MAT 202 : Calcul différentiel et intégral des fonctions à plusieurs variables

Sylvain Lavau

Galatasaray Üniversitesi

Table des matières

Remerciements. *Ces notes de cours s'appuient sur plusieurs manuscrits, dont :*

— *F. Liret et D. Martinais,* Analyse 2ème année, *Dunod.*

— *Susan Jane Colley,* Vector Calculus, *Pearson.*

— *le cours de fonctions de plusieurs variables d'EXO 7*

— *les cours de F. Fayard et de D. Choimet au Lycée du Parc sur les années 2006-08*

Le présent manuscrit contient plusieurs images tirées de ces ressources, ainsi que de Wikipedia (entre autres) que je remercie ici. Il y a bien sûr encore des fautes et des approximations donc n'hésitez pas à m'en faire part.

1 Géométrie des fonctions à deux variables

Dans ce cours nous étudions des fonctions continues (pas forcément linéaires, donc pas forcément représentables par des matrices) entre deux espaces de dimension finie E et E'. Nous commençons par rappeler la notion de continuité pour une fonction à plusieurs variables, et un résultat que nous admettrons.

Définition 1.1. *Soit (E, N) et (E', N') deux espaces vectoriels normés (de dimension finie ou infinie). Soit $D \subset E$ une partie de E, et $f : D \to E'$ une application. On dit que f* est continue *en $a \in D$ si :*

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \text{ tel que } \forall\, x \in D,\ \text{si } N(x-a) < \delta \text{ alors } N(f(x) - f(a)) < \epsilon$$

On dit que f est continue sur *D si f est continue en tout point de D.*

Exemple 1.2. Si on définit la fonction à deux variables suivante :

$$\begin{aligned} f :\ \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto f(x,y) = \begin{cases} \frac{y^3}{\sqrt{x^2+y^4}} \text{ si } (x,y) \neq (0,0) \\ 0 \text{ si } (x,y) \neq (0,0) \end{cases} \end{aligned}$$

Elle est continue en tout point différent de l'origine, mais étudions sa continuité à l'origine. Nous observons que pour $0 < |x| \leq 1$, on a $x^4 \leq x^2$. On a donc légitimement $\sqrt{x^4+y^4} \leq \sqrt{x^2+y^4}$ pour x suffisamment proche de 0. Nous nous plaçons dans ce cas, nous avons alors pour tout $(x,y) \neq (0,0)$:

$$|f(x,y) - f(0,0)| = \frac{|y|^3}{\sqrt{x^2+y^4}} \leq \frac{|y|^3}{\sqrt{x^4+y^4}}$$

Si on effectue un changement de variable en écrivant $x = r\cos(\theta)$ et $y = r\sin(\theta)$, avec $r \in]0, +\infty[$, nous obtenons :

$$|f(x,y) - f(0,0)| \leq \frac{r^3|\sin(\theta)|^3}{\sqrt{r^4\cos^4(\theta) + r^4\sin^4(\theta)}} = \frac{r|\sin(\theta)|^3}{\sqrt{\cos^4(\theta) + \sin^4(\theta)}}$$

Le terme de droite tend vers 0 lorsque r tend vers 0, quel que soit le choix d'angle (qui peut changer). On voit donc que $f(x,y) \underset{(x,y)\to(0,0)}{\longrightarrow} f(0,0)$ c'est à dire : f est continue en $(0,0)$.

Exemple 1.3. Si on définit la fonction à deux variables suivante :

$$\begin{aligned} f :\ \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases} \end{aligned}$$

La fonction est bien entendue continue sur le plan épointé, c'est à dire pour tout $(x,y) \neq (0,0)$, mais est-elle continue en l'origine ? Passons aux polaires, on prend $r > 0$ et $\theta \in \mathbb{R}$, et on pose $x = r\cos(\theta)$ et $y = r\sin(\theta)$. On a alors :

$$f\,(x,y) = f(r\cos(\theta), r\sin(\theta)) = \sin(\theta)\cos(\theta) = \frac{\sin(2\theta)}{2}$$

On voit que le résultat ne dépend que de l'angle θ, pas de la distance. Donc, si on approche de l'origine $(0,0)$ en venant de la direction θ, c'est à dire en faisant tendre r vers 0 tout en gardant θ constant, on voit que la valeur de la fonction f est constante selon ce chemin :

$$\forall\, r, r' > 0 \qquad f\,(r\cos(\theta), r\sin(\theta)) = \frac{\sin(2\theta)}{2} = f\,(r'\cos(\theta), r'\sin(\theta))$$

Et donc à la limite $r \to 0$, on voit que :

$$\lim_{\substack{r\to 0 \\ r\neq 0}} f\,(r\cos(\theta), r\sin(\theta)) = \frac{\sin(2\theta)}{2}$$

On ne tend donc certainement pas vers 0 (la valeur de f en $(0,0)$). La fonction f n'est pas continue à l'origine.

En dimension finie, nous avons l'avantage d'avoir des bases de cardinal fini, sur lesquels on peut décomposer les vecteurs. Dans ce cas, on peut décomposer une application $f : E \to E'$ en ses composantes. Soit $m = \dim(E')$, de façon à ce que, si $e'_1, \ldots, e'_m$ est une base de E', alors il existe m *fonctions composantes* $f^i : E \to \mathbb{R}$ telles que pour tout $x \in E$, $f(x) = \sum_{i=1}^{m} f^i(x)e'_i$. On a alors le même résultat intuitif que pour les suites, qui nous permet de restreindre beaucoup de nos analyses à l'études de fonctions de E dans $\mathbb{R}$.

Proposition 1.4. *Soit (E, N) et (E', N') deux espaces vectoriels normés (de dimension finie). L'application $f : E \to E'$ est continue en $a \in E$ si et seulement si chaque fonction composante $f^i : E \to \mathbb{R}$ de f pour $1 \leq i \leq \dim(E')$ est continue en a.*

Démonstration. En effet en dimension finie toutes les normes sont équivalentes donc on peut prendre la norme 1 sur E'. Cela signifie que la définition de la continuité en un point $a \in D$ se récrit :

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \text{ tel que } \forall\, x \in D,\ \text{si } N(x-a) < \delta \text{ alors } \|f(x) - f(a)\|_1 < \epsilon$$

Autrement dit :

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \text{ tel que } \forall\, x \in D,\ \text{si } N(x-a) < \delta \text{ alors } \sum_{i=1}^{m} |f^i(x) - f^i(a)| < \epsilon$$

De là on voit que l'inégalité de droite est satisfaite si et seulement si chaque terme $|f^i(x) - f^i(a)|$ peut être aussi petit que l'on veut dès qu'on est assez proche de a, c'est donc que chaque composante f^i est continue. □

En dimension finie, si $\dim(E) = n$ et $\dim(E') = m$ et si on a une base $e_1, \ldots, e_n$ de E et $e_1, \ldots, e_m$ une base de E', on peut décomposer une application linéaire en composantes :

$$\forall\, x = \sum_{j=1}^{n} x^j e_j \in E, \qquad f\left(\sum_{j=1}^{n} x^j e_j\right) = \sum_{j=1}^{n} x^j f(e_j) = \sum_{i=1}^{m}\sum_{j=1}^{n} x^j f^i_j e'_i \tag{1.1}$$

où on a noté $f^i_j = f^i(e_j) \in \mathbb{R}$. Nous voyons dans l'Equation (1.1) que l'application linéaire f définit une matrice M de taille $m \times n$ dont les coefficients sont $M_{ij} = f^i_j$. Attention à l'ordre des indices ! ! L'indice des colonnes est l'indice $1 \leq j \leq \dim(E)$ et l'indice des lignes est $1 \leq i \leq \dim(E')$. C'est comme cela qu'on montre que l'espace des applications linéaires de E dans E' est isomorphe à l'espace des matrices réelles de taille $m \times n$, c'est à dire $\mathcal{L}(E, E') \simeq \mathcal{M}_{m\times n}(\mathbb{R})$. Nous pourrons donc naviguer facilement entre les deux mondes (applications linéaires et matrices) car ce sont les mêmes objets (en dimension finie). Les applications linéaires satisfont la propriété suivante qu'on admet ici :

Proposition 1.5. *En dimension finie, toute application linéaire est continue.*

Si $\dim(E) = n$ et $\dim(E') = m$, alors $E \simeq \mathbb{R}^n$ et $E' \simeq \mathbb{R}^m$. Une application quelconque $f : E \to E'$ peut donc s'interpréter comme une application $f : \mathbb{R}^n \to \mathbb{R}^m$, ou plus généralement définies sur un sous-ensemble $D \subset \mathbb{R}^n$. Pour certaines valeurs de m et n nous avons des représentations graphiques de certaines classes de fonctions de plusieurs variables, que nous avons déjà étudiées ou que nous allons étudier dans ce cours :

1. si $n = 1, m = 1$, alors f une fonction réelle à une variable ;
2. si $n = 1, m > 1$, alors f est une courbe paramétrée dans $\mathbb{R}^m$
3. si $n = 2, m = 1$, alors f est une *fonction à deux variables.*

Nous allons largement étudier ce dernier cas dans le cours. On représenter cette fonction par son graphe, qui est une surface dans $\mathbb{R}^3$. Plus généralement, on appelle

1. *fonctions scalaires* les fonctions $f : \mathbb{R}^n \to \mathbb{R}$;
2. *fonctions vectorielles* les fonctions $f : \mathbb{R}^n \to \mathbb{R}^m$.

