose la fonction f suggérée. Pour $(p, q) = (1, 2)$, une solution évidente (et unique) est $x = 1$. C'est la solution "à l'ordre 0" de (E), précise à 10^{-1} près.

Pour appliquer le théorème des fonctions implicites, il faut que $\partial_x f(p, q) = 7x^6 + p \neq 0$. Or cela n'est possible que si $p \leq 0$ ce qui n'est pas le cas ici. Donc il existe un voisinage ouvert de $(1, 2)$ et une fonction unique $\varphi(p, q)$ définie sur ce voisinage qui vérifie $f(\varphi(p, q), p, q) = 0$. D'après ce qui précède, on a $\varphi(1, 2) = 1$ et on va essayer de déterminer approximativement $\varphi(0,99; 2,03)$ grâce à un développement limité.

À présent posons $(p, q) = (1 + \delta p, 2 + \delta q)$, et notons symboliquement

$$\tilde{x} = x + dx + d^2x + \mathcal{O}(dx^3) = \varphi(1, 2) + d\varphi(1, 2)(\delta p, \delta q) + (1/2)d^2\varphi(1, 2)(\delta p, \delta q) + \mathcal{O}(\|(\delta p, \delta q)\|^3) \quad (129)$$

afin de simplifier les notations. Différentiation l'équation à résoudre.

$$df(x, p, q) = 7x^6 dx + x dp + p dx - dq = 0 \quad (130)$$

En substituant le résultat à l'ordre 0 en x, p , et q , les termes d'ordre 0 se sont implicitement éliminés pour $x = 1, p = 1, q = 2$ et il ne reste que les termes d'ordre 1, ce qui permet de déduire dx :

$$dx = -\frac{x}{7x^6 + p} dp + \frac{1}{7x^6 + p} dp \quad (131)$$

En remplaçant x par 1, p par 1, q par 2, dp par $\delta p = -0.001$, dq par $\delta q = 0.02$, on obtient $\delta x = 0.005$ ce qui donne $\tilde{x} \approx x + \delta x = 1.005$, résultat précis à 10^{-3} .

Pour augmenter la précision il faut aller chercher le résultat à l'ordre 2. Si l'on fait varier l'équation (E) à l'ordre 2 on trouve ceci

$$(x + dx + d^2x)^7 + xp + xdp + pdx - q - dq = x^7 + 7x^6(dx + d^2x) + 21x^5(dx)^2 + xp + xdp + pdx - q - dq + \mathcal{O}(\|(dp, dq)\|^3) \quad (132)$$

mais on sait que $x^7 + px - q = 0$ et que $7x^6dx + xdp + pdx - dq = 0$, il ne reste donc que les termes d'ordre 2, ce qui donne

$$d^2x = -\frac{3dx^2}{x} - \frac{dxdp}{7x^6} \quad (133)$$

et en remplaçant par les bonnes valeurs (notamment $dx = 0.005$) on trouve $d^2x = -6 \cdot 10^{-5}$ ce qui donne un résultat précis à 10^{-5} : $x \approx 1 + 0.05 - 0.00006 = 1.00494$.

3. Pour les dérivées d'ordre 1 il suffit d'identifier les termes de la différentielle de l'équation (131). Cela nous donne donc

$$\partial_p \varphi(p, q) = -\frac{x}{7x^6 + p} = -\frac{\varphi(p, q)}{7\varphi(p, q)^6 + p} \quad (134)$$

$$\partial_q \varphi(p, q) = \frac{1}{7x^6 + p} = \frac{1}{7\varphi(p, q)^6 + p} \quad (135)$$

Pour les dérivées d'ordre 2 il faut développer l'équation (133) en y injectant l'expression de dx . Après calculs on trouve ceci

$$d^2x = \frac{x}{7x^6 + p} \left(-\frac{3}{7x^6 + p} + \frac{1}{7x^6} \right) dp^2 - \frac{3}{x(7x^6 + p)} dq^2 + \frac{1}{7x^6 + p} \left(\frac{6}{7x^6 + p} - \frac{1}{7x^6} \right) dpdq \quad (136)$$