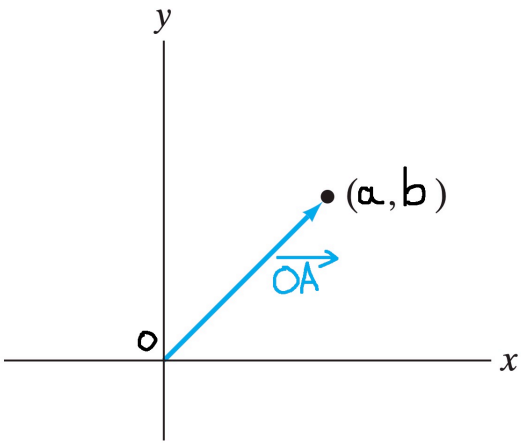
Nous allons maintenant restreindre notre focus aux fonctions de $\mathbb{R}^2$ dans $\mathbb{R}$, car nous avons la possibilité de dessiner ces fonctions. En particulier, nous sommes intéressés par décrire la géométrie de leur graphe, qui sont des surfaces de $\mathbb{R}^3$. Dans ce cadre, nous avons les comparaisons suivantes en dessinant deux fonctions :

Fonction	$\mathbb{R} \to \mathbb{R}$	$\mathbb{R}^2 \to \mathbb{R}$
Domaine de définition	union d'intervalles ou de segments	domaine ouvert ou compact
Continuité	$\lim_{x \to a} \lvert f(x) - f(a) \rvert = 0$	$\lim_{(x,y) \to \vec{a}} \lVert f(x,y) - f(\vec{a}) \rVert = 0$
Graphe	courbe dans $\mathbb{R}^2$	surface dans $\mathbb{R}^3$
Différentiabilité	$f(a+h) = f(a) + hf'(a) + o(\lvert h \rvert)$	$f(\vec{a} + (h,k)) = f(\vec{a})$ $+h\frac{\partial f}{\partial x}(\vec{a}) + k\frac{\partial f}{\partial y}(\vec{a}) + o(\lVert (h,k) \rVert)$
Tangence au graphe	droite tangente en a	plan tangent en $\vec{a}$
Intégrabilité	$\int_I f(t)dt$	$\iint_D f(s,t)dsdt \stackrel{?}{=} \iint_D f(s,t)dtds$
Intégrale	surface sous le graphe	volume sous le graphe

Dans le chapitre qui suit, nous allons étudier en détail les fonctions à deux variables, mais donner les résultats généraux sur les fonctions vectorielles de $\mathbb{R}^n$ dans $\mathbb{R}^m$.

1.1 Introduction au calcul différentiel

Nous rappelons quelques notions de topologie pour $\mathbb{R}^2$ (valide pour $\mathbb{R}^n$). Le zéro est noté O. Nous pouvons voir les éléments de $\mathbb{R}^2$ soit comme des points A soit comme leur vecteur associé $\overrightarrow{OA}$. Dans $\mathbb{R}^2$ nous pouvons choisir un vecteur particulier – qu'on dénote $\vec{i}$ – et son image sous une rotation de 90° dans le sens direct – qu'on note $\vec{j}$. Le triplet $(O, \vec{i}, \vec{j})$ est appelé une *base*. Les coordonnées d'un point A de $\mathbb{R}^2$ – coordonnées qu'on note (a, b) – correspondent aux composantes du vecteur $\overrightarrow{OA}$ dans la base $(O, \vec{i}, \vec{j})$: $\overrightarrow{OA} = a\vec{i} + b\vec{j}$. Un point A est représenté par ses coordonnées (a, b) dans cette base.



Par convention, le vecteur $\vec{i}$ (resp. $\vec{j}$) est le vecteur reliant l'origine $(0, 0)$ au point $(1, 0)$ (resp. $(0, 1)$). La notation pour un vecteur en algèbre linéaire est une représentation par un vecteur colonne :

$$\overrightarrow{OA} = \begin{pmatrix} a \\ b \end{pmatrix}, \qquad \vec{i} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{et} \qquad \vec{j} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Etant donné deux points $A = (a, b)$ et $A' = (a', b')$, le vecteur $\overrightarrow{AA'}$ a pour coordonnées $\begin{pmatrix} a'-a \\ b'-b \end{pmatrix}$, en effet cela peut se voir directement comme suit :

$$\overrightarrow{AA'} = \overrightarrow{OA'} - \overrightarrow{OA} = a'\vec{i} + b'\vec{j} - (a\vec{i} + b\vec{j}) = (a - a')\vec{i} + (b - b')\vec{j}$$

Dans tout le cours, on pourra bien souvent identifier le point $A = (a, b)$ avec le vecteur $\overrightarrow{OA} = \binom{a}{b}$. Le point A existe indépendamment de la base choisie mais ses coordonnées en dépendent.

On peut prendre une autre base de $\mathbb{R}^2$, par exemple $(\vec{i}', \vec{j}')$, obtenue à partir de la base $(\vec{i}, \vec{j})$ à l'aide d'une matrice de changement de base (inversible) :

$$\vec{i}' = P \times \vec{i} \quad \text{et} \quad \vec{j}' = P \times \vec{j}$$

Si on a un vecteur $\vec{v}$, le vecteur existe indépendamment du choix de la base mais ses composantes dépendent du choix de la base ! ! Par exemple dans le première base on a $\vec{v} = h\vec{i} + k\vec{j}$, et dans la deuxième base on a $\vec{v} = h'\vec{i}' + k'\vec{j}'$, ce qui donne

$$\vec{v}' = P \times \vec{v} \quad \text{c'est à dire} \quad \begin{pmatrix} h' \\ k' \end{pmatrix} = P \times \begin{pmatrix} h \\ k \end{pmatrix}$$

Si on note $P = \left(\begin{smallmatrix} r & s \\ p & q \end{smallmatrix}\right)$ on a $h' = rh + sk$ et $k' = ph + qk$. De même, les coordonnées d'un point A, qui sont (a, b) par rapport à la base canonique, deviennent $(a', b') = (ra + sb, pa + qb)$ dans la nouvelle base. Le point existe indépendamment de la base, mais ses coordonnées dépendent de la base choisie.

Une application $\alpha : \mathbb{R}^2 \to \mathbb{R}$ est dite *linéaire* si pour tout $\lambda \in \mathbb{R}$ et pour tout vecteur $\vec{v} \in \mathbb{R}^2$, on a :

$$\alpha(\lambda\vec{v}) = \lambda\alpha(\vec{v})$$

L'espace vectoriel de toutes les applications linéaires sur $\mathbb{R}^2$ forme un espace vectoriel qu'on appelle *espace dual* et qu'on note $(\mathbb{R}^2)^*$. Il est de dimension 2 et son zéro est l'application nulle. Ses éléments sont parfois appelés *covecteurs*. On peut donc représenter toute forme linéaire (covecteur) α par deux composantes, comme un vecteur. Il existe deux formes linéaires particulières, i^* et j^*, qui sont duales des vecteurs de base $\vec{i}$ et $\vec{j}$, dans le sens où :

$$\forall\ \vec{v} = \begin{pmatrix} h \\ k \end{pmatrix} \qquad i^*(\vec{v}) = h \quad \text{et} \quad j^*(\vec{v}) = k$$

Donc en particulier, on a

$$i^*(\vec{i}) = 1, \quad i^*(\vec{j}) = 0, \quad j^*(\vec{i}) = 0, \quad j^*(\vec{j}) = 1 \tag{1.2}$$

Les deux formes linéaires i^* et j^* forme une base de $(\mathbb{R}^2)^*$, c'est à dire qu'on peut projeter toute forme linéaire α sur ces vecteurs de base, et ce sont les coordonnées de α dans $(\mathbb{R}^2)^*$: $\alpha = ri^* + sj^*$. Par convention, une forme linéaire est représentée par un vecteur ligne, donc $\alpha = (r\ s)$ (ne pas confondre avec les coordonnées des points/vecteurs de $\mathbb{R}^2$!). En particulier, une telle application linéaire, évaluée sur un vecteur $\vec{v} = (h, k)$ revient à faire le produit matriciel :

$$\alpha(\vec{v}) = \begin{pmatrix} r & s \end{pmatrix} \cdot \begin{pmatrix} h \\ k \end{pmatrix} = rh + sk$$

En outre, dans ce cadre la transposée d'un vecteur donne une application linéaire, et inversement autrement dit :

$$\vec{v}^t = \begin{pmatrix} h & k \end{pmatrix} \qquad \text{et} \qquad \alpha^t = \begin{pmatrix} r \\ s \end{pmatrix}$$

Exemple 1.6. Montrer que (i^*, j^*) est une famille libre et génératrice, c'est à dire une base du dual $(\mathbb{R}^2)^*$, en l'évaluant sur la base $(\vec{i}, \vec{j})$ et en utilisant les identités (1.2).

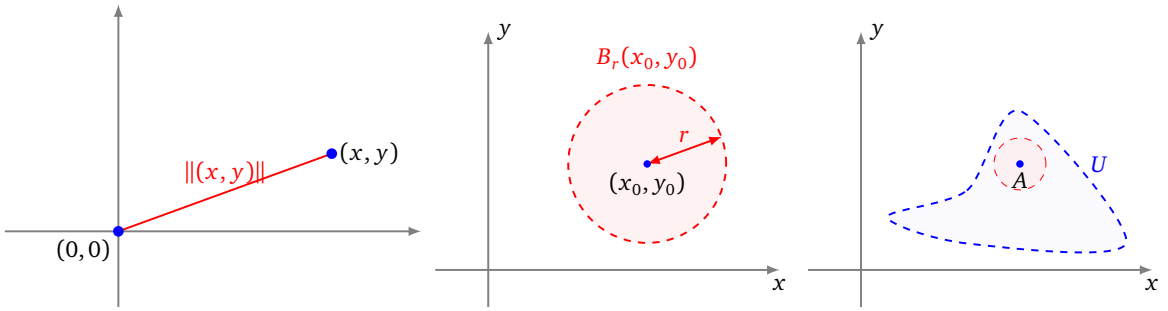


FIGURE 1 – A gauche, la norme euclidienne ; au centre, une boule ouverte de rayon r centrée au point (x_0, y_0) ; à droite un voisinage (ouvert, puisque la frontière est en pointillé) du point A.

L'espace $\mathbb{R}^2$ possède un produit scalaire, qui est linéaire en ses deux entrées :

$$\forall\, \vec{v} = \begin{pmatrix} h \\ k \end{pmatrix},\ \vec{v'} = \begin{pmatrix} h' \\ k' \end{pmatrix}, \qquad \langle \vec{v}, \vec{v'} \rangle = hh' + kk'$$

On observe, par la propriété de linéarité, que le produit scalaire $\langle \vec{v}, \cdot \rangle$ est une forme linéaire :

$$\begin{array}{rcl} \langle \vec{v}, \cdot \rangle : \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ \vec{v'} & \longmapsto & \langle \vec{v}, \vec{v'} \rangle \end{array}$$

La norme la plus pratique sur $\mathbb{R}^2$ est la norme euclidienne (c'est la norme 2), qui est induite par le produit scalaire (on n'écrit pas le petit 2 en bas à droite par commodité) :

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v} \rangle} = \sqrt{h^2 + k^2}$$

La norme permet de définir la distance entre deux points (c'est une fonction symmétrique) :

$$d(A, A') = \left\| \overrightarrow{AA'} \right\| = \sqrt{(a' - a)^2 + (b' - b)^2}$$

Une boule ouverte de centre A et de rayon $r > 0$ (par rapport à la norme 2) est un disque ouvert :

$$B_2(A, r) = \left\{ M \in \mathbb{R}^2 \ \text{ tel que } \ d(A, M) < r \right\}$$

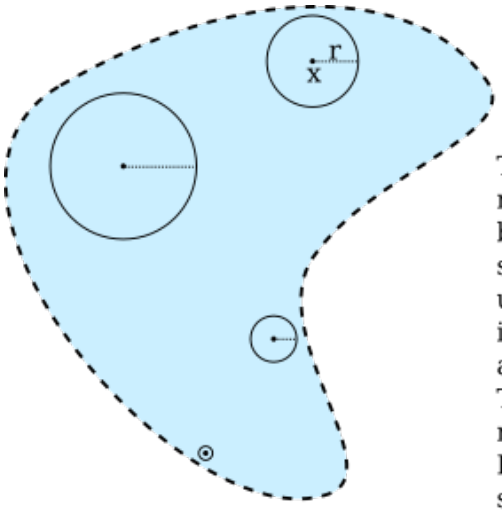
Définition 1.7. *Soit U une partie de $\mathbb{R}^2$. On dit que :*

- *U est un* voisinage *du point A si U contient une boule ouverte centrée en A (donc en particulier $A \in U$) ;*
- *U est un ouvert de $\mathbb{R}^2$ (par rapport à la norme 2) si U est un voisinage de chacun de ses points. C'est à dire, pour tout point M de U, U contient une boule ouverte centrée en M ;*
- *U est un fermé de $\mathbb{R}^2$ (par rapport à la norme 2) si son complémentaire $U^c = \mathbb{R}^2 \backslash U$ est ouvert ;*
- *U est compact si U est fermé borné (cela n'est valable qu'en dimension finie).*

Remarque 1.8. Nous rappelons qu'en dimension finie, toutes les normes sont équivalentes, c'est à dire que si une suite est convergente par rapport à une norme, elle l'est par rapport à toutes les autres normes. Un ensemble ouvert par rapport à une norme est donc ouvert par rapport à toutes les normes.

Soit P une partie de $\mathbb{R}^2$, on appelle *intérieur* de P le plus grand sous-ensemble ouvert de $\mathbb{R}^2$ contenu dans P – on le dénote $\mathring{P}$. On appelle *adhérence* de P le plus petit sous-ensemble fermé de $\mathbb{R}^2$ qui contient P – on le dénote $\overline{P}$. On appelle *frontière* de P le sous-ensemble formé de la différence $\partial P = \overline{P} \backslash \mathring{P}$.

Exemple 1.9. Une boule (disque) ouverte est un ouvert. Un rectangle $]a, b[\times]c, d[$ de $\mathbb{R}^2$ est un ouvert. Une boule fermée est un fermée, un rectangle $[a, b] \times [c, d]$ est un fermé. Le rectangle $[a, b[\times]c, d]$ n'est ni ouvert ni fermé.



The set shown in blue, not including the dashed boundary, is an open subset of $\mathbf{R}^2$, with its usual metric, because it includes an open ball around each of its points. The open ball of radius r about a point x is shown. For other points in the set, different radii might be used.

Remarque 1.10. L'ensemble des sous-ensembles ouverts de $\mathbb{R}^2$ est appelé la *topologie (métrique)* de $\mathbb{R}^2$ et est notée $\mathcal{T}(\mathbb{R}^2)$. Par convention on pose que l'ensemble vide $\emptyset$ est ouvert. En particulier, nous avons que :

$$\mathcal{T}(\mathbb{R}^2) \subset \mathcal{P}(\mathbb{R}^2)$$

C'est à dire que la topologie est un sous-ensemble de l'ensemble des parties de $\mathbb{R}^2$. Comme toutes les normes sont équivalentes en dimension finie, la topologie (métrique) ne dépend pas de la norme choisie.

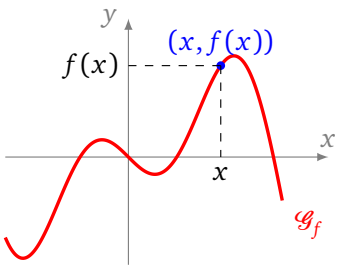
Le *domaine de définition* d'une fonction f est le plus grand sous-ensemble de $\mathbb{R}^2$ – noté D – tel que pour tout $(x, y) \in D$, $f(x, y)$ est bien défini (existe). Attention, il est possible d'étendre la fonction à certains points de la frontière de l'ensemble de définition si $f(x, y)$ tend vers une unique limite finie en ce point. L'*image* de f est le sous-ensemble de $\mathbb{R}$ défini par

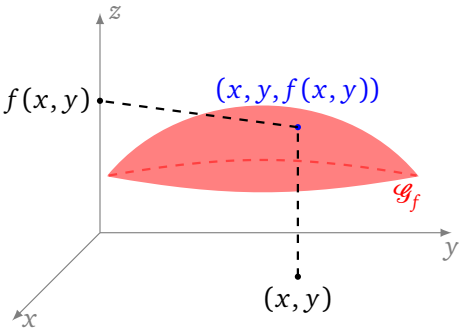
$$\begin{aligned} \mathrm{Im}(f) = f(D) &= \Big\{ z \in \mathbb{R} \ \text{ tel que } \ \exists\, (x, y) \ \in \mathbb{R}^2 \ \text{ tel que } \ z = f(x, y) \Big\} \\ &= \{ f(x, y) \ \in \mathbb{R} \ \text{ tel que } \ (x, y) \in D \} \end{aligned}$$

Le *graphe* de f est le sous-ensemble de $\mathbb{R}^3$ défini par

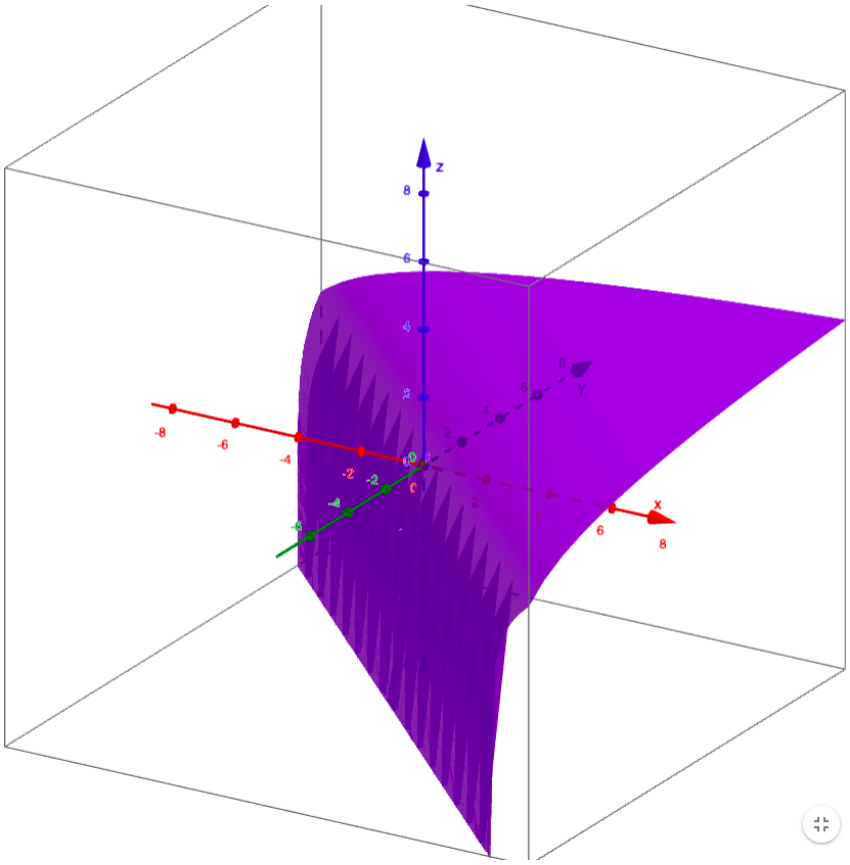
$$\mathrm{Gr}(f) = \Big\{ (x, y, z) \ \in \mathbb{R}^3 \ \text{ tel que } \ z = f(x, y) \Big\}$$

C'est donc une surface de $\mathbb{R}^3$ (tandis que pour une fonction $f : \mathbb{R} \to \mathbb{R}$ c'est une ligne de $\mathbb{R}^2$).