De là il suffit d'identifier avec

$$d^2x = \frac{1}{2} d^2 \varphi(p, q) = \frac{1}{2} (\partial_p^2 \varphi(p, q) dp^2 + \partial_q^2 \varphi(p, q) dq^2 + 2\partial_p \partial_q \varphi(p, q) dpdq) \quad (137)$$

que l'on déduit du corollaire 3.2. Soit

$$\partial_p^2 \varphi(p, q) = 2 \frac{x}{7x^6 + p} \left(-\frac{3}{7x^6 + p} + \frac{1}{7x^6} \right) \quad (138)$$

$$\partial_q^2 \varphi(p, q) = -\frac{6}{x(7x^6 + p)} \quad (139)$$

$$\partial_p \partial_q \varphi(p, q) = \frac{1}{7x^6 + p} \left(\frac{6}{7x^6 + p} - \frac{1}{7x^6} \right) \quad (140)$$

où l'on substitue $\varphi(p, q)$ à x (mentalement seulement pour éviter d'alourdir les notations).

Thermodynamique En thermodynamique des gaz, les trois variables p, v, t sont souvent reliées par une équation d'état de la forme $f(p, v, t) = 0$. Notamment, l'équation des gaz parfait correspond à $f(p, v, t) = pv - t$ en unités naturelles. Si l'on suppose que $p, v, t > 0$ et que toutes les dérivées partielles de f sont non nulles en p_0, v_0, t_0 , alors on peut appliquer le théorème des fonctions implicites trois fois : il existe trois uniques fonctions de deux variables π, ν, τ telles qu'on ait, au voisinage de (p_0, v_0, t_0) :

$$f(\pi(v, t), v, t) = 0 \quad (141)$$

$$f(p, \nu(p, t), t) = 0 \quad (142)$$

$$f(p, v, \tau(p, v)) = 0 \quad (143)$$

Lorsque les physiciens choisissent de considérer l'une de ces trois fonctions, cela signifie qu'ils choisissent de considérer que toutes les autres fonctions ne dépendront que des deux autres variables.

Les physiciens considèrent souvent des fonctions d'état $u : \mathbb{R}_+^{*3} \rightarrow \mathbb{R}, (p, v, t) \mapsto u(p, v, t)$. Cependant, comme les variables (p, v, t) sont reliées par l'équation d'état, les physiciens notent souvent par exemple $u(p, v, t) = u(p, v)$, ou encore $u(p, v, t) = u(p, t)$ ou $u(p, v, t) = u(v, t)$ pour signifier qu'ils choisissent un jeu de variables, alors qu'ils considèrent en fait respectivement les fonctions $\alpha(p, v) = u(p, v, \tau(p, v)), \beta(p, t) = u(p, \nu(p, t), t)$, et $\gamma(v, t) = u(\pi(v, t), v, t)$. L'avantage qu'ils en tirent est que les mêmes quantités physiques gardent les mêmes lettres. L'inconvénient est qu'ils se mélangent les pinceaux lorsqu'ils calculent les dérivées partielles. Pour éviter de s'embrouiller, ils ont trouvé une astuce et notent les dérivées partielles ainsi

$$\partial_p \alpha = \left(\frac{\partial u}{\partial p} \right)_v \quad (144)$$

$$\partial_p \beta = \left(\frac{\partial u}{\partial p} \right)_t \quad (145)$$

et ainsi de suite.

1. Calculer les dérivées partielles de π, ν, τ , dans le cas du gaz parfait et dans le cas général.
2. On considère le gaz parfait. Montrer que

$$\left(\frac{\partial u}{\partial p} \right)_t - \left(\frac{\partial u}{\partial p} \right)_v = -v \left(\frac{\partial u}{\partial t} \right)_p = -\frac{v}{p} \left(\frac{\partial u}{\partial v} \right)_p \quad (146)$$

3. Cas général : montrer que

$$\left(\frac{\partial u}{\partial p} \right)_t - \left(\frac{\partial u}{\partial p} \right)_v = \frac{\partial_p f}{\partial_t f} \left(\frac{\partial u}{\partial t} \right)_p = -\frac{\partial_p f}{\partial_v f} \left(\frac{\partial u}{\partial v} \right)_p \quad (147)$$

4. Vérifier que

$$\left(\frac{\partial p}{\partial v} \right)_t \left(\frac{\partial v}{\partial t} \right)_p \left(\frac{\partial t}{\partial p} \right)_v = -1 \quad (148)$$

Solution.

1. Dans le cas du gaz parfait on a $\pi(v, t) = t/v$ d'où l'on tire aisément $\partial_v \pi = -t/v^2$ et $\partial_t \pi = 1/v$. De même, on trouve $\nu(p, t) = t/p$ d'où $\partial_p \nu = -t/p^2$, $\partial_t \nu = 1/p$, et de même $\tau(p, v) = pv$ d'où $\partial_p \tau = v$ et $\partial_v \tau = p$.