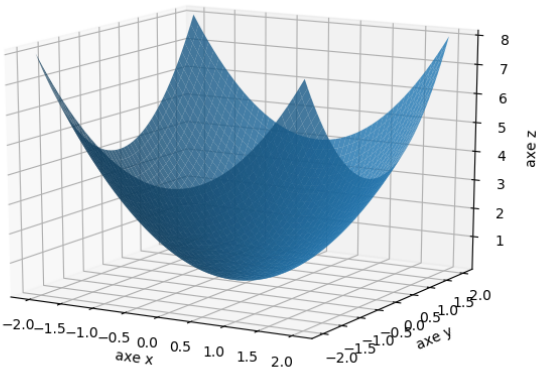


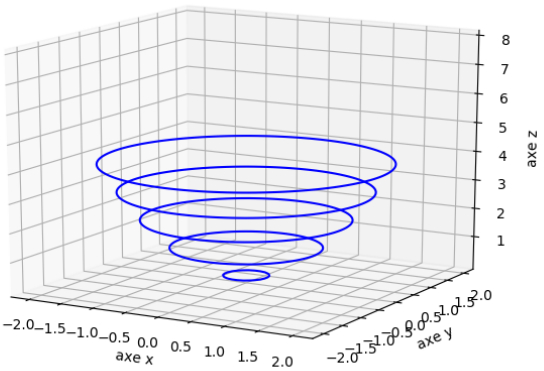
Exemple 1.11. Soit f une fonction à deux variables définie par $f(x,y) = \ln(1+x+y)$. Son domaine de définition est la partie de $\mathbb{R}^2$ placée strictement au dessus de la droite d'équation $y = -x-1$ (non-incluse) : $D = \{(x,y) \text{ tel que } y > -x-1\}$. C'est un ouvert. L'image de f est $\mathbb{R}$, en effet si $x = e^z$ et $y = 1$, alors $(x,y) \in D$ et on a $f(e^z, 1) = \ln(e^z) = z$. Le graphe de f c'est une nappe formée par translatant le long de l'antidiagonale le graphe de la fonction logarithme $\ln(1+u)$ définie le long de la diagonale $y = x$.

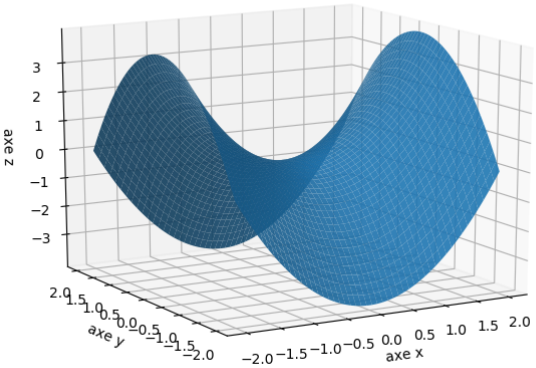


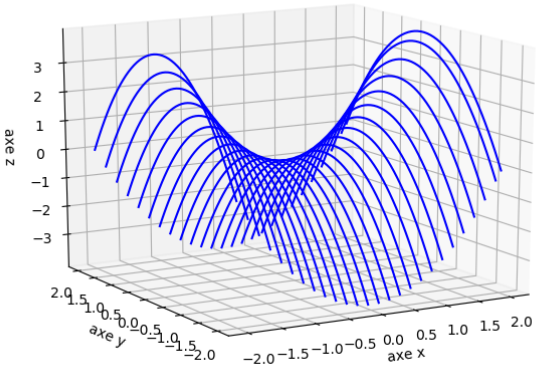
Exemple 1.12. Soit g une fonction à deux variables définie par $g(x,y) = \exp\left(\frac{x+y}{x^2-y}\right)$. Son domaine de définition est $D = \{(x,y) \text{ tel que } x^2 \neq y$ c'est donc $\mathbb{R}^2$ privé de la parabole $y = x^2$ (c'est donc un ouvert). Son image est $\mathbb{R}_+^*$ car si on évalue g sur la ligne diagonale d'équation $y = x$, on a $g(x,x) = \exp(\frac{2}{x-1})$ et on peut faire $x \to 1^{\pm}$ ce qui fait tendre l'exponentielle vers $\pm\infty$. Son graphe est un peu compliqué à dessiner (checker Geogebra).

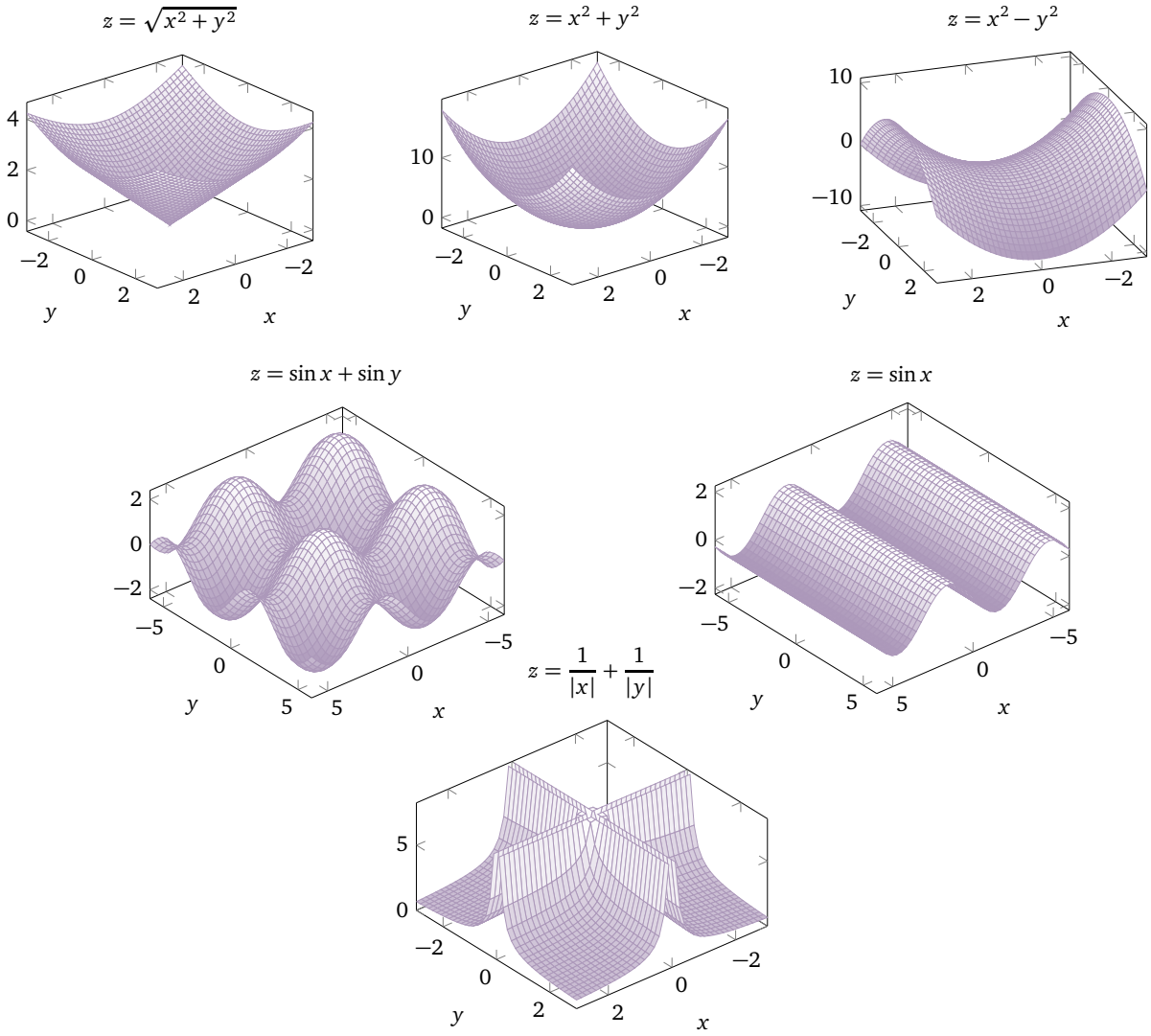
Exemple 1.13. La fonction $h(x,y) = x^2+y^2$ est définie sur le plan tout entier et ressemble à une parabole qu'on a fait tourner sur elle-même : c'est un paraboloïde elliptique (un peu comme un döner). Au contraire, la fonction $k(x,y) = x^2 - y^2$ ressemble à un chipster, c'est un paraboloïde hyperbolique.


axe z
axe x
axe y


axe z
axe x
axe y


axe z
axe y
axe x

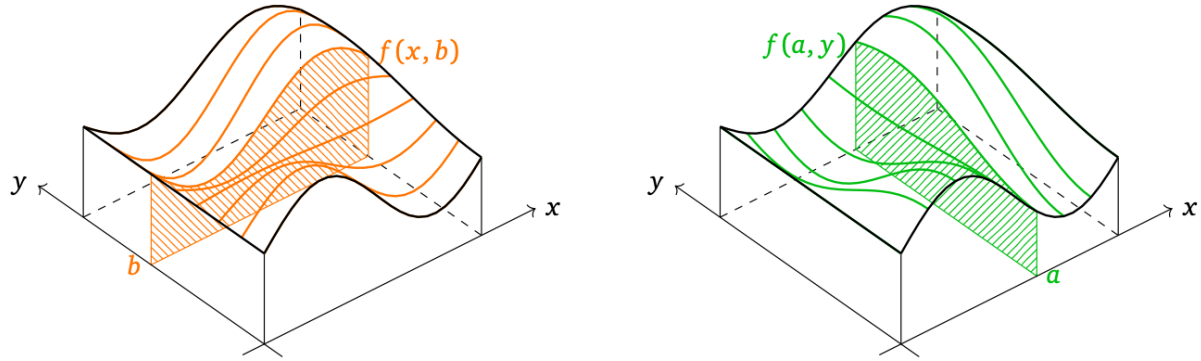

axe z
axe y
axe x



Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables, d'ensemble de définition $D \subset \mathbb{R}^2$. qu'une fonction à plusieurs variables. Soit $A = (a, b) \in D$ un point du domaine de définition, alors on peut poser :

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

pour toute paires de points $(x, b), (a, y) \in D$. Ce sont des fonctions à une variable : leur graphe est donc une courbe. La fonction f_1 est définie au dessus de la droite d'équation $y = b$ et donc son graphe est une courbe au dessus de cette droite, courbe qui est incluse dans le graphe de f (qui est une surface). La fonction f_2 est définie au dessus de la droite d'équation $x = a$ et donc son graphe est une courbe au dessus de cette droite, courbe qui est incluse dans le graphe de f. Ces deux courbes se croisent exactement au point $(a, b, f(a, b))$. Les graphes respectifs des deux fonctions définissent des chemins sur la surface de $\mathbb{R}^3$ représentée par le graphe de f. Le graphe de f_1 est l'intersection du graphe de f et du plan vertical d'équation $y = b$, tandis que le graphe de f_2 est l'intersection du graphe de f et du plan vertical d'équation $x = a$.



Notons que la définition des fonctions f_1 et f_2 dépend du choix des coordonnées, et donc de la base choisie, ici en l'occurrence on a pris la base canonique $(\vec{i}, \vec{j})$. Si on prend une autre base de $\mathbb{R}^2$, par exemple $(\vec{i}', \vec{j}')$, obtenue à partir de la base $(\vec{i}, \vec{j})$ à l'aide d'une matrice de changement de base (inversible) $P = \left(\begin{smallmatrix} r & s \\ p & q \end{smallmatrix}\right)$, la fonction $f : (x, y) \mapsto f(x, y)$ peut être réécrite dans la nouvelle base comme suit :

$$\overline{f}(w, z) = f(rx + sy, px + qy)$$

En posant $(a', b') = (ra + sb, pa + qb)$ qui sont les coordonnées du point A dans la nouvelle base, on obtient que les fonctions partielles $\overline{f}_1 : w \mapsto \overline{f}(w, b')$ et $\overline{f}_2 : z \mapsto \overline{f}(a', z)$ de la fonction $\overline{f}$ sont donc différentes des fonctions partielles f_1 et f_2.

Définition 1.14. Fonctions partielles en dimension finie. *Soit $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ une fonction et $A \in D$ un point du domaine de définition de f. Soit $e_1, \ldots, e_n$ une base de $\mathbb{R}^n$ et $(A^1, \ldots, A^n)$ les coordonnées du point A dans cette base, c'est à dire que $A = \sum_{j=1}^n A^j e_j$. Pour tout $1 \leq j \leq n$ on définit la j-ème* fonction partielle de f en A *comme la courbe paramétrée à valeurs dans $\mathbb{R}^m$ suivante :*

$$\begin{aligned} f_j : \ \mathbb{R} &\longrightarrow \mathbb{R}^m \\ t &\longmapsto f(A^1, \ldots, A^{j-1}, A^j + t, A^{j+1}, \ldots, A^n) \end{aligned}$$

pour tout t tel que $(A^1, \ldots, A^{j-1}, A^j + t, A^{j+1}, \ldots, A^n) \in D$.

Proposition 1.15. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction et $A \in D$. Si f est continue en A, alors toutes ses fonctions partielles de f en A (par rapport à n'importe quelle base donnée) sont continues en* 0.

Démonstration. En dimension finie toutes les normes sont équivalentes donc si la fonction f est continue en A par rapport à une norme donnée, elle l'est par rapport à toutes les autres normes. Prenons la norme 1 sur $\mathbb{R}^m$ et n'importe quelle norme N sur $\mathbb{R}^n$, ainsi qu'un entier $1 \leq j \leq n$. Soit $\epsilon > 0$, alors il existe $\delta > 0$ tel que pour tout point $M \in B_1(A, \delta)$, on a $N(f(M) - f(A)) < \epsilon$. En particulier, cela est vrai pour tout $M \in B_1(A, \delta)$ tel que $M^k = A^k$ dès que $k \neq j$, où $(M^1, \ldots, M^n)$ sont les coordonnées de M dans la base donnée. Pour tout $t \in]A^j - \delta, A^j + \delta[$ on pose $M_t = (A^1, \ldots, A^{j-1}, t + A^j, A^{j+1, \ldots, A^n} \in B_1(A, \delta)$. On a alors par la propriété de la norme 1 :

$$\|M_t - A\|_1 = \sum_{k=1}^{n} \left| M_t^k - A^k \right| = |t| < \delta$$

Ce qui nous dit que $M_t \in B_1(A, \delta)$, pour tout $t \in]A^j - \delta, A^j + \delta[$. Alors nous avons que $Nf(M_t) - f(A)) < \epsilon$, ce qui s'écrit en réalité par construction de M_t : $N(f_j(t) - f(A)) < \epsilon$. Autrement dit nous avons montré que :

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \text{ tel que } \forall\, t \in]A^j - \delta, A^j + \delta[,\ N(f_j(t) - f(A)) < \epsilon$$

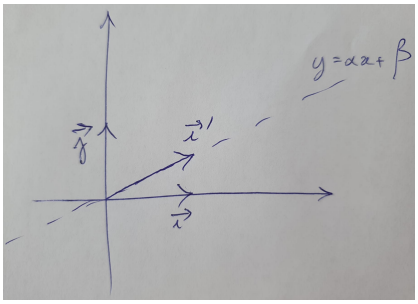
C'est la continuité de la j-ème fonction partielle de f en A. □

Exemple 1.16. La Proposition 1.15 n'a pas de réciproque. Prenons la fonction

$$\begin{aligned} f: \ \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto f(x,y) = \begin{cases} e^{-\frac{y^3}{(y-x^2)^2}} \text{ si } y \neq x^2 \\ 0 \text{ si } y = x^2 \text{ et } x \neq 0 \\ 1 \text{ si } x = y = 0 \end{cases} \end{aligned}$$

La fonction est continue sur le plan épointé : c'est à dire en tout point $(x,y) \neq (0,0)$. Un doute subsiste à l'origine. Prenons $(a,b) = (0,0)$. La fonction partielle f_1 en $(0,0)$ est donnée par $f_1(x) = f(x,0) = e^0 = 1$ qui tend donc vers $1 = f(0,0)$ lorsque x tend vers 0. La fonction partielle f_2 en $(0,0)$ est donnée par $f_2(y) = f(0,y) = e^{-\frac{y^3}{y^2}} = e^{-y}$ qui tend vers $1 = f(0,0)$ lorsque y tend vers 0. Donc les deux fonctions partielles f_1 et f_2 sont continue en 0.

Approchons nous de l'origine le long de la droite d'équation $y = \alpha x$, pour un $\alpha \in \mathbb{R}^*$ donné. Effectuons un changement de base en prenant comme nouveaux vecteurs de base $\vec{i'} = \vec{i} + \alpha\vec{j}$ et $\vec{j'} = \vec{j}$. Un point de coordonnée (x,y) dans la première base a pour coordonnées $(s,t) = (x, y-\alpha x)$ dans la deuxième base. Inversement, un point de coordonnées (s,t) dans la deuxième base a pour coordonnées $(x,y) = (s, t+\alpha s)$ dans la première base.



La fonction f prend alors une nouvelle expression dans la nouvelle base :

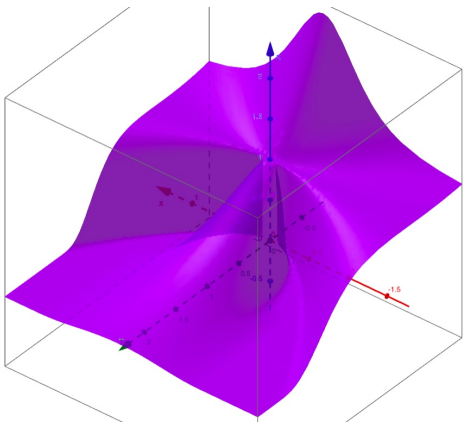
$$\begin{aligned} f_\alpha: \ \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (s,t) &\longmapsto f_\alpha(s,t) = \begin{cases} e^{-\frac{(t+\alpha s)^3}{(t+\alpha s - s^2)^2}} \text{ si } t \neq s^2 - \alpha s \\ 0 \text{ si } t = s^2 - \alpha s \text{ et } s \neq 0 \\ 1 \text{ si } s = t = 0 \end{cases} \end{aligned}$$

La fonction partielle par rapport à la première variable en $(0,0)$ est :

$$f_{\alpha,1}(s) = e^{-\frac{(\alpha s)^3}{(\alpha s - s^2)^2}} = e^{-\frac{\alpha s}{(1-\frac{s}{\alpha})^2}}$$

qui tend vers $1 = f(0,0)$ quand s tend vers 0. La deuxième fonction partielle coincide avec f_2 calculée précédemment.

Donc toutes les fonctions partielles sont continues en 0, quelle que soit la direction de la droite choisie pour tendre vers 0 (c'est à dire quelle que soit la base choisie la base choisie). Par contre, si on tend vers 0 en suivant la parabole $y = x^2$, on aura $f(x,x^2) = 0$ pour tout $x \neq 0$, et on voit que si on fait tendre x vers 0, la limite de $f(x,x^2)$ est nulle, mais la valeur de la fonction en $(0,0)$ est 1. La fonction n'est donc pas continue à l'origine, par contre, on ne pouvait pas le voir à partir des seules fonctions partielles. Il fallait tendre vers l'origine en suivant la parabole. C'est pourquoi l'utilisation des fonctions partielles est un peu limitée et seule la contraposée suivante est utile :



Proposition 1.17. Contraposée de la Proposition 1.15. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. S'il existe une base de $\mathbb{R}^n$ par rapport à laquelle une fonction partielle de f en A n'est pas continue en 0, alors f n'est pas continue en A.*

Exemple 1.18. En revenant au cas bidimensionnel, prenons la fonction $f : \mathbb{R}^2 \to \mathbb{R}$ définie par :

$$f(x,y) = \begin{cases} \frac{x^2+y^2}{\sqrt{x^4+y^4}} \text{ si } (x,y) \neq (0,0) \\ 1 \text{ si } (x,y) = (0,0) \end{cases}$$

Les fonctions partielles $f_1 : x \mapsto f(x,0)$ et $f_2 : y \mapsto f(0,y)$ sont continues en 0 car, par exemple, on a $f_1(x) = \frac{x^2+0}{\sqrt{x^4+0}} = \frac{x^2}{x^2} = 1$ pour tout x. Cependant, cela ne veut rien dire sur la continuité de f, au contraire, nous pouvons changer de base et définir les fonctions partielles par rapport à la nouvelle base, et voir si elles sont continues (il faut choisir la base intelligemment).