Dans le cas général il suffit d'écrire $df = \partial_p f dp + \partial_v f dv + \partial_t f dt = 0$ et pour obtenir par exemple les dérivées partielles de π , on aurait

$$d(f(\pi(v, t), v, t)) = \partial_p f d\pi(v, t) + \partial_v f dv + \partial_t f dt = 0 \quad (149)$$

ce qui nous donne

$$d\pi = -\frac{\partial_v f}{\partial_p f} dv - \frac{\partial_t f}{\partial_p f} dt \quad (150)$$

d'où

$$\partial_v \pi = -\frac{\partial_v f}{\partial_p f} \quad \partial_t \pi = -\frac{\partial_t f}{\partial_p f} \quad (151)$$

De même

$$\partial_p \nu = -\frac{\partial_p f}{\partial_v f} \quad \partial_t \nu = -\frac{\partial_t f}{\partial_v f} \quad \partial_p \tau = -\frac{\partial_p f}{\partial_t f} \quad \partial_v \tau = -\frac{\partial_v f}{\partial_t f} \quad (152)$$

2. Dans le cas du gaz parfait, on a $\pi(v, t) = v/t$, $\nu(p, t) = t/p$, $\tau(p, v) = pv$. On a

$$\left(\frac{\partial u}{\partial p}\right)_t = \partial_p \beta = \partial_p(u(p, t/p, t)) = \partial_1 u - \frac{t}{p^2} \partial_2 u = \partial_1 u - \frac{v}{p} \partial_2 u \quad (153)$$

$$\left(\frac{\partial u}{\partial p}\right)_v = \partial_p \alpha = \partial_p(u(p, v, pv)) = \partial_1 u + v \partial_3 u \quad (154)$$

$$\left(\frac{\partial u}{\partial t}\right)_p = \partial_t \beta = \partial_t(u(p, t/p, t)) = \frac{1}{p} \partial_2 u + \partial_3 u \quad (155)$$

$$\left(\frac{\partial u}{\partial v}\right)_p = \partial_v \alpha = \partial_v(u(p, v, pv)) = \partial_2 u + p \partial_3 u \quad (156)$$

En combinant les bonnes équations on en déduit aisément le résultat demandé.

3. Avec la même méthode on a

$$\left(\frac{\partial u}{\partial p}\right)_t = \partial_p \beta = \partial_p(u(p, \nu(p, t), t)) = \partial_1 u + \partial_p \nu \partial_2 u = \partial_1 u - \frac{\partial_p f}{\partial_v f} \partial_2 u \quad (157)$$

$$\left(\frac{\partial u}{\partial p}\right)_v = \partial_p \alpha = \partial_p(u(p, v, \tau(p, v))) = \partial_1 u + \partial_p \tau \partial_3 u = \partial_1 u - \frac{\partial_p f}{\partial_t f} \partial_3 u \quad (158)$$

$$\left(\frac{\partial u}{\partial t}\right)_p = \partial_t \beta = \partial_t(u(p, \nu(p, t), t)) = -\frac{\partial_t f}{\partial_v f} \partial_2 u + \partial_3 u \quad (159)$$

$$\left(\frac{\partial u}{\partial v}\right)_p = \partial_v \alpha = \partial_v(u(p, v, \tau(p, v))) = \partial_2 u + \partial_v \tau \partial_3 u = \partial_2 u - \frac{\partial_v f}{\partial_t f} \partial_3 u \quad (160)$$

En combinant les bonnes équations on en déduit aisément le résultat demandé.

4. D'après 1. on a

$$\left(\frac{\partial p}{\partial v}\right)_t \left(\frac{\partial v}{\partial t}\right)_p \left(\frac{\partial t}{\partial p}\right)_v = \partial_v \pi \partial_t \nu \partial_p \tau = \left(-\frac{\partial_v f}{\partial_p f}\right) \left(-\frac{\partial_t f}{\partial_v f}\right) \left(-\frac{\partial_p f}{\partial_t f}\right) = -1 \quad (161)$$

5 Annexe : preuves techniques

5.1 Différentiabilité

5.2 Théorèmes d'inversion

5.2.1 Théorème d'inversion locale

5.2.2 Théorème des fonctions implicites