L'expression de la fonction f se fait à partir de la base orthonormale canonique $(\vec{i}, \vec{j})$, mais on peut choisir une autre base, par exemple $(\vec{i}, i + \vec{j})$. Dans cette base, les coordonnées sont données par un couple $(x-y, y)$ car en effet, si on a un vecteur $x\vec{i} + y\vec{j}$ on peut l'écrire dans la deuxième base par $(x-y)\vec{i} + y(\vec{i} + \vec{j})$. Donc un vecteur de coordonnées (w, z) dans la deuxième base s'exprime comme $(w+z, z)$ dans la première base :

$$x\vec{i} + y\vec{j} \longmapsto (x-y)\vec{i} + y(\vec{i} + \vec{j}) \qquad \text{et} \qquad w\vec{i} + z(\vec{i} + \vec{j}) \longmapsto (w+z)\vec{i} + z\vec{j}$$

Donc on peut récrire la fonction dans la seconde base comme :

$$\overline{f}(w,z) = \begin{cases} \frac{(w+z)^2+z^2}{\sqrt{(w+z)^4+z^4}} \text{ si } (x,y) \neq (0,0) \\ 1 \text{ si } (x,y) = (0,0) \end{cases}$$

Les fonctions f et $\overline{f}$ définissent la même fonction, mais dans deux bases différentes. Leur valeur en un point de $\mathbb{R}^2$ est indépendante de la base et identique.

On veut évaluer la continuité de la fonction en $(0,0)$, par rapport à voir si ses fonctions partielles sont continues. On a donc les deux fonctions partielles suivantes :

$$\overline{f}_1 : w \mapsto \overline{f}(w,0) = \frac{w^2}{\sqrt{w^4}} = 1 \quad \text{et} \quad \overline{f}_2 : z \mapsto \overline{f}(0,z) = \frac{2z^2}{\sqrt{2z^4}} = \sqrt{2}$$

La première fonction tend vers 1 quand w tend vers 0. La deuxième fonction tend vers $\sqrt{2}$ quand z tend vers 0. La première fonction partielle est continue en 0 mais la deuxième ne coincide pas avec $f(0,0)$ en 0, donc la fonction $\overline{f} = f$ n'est pas continue en $(0,0)$.

On aurait atteint le même résultat dans la base originale si on avait défini $\alpha(t) = (t, 0)$ et $\beta(t) = (t, t)$. Alors $f \circ \alpha(t) = \frac{t^2}{\sqrt{t^4}}$ et $f \circ \beta(t) = \frac{2t^2}{\sqrt{2t^4}}$. Cela revient à suivre le même chemin qu'on a suivi juste au dessus. Donc plus généralement, pour montrer qu'une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ n'est pas continue au point $A = (a, b)$, il *suffit* de trouver deux fonctions $u, v : \mathbb{R} \to \mathbb{R}$ telles que $(u(0), v(0)) = A$ et telle que la courbe paramétrée $t \mapsto f(u(t), v(t))$ ne tende pas vers $f(A)$ lorsque t tend vers 0.

Exemple 1.19. Autre exemple de fonction non continue en dimension 2. En effet prenons :

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

et supposons qu'on approche de l'origine par la direction $(x, \alpha x)$, c'est à dire par la droite de pente α. On a alors $f(x, \alpha x) = \frac{\alpha}{1+\alpha^2}$ qui est constante. On peut réexprimer cette fonction avec

des coordonnées polaires. En effet nous avons une bijection :

$$\begin{array}{ccc} \mathbb{R}_+^* \times]-\pi,\pi[& \longrightarrow & \mathbb{R}^2 \backslash \{(x,y) \text{ t.q. } x \leq 0 \text{ et } y = 0\} \\ (r,\theta) & \longmapsto & (x,y) = (r\cos(\theta), r\sin(\theta)) \end{array}$$

La fonction réciproque nous donne (r,θ) en fonction de (x,y), avec

$$\begin{array}{ccc} \mathbb{R}^2 \backslash \{(x,y) \text{ t.q. } x \leq 0 \text{ et } y = 0\} & \longrightarrow & \mathbb{R}_+^* \times]-\pi,\pi[\\ (x,y) & \longmapsto & (r,\theta) = \left(\sqrt{x^2+y^2}, 2\arctan\left(\frac{y}{x+\sqrt{x^2+y^2}} \right) \right) \end{array}$$

où l'arctangente est la fonction réciproque de la tangente. La fonction ci dessus se lit maintenant $f(r,\theta) = \frac{\sin(2\theta)}{2}$.

Remarque 1.20. Pour résumer, l'idée de la continuité pour les fonctions à deux variables c'est qu'une fonction est continue en (a,b) si et seulement si on tend vers la valeur $f(a,b)$ lorsqu'on s'approche du point (a,b) selon n'importe quelle droite ou courbe. Regarder les droites ne suffit pas à dire si une fonction est continue comme on a vu dans l'Exemple 1.16, mais ça suffit à dire si une fonction n'est pas continue comme on la vue dans la Contraposée 1.17.

L'intérêt des fonctions partielles et de définir des *dérivées partielles*. En effet, rappelons que dans $\mathbb{R}$, la définition du nombre dérivé d'une fonction f en un point a fait intervenir le taux d'accroissement $\frac{f(x)-f(a)}{x-a}$. Le calcul de la dérivée implique donc nécessairement de pouvoir diviser par $x-a$. Mais dans $\mathbb{R}^n$ ça n'a pas de sens car la division par un vecteur n'est pas définie. Comment alors généraliser la dérivée d'une fonction d'une variable à toutes les fonctions vectorielles ? Nous allons introduire pour cela la notion de *différentiabilité*.

Définition 1.21. Dérivées partielles en dimension finie. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. Soit $A \in \mathring{D}$ un point de l'intérieur du domaine de définition de f. Pour tout $1 \leq j \leq n$, on appelle* dérivée partielle de f en A par rapport à x^j *– et on note $\frac{\partial f}{\partial x^j}(A)$ le nombre dérivé – si il existe – de la j-ème fonction partielle de f en A au temps $t = 0$, c'est à dire :*

$$\frac{\partial f}{\partial x^j}(A) = \lim_{t \to 0} \frac{f(A^1, \ldots, A^{j-1}, A^j + t, A^{j+1}, \ldots, A^n) - f(A)}{t}$$

Si en tous points A de $\mathring{D}$ la fonction f admet une dérivée partielle par rapport à la j-ème variable en A, alors on peut définir une fonction "dérivée partielle par rapport à x^j" :

$$\begin{array}{rccc} \frac{\partial f}{\partial x^j} : & \mathring{D} & \longrightarrow & \mathbb{R}^m \\ & A & \longmapsto & \frac{\partial f}{\partial x^j}(A) \end{array}$$

On dit que la fonction f est de classe $\mathcal{C}^1$ *si toutes ses fonctions dérivées partielles $\frac{\partial f}{\partial x^j}$ pour $1 \leq j \leq n$ sont continues sur $\mathring{D}$.*

Exemple 1.22. Calculons les dérivées partielles de la fonction $f(x,y,z) = yz^x$. Le domaine de définition de cette fonction est donné par tous les points (x,y,z) tels que $e^{x\ln(z)}$ est défini, c'est à dire pour $z > 0$. Donc $D \subset \mathbb{R}^3$ est le demi-espace ouvert supérieur. Soit $u = (a,b,c) \in D$ donc $c > 0$. Prenons la dérivée par rapport à la première variable, la fonction partielle est $x \mapsto be^{x\ln(c)}$ donc la dérivée est $\frac{\partial f}{\partial x}(a,b,c) = \ln(c)be^{x\ln(c)}$. Et quand on dérive par rapport à la deuxième variable, on a $\frac{\partial f}{\partial y}(a,b,c) = c^a$.

Le choix d'une base est arbitraire pour la définition des fonctions partielles, et donc pour la définition des dérivées partielles, et cela n'est pas idéal en mathématiques. En effet, les dérivées partielles peuvent exister dans une base, mais pas dans une autre, comme par exemple dans l'Exemple 1.18, les dérivées partielles de f dans la base canonique $(\vec{i}, \vec{j})$ existent, mais pas celles dans la base $(\vec{i}, \vec{i}+\vec{j})$ car la fonction partielle $\overline{f}_2$ n'est pas continue en 0 dans cette base donc certainement pas dérivable. Nous avons donc très logiquement la notion suivante, qui permet de se séparer de l'arbitrarité du choix de base.

Définition 1.23. Dérivée directionnelle en dimension finie. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. Soit $A \in \mathring{D}$ un point de l'intérieur du domaine de définition de f, et $v \in \mathbb{R}^n$ (alors pour tout t assez petit, on a $A + tv \in \mathring{D}$ car c'est un ouvert). La* dérivée directionnelle de f au point $A \in D$ par rapport au vecteur v *est la limite suivante, si elle existe :*

$$\lim_{t \to 0} \frac{f(A+tv) - f(A)}{t}$$

On la note $D_v f(A)$ ou $df_A(v)$.

Remarque 1.24. De cette définition, on déduit que, dans une base donnée, la dérivée directionnelle dans la direction e_j est la dérivée partielle par rapport à la j-ème variable.

Regardons en détail ce qu'il se passe avec les fonctions de deux variables. Les dérivées partielles d'une fonction f par rapport aux variables x et y sont les dérivées respectives des fonctions partielles f_1 et f_2 au point $A = (a, b)$, ce qui s'écrit d'après le formalisme général :

$$\frac{\partial f}{\partial x}(a,b) = \lim_{h \to 0} \frac{f(a+h,b) - f(a,b)}{h} \qquad \text{et} \qquad \frac{\partial f}{\partial y}(a,b) = \lim_{h \to 0} \frac{f(a,b+h) - f(a,b)}{h}$$

ou de façon équivalente :

$$\frac{\partial f}{\partial x}(a,b) = f_1'(a) = \lim_{x \to a} \frac{f(x,b) - f(a,b)}{x-a} \qquad \text{et} \qquad \frac{\partial f}{\partial y}(a,b) = f_2'(b) = \lim_{y \to b} \frac{f(a,y) - f(a,b)}{y-b}$$

Si on voit le plan d'équation $y = b$ (resp. $x = a$) en coupe, et comme on connait la signification du nombre dérivé, on en déduit que $\frac{\partial f}{\partial x}(a,b)$ (resp. $\frac{\partial f}{\partial y}(a,b)$) est la pente de la tangente à la courbe paramétrée $x \mapsto f(x,b)$ (resp. $y \mapsto f(a,y)$) au point $A = (a,b)$.

Plus généralement, comme en dimension n, on peut définir des fonctions partielles pour tout choix de direction donnée par un vecteur $\vec{v} = \binom{h}{k}$. En effet, si $A = (a,b)$ alors on a une fonction partielle dans la direction $\vec{v}$:

$$\begin{aligned} f_{\vec{v}} : \mathbb{R} &\longrightarrow \mathbb{R} \\ t &\longmapsto f(A + t\vec{v}) = f(a+th, b+tk) \end{aligned}$$

Ce qui fait qu'on a $f_{\vec{i}}(t) = f_1(a+t)$ et $f_{\vec{j}}(t) = f_2(b+t)$. Le graphe de la fonction partielle dans la direction $\vec{v}$ est une courbe paramétrée, incluse dans le graphe de f (une surface), voir le dessin plus bas. On peut dériver la fonction $f_{\vec{v}}$ par rapport à t et cela donne la *dérivée directionnelle* de f au point $A = (a,b)$ selon le vecteur $\vec{v} = \binom{h}{k}$.

Définition 1.25. *Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction, (a,b) un point et $\vec{v} = \binom{h}{k}$ un vecteur. La* dérivée directionnelle *de f au point (a,b) selon le vecteur $\vec{v}$ est la dérivée de la fonction partielle dans la direction $\vec{v}$ au temps 0, c'est à dire :*

$$D_{\vec{v}} f(a,b) = \lim_{t \to 0} \frac{f((a,b) + t\vec{v}) - f(a,b)}{t} = \lim_{t \to 0} \frac{f(a+th, b+tk) - f(a,b)}{t}$$

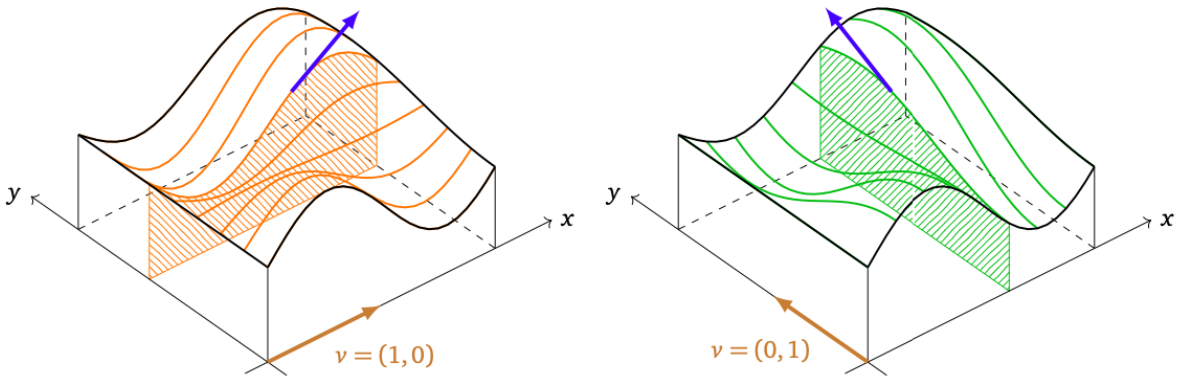
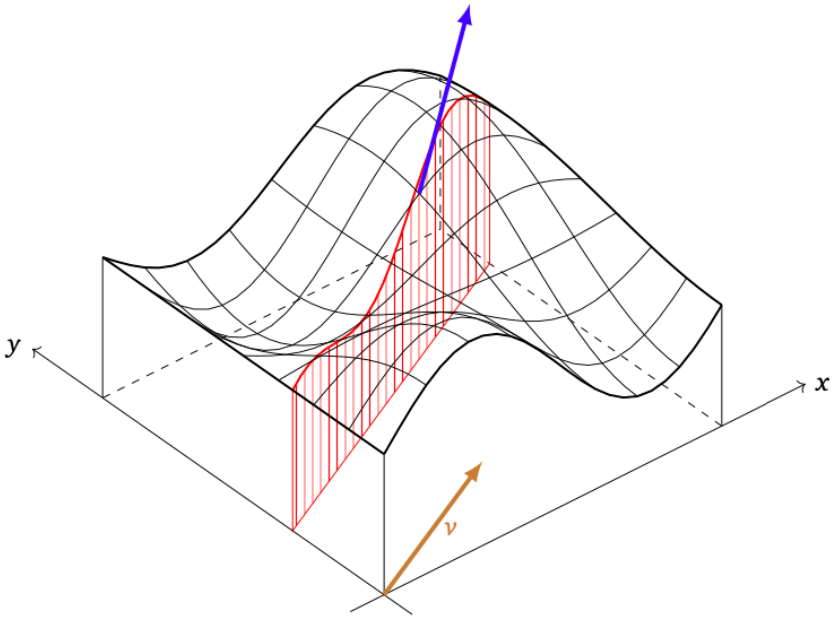


FIGURE 2 – Le vecteurs orange à gauche (resp. à droite) est le vecteur de base $\vec{i}$ (resp. $\vec{j}$), et le vecteur bleu est le vecteur directeur canonique de la tangente au graphe de la fonction partielle f_1 (resp. f_2) au point $(a, b, f(a, b))$, c'est à dire le vecteur $\begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(a,b) \end{pmatrix}$ (resp. $\begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$).

On observe que si $\vec{v} = \vec{i}$ (resp. $\vec{j}$), alors la dérivée directionnelle selon le vecteur $\vec{v}$ est la dérivée partielle selon x (resp. y) :

$$D_{\vec{i}} f(a, b) = \lim_{t \to 0} \frac{f((a, b) + t\vec{i}) - f(a, b)}{t} = \lim_{t \to 0} \frac{f(a + t, b) - f(a, b)}{t} = \frac{\partial f}{\partial x}(a, b)$$

Graphiquement, lorsque $\vec{v}$ est un vecteur *unitaire*, c'est à dire lorsque $||\vec{v}||_2 = 1$, alors la dérivée directionnelle dans la direction $\vec{v}$ indique la pente au point $A = (a, b)$ de la tangente au graphe de la fonction partielle directionnelle $f_{\vec{v}}$ dans la direction du vecteur $\vec{v}$.



Exemple 1.26. Reprenons l'Exemple 1.16 avec cette approche. Soit $(a, b) = (0, 0)$ et $\vec{v} = \begin{pmatrix} h \\ k \end{pmatrix}$ un vecteur. On peut donc écrire la fonction partielle $f_{\vec{v}}$ construite à partir de la fonction de

l'Exemple 1.16 (quand t est assez petit) :

$$f\big((0,0)+t\vec{v}\big) = e^{-\frac{(tk)^3}{(tk-(th)^2)^2}} = e^{-\frac{(tk)^3}{t^2k^2(1-t\frac{h^2}{k})^2}} = e^{-\frac{tk}{1-t\frac{h^2}{k})^2}} \underset{t\to 0}{\longrightarrow} e^0 = 1$$

On retrouve que la fonction f est continue si on s'approche de l'origine en ligne droite.

1.2 Différentielle et points critiques

Soit $A=(a,b)$ un point, $\vec{v}$ un vecteur et $f:\mathbb{R}^2\to\mathbb{R}$ une fonction. Soit $\lambda\neq 0$ et soit $\vec{w}=\lambda\vec{v}$. Reprenons la définition de la dérivée directionnelle :

$$\begin{aligned} D_{\vec{w}}f(a,b) &= \lim_{t\to 0}\frac{f(A+t\vec{w})-f(A)}{t} = \lim_{t\to 0}\frac{f(A+t\lambda\vec{v})-f(A)}{t} \\ &= \lim_{t\to 0}\lambda\frac{f(A+t\lambda\vec{v})-f(A)}{t\lambda} = \lambda\left(\lim_{s\to 0}\frac{f(A+s\vec{v})-f(A)}{s}\right) = \lambda D_{\vec{v}}f(a,b) \end{aligned}$$

Nous voyons donc que si la dérivée directionnelle existe dans une direction $\vec{v}$ donnée en un point $A=(a,b)$ donné, alors l'opérateur $\lambda\mapsto D_{\lambda\vec{v}}f(a,b)$ est une application linéaire de $\mathbb{R}$ dans $\mathbb{R}$, c'est à dire que $D_{\lambda\vec{v}}f(a,b)=\lambda D_{\vec{v}}f(a,b)$.

On peut se demander si nous avons la linéarité selon l'addition, c'est à dire que si $\vec{v},\vec{w}$ sont deux vecteurs quelconques de $\mathbb{R}^2$, a-t-on $D_{\vec{v}+\vec{w}}f(a,b) = D_{\vec{v}}f(a,b)+D_{\vec{w}}f(a,b)$? ? La réponse est a priori non. En effet rappelons qu'il existe des fonctions qui n'admettent pas de dérivée directionnelle dans certaines directions. Plus précisément, il n'est pas vrai que si f admet des dérivées partielles selon les directions $\vec{i}$ et $\vec{j}$, alors elle admet une dérivée directionnelle selon toutes les directions $\vec{v}=\binom{h}{k}=h\vec{i}+k\vec{j}$. En effet dans l'Exemple 1.18 on a vu une fonction qui n'est pas continue selon la direction $\vec{v}=\vec{i}+\vec{j}$ et donc elle n'est certainement pas dérivable dans cette direction et donc n'admet pas de dérivée directionnelle selon cette direction, même si elle en admet une dans les directions canoniques $\vec{i}$ et $\vec{j}$.

Cependant, le problème n'est pas relié à la continuité, car il existe des fonctions f qui ne sont pas continues en un point (a,b), mais dont toutes les dérivées directionnelles existent en ce point, et elles satisfont la condition de linéarité totale :

$$\forall\,\lambda,\mu\in\mathbb{R},\ \forall\,\vec{v},\vec{w}\in\mathbb{R}^2 \qquad D_{\lambda\vec{v}+\mu\vec{w}}f(a,b)=\lambda D_{\vec{v}}f(a,b)+\mu D_{\vec{w}}f(a,b) \tag{1.3}$$

En effet, dans l'Exemple 1.16, les fonctions partielles sont continues et dérivables dans toutes les directions et on a :

$$D_{\vec{i}+\alpha\vec{j}}f(0,0)=-\alpha \qquad \text{et} \qquad D_{\vec{j}}f(0,0)=-1$$

Donc en particulier $D_{\vec{i}}f(0,0)=0$. Nous voyons que pour tout vecteur $\vec{v}=h\vec{i}+k\vec{j}$ tel que $h\neq 0$, nous avons

$$D_{h\vec{i}+k\vec{j}}f(0,0)=D_{h(\vec{i}+\frac{k}{h}\vec{j})}f(0,0)=hD_{\vec{i}+\frac{k}{h}\vec{j}}f(0,0)=h\left(-\frac{k}{h}\right)=-k=hD_{\vec{i}}f(0,0)+kD_{\vec{j}}f(0,0)$$

D'autre part, si $h=0$ alors nous avons par linéarité $D_{k\vec{j}}f(0,0)=kD_{\vec{j}}f(0,0)$. Cela montre que les dérivées directionnelles de la fonction de l'exercice 1.16 satisfont la condition de linéarité (1.3) en $(0,0)$. Cependant elle n'est pas continue en $(0,0)$ par rapport aux vecteurs de base $\vec{i}$ et $\vec{j}$ (et donc par rapport à toute base). Cela montre que les dérivées directionnelles ne sont *pas* la généralisation correcte de la dérivée d'une fonction réelle. Car nous aimerions avoir un résultat du type "si une fonction à deux variable est dérivable, alors elle est continue", mais avec

l'Exemple 1.16 nous voyons qu'une fonction à deux variables peut être dérivable dans toutes les directions, sans être pour autant continue.

Il existe aussi des fonctions à deux variables qui sont continues, qui admettent des dérivées dans toutes les directions, mais qui ne satisfont pas la condition de linéarité totale (1.3). Prenons l'Exemple 1.2. La fonction f est continue et elle admet une dérivée directionnelle dans toutes les directions. En effet, pour tout vecteur $\vec{v} = h\vec{i} + k\vec{j}$ tel que $h \neq 0$, nous avons

$$\begin{aligned} D_{\vec{v}} f_{(0,0)} = \lim_{t\to 0} \frac{f((0,0)+t\vec{v}) - f(0,0)}{t} &= \lim_{t\to 0} \frac{f(th,tk)-0}{t} \\ &= \lim_{t\to 0} \frac{t^2k^3}{\sqrt{t^2h^2+t^4k^4}} = \lim_{t\to 0} \frac{tk^3}{\sqrt{1+t^2\frac{k^4}{h^2}}} = 0 \end{aligned}$$

Lorsque $h = 0$ c'est à dire si $\vec{v} = k\vec{j}$, nous avons :

$$D_{k\vec{j}} f_{(0,0)} = \lim_{t\to 0} \frac{f(0,tk)-0}{t} = \lim_{t\to 0} \frac{t^2k^3}{\sqrt{t^4k^4}} = k$$

Autrement dit les dérivées directionnelles sont toutes nulles sauf si $\vec{v}$ est vertical, et dans ce cas particulier nous avons $D_{\vec{j}} f_{(0,0)} = 1$. On peut vérifier facilement que l'identité de linéarité (1.3) n'est pas satisfaite. La fonction est pourtant continue. Ceci montre qu'une fonction continue, dont les dérivées partielles existent dans toutes les directions, ne satisfait pas forcément la condition de linéarité totale.

Pour trouver la notion correcte à 2 variables (ou plus) qui généralise la notion de dérivée, nous ne pouvons pas compter sur la notion de dérivée directionnelle. Nous devons généraliser la dérivée en prenant une définition plus profonde. Rappelons que nous avons le résultat suivant pour les fonctions réelles dérivables, qui caractérise la dérivée d'une fonction :

Proposition 1.27. Rappel de la dérivabilité d'une fonction réelle. *La fonction réelle à une variable $f : D \to \mathbb{R}$ est dérivable en $a \in D$ si et seulement si il existe un nombre réel α tel que :*

$$f(x) = f(a) + \alpha(x-a) + o(|x-a|)$$

Dans ce cas, le nombre α est égal à $f'(a)$.

L'avantage de cette définition/proposition est que nous ne divisions pas par un nombre, donc cette propriété peut vraisemblablement se généraliser au cas à plusieurs variables (rappelons qu'avec plusieurs variables nous ne pouvons pas diviser par un vecteur). Nous avons effectivement les deux définitions suivantes équivalentes :

Définition 1.28. Différentiabilité en dimension finie. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. Soit $A \in \mathring{D}$ un point de l'intérieur du domaine de définition de f. La fonction $f : D \to \mathbb{R}^m$ est differentiable en A s'il existe une application linéaire (donc continue) $\alpha : \mathbb{R}^n \to \mathbb{R}^m$ telle que pour tout M assez proche de A :*

$$f(M) = f(A) + \alpha\left(\overrightarrow{AM}\right) + o\left(\left\|\overrightarrow{AM}\right\|\right)$$

ou bien de manière équivalente :

$$\lim_{M\to A} \frac{\left\|f(M) - f(A) - \alpha\left(\overrightarrow{AM}\right)\right\|}{\left\|\overrightarrow{AM}\right\|} = 0$$

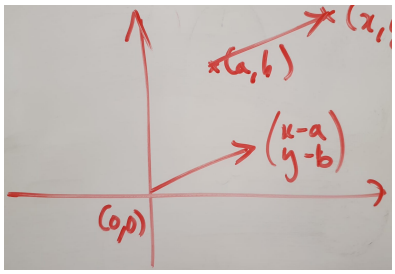
Dans ce cas on dit que α est la différentielle de f en A, *et on la note $df_A : \mathbb{R}^n \to \mathbb{R}^m$.*

Remarque 1.29. Notons que dans la deuxième formule, nous ne spécifions pas comment M tend vers A. Il peut converger en ligne droite ou en spirale, ou selon n'importe quelle courbe, et même de façon discrète. L'important est justement qu'on peut converger de n'importe quelle façon vers A, et que l'application linéaire $\alpha = df_A$ est bien définie dans tous les cas.

Remarque 1.30. Pour $n = 2$ et $m = 1$, la première formule se lit :

$$f(x,y) = f(a,b) + \alpha\begin{pmatrix} x-a \\ y-b \end{pmatrix} + o(\|(x,y)-(a,b)\|_2)$$

où on a pris $A = (a,b)$, $M = (x,y) = A + \begin{pmatrix} x-a \\ y-b \end{pmatrix}$, et $\alpha : \mathbb{R}^2 \to \mathbb{R}$ est une matrice ligne à deux colonnes.



Proposition 1.31. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction différentiable en un point A. Alors la différentielle de f en A est unique et elle est définie par :*

$$\begin{aligned} df_A : \mathbb{R}^n &\longrightarrow \mathbb{R}^m \\ \vec{v} &\longmapsto D_{\vec{v}} f(A) \end{aligned}$$

D'autre part si f est une application linéaire alors $df_A = f$.

Démonstration. En effet la deuxième formule de la Definition 1.28 nous dit que, si on prend un vecteur $\vec{v}$, si f est différentiable en A alors (ce n'est pas une équivalence mais une implication) :

$$\lim_{t\to 0} \frac{\|f(A+t\vec{v}) - f(A) - df_A(t\vec{v})\|}{|t|} = 0$$

ce qui se récrit, en faisant rentrer $|t|$ à l'intérieur de la norme du haut, et par linéarité de la différentielle de f en A :

$$\lim_{t\to 0} \left\| \frac{f(A+t\vec{v}) - f(A) - t\, df_A(\vec{v})}{t} \right\| = 0$$

ce qui donne comme résultat :

$$\lim_{t\to 0} \left\| \frac{f(A+t\vec{v}) - f(A)}{t} - df_A(\vec{v}) \right\| = 0$$

Le premier terme dans la limite tend vers $D_{\vec{v}} f(A)$ et le deuxième terme ne dépend pas de t, donc on a à la limite à la limite $t = 0$: $\|D_{\vec{v}} f(A) - df_A(\vec{v})\| = 0$, c'est à dire $df_A(\vec{v}) = D_{\vec{v}} f(A)$.

Si f est une application linéaire, alors la dérivée directionnelle dans la direction $\vec{v}$ est donnée par :

$$D_{\vec{v}} f(A) = \lim_{t\to 0} \frac{f(A+t\vec{v}) - f(A)}{t} = \lim_{t\to 0} \frac{f(A) + tf(\vec{v}) - f(A)}{t} = \lim_{t\to 0} f(\vec{v}) = f(\vec{v})$$

Nous en déduisons que si f est une application linéaire alors $df_A = f$. □

Corollaire 1.32. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction différentiable en un point A. Alors nous avons les résultats suivants :*

1. *la fonction f est continue en A ;*
2. *les dérivées directionnelles de f existent dans toutes les directions, et*
3. *la condition de linéarité des dérivées directionnelles (1.3) est satisfaite pour f :*

$$\forall\, \lambda, \mu \in \mathbb{R},\ \forall\, \vec{v}, \vec{w} \in \mathbb{R}^n \qquad D_{\lambda\vec{v}+\mu\vec{w}} f(A) = \lambda D_{\vec{v}} f(A) + \mu D_{\vec{w}} f(A)$$

Démonstration. Le premier item se voit sur la définition de la différentielle, et les deux derniers items viennent de la Proposition 1.31 qui nous donne l'expression de la différentielle de f en A, qui est une application linéaire. □

Remarque 1.33. Nous voyons que la définition de la différentielle est plus générale que juste l'existence des dérivées directionnelles dans toutes les directions, ainsi que leur linéarité. Autrement la différentiabilité répond aux exemples du début de cette section qui montraient des cas où la fonction ne satisfaisait pas l'une des trois propriétés de la Proposition 1.32.

Remarque 1.34. La Définition 1.28 est cohérente avec la dérivation normale d'une fonction à une variable réelle. En effet, si $n = m = 1$, et si $f : \mathbb{R} \to \mathbb{R}$ est dérivable en un point a de son domaine de définition, alors on a d'après le cours de première année qu'il existe une fonction $\epsilon : \mathbb{R} \to \mathbb{R}$ qui tend vers 0 quand x tend vers a et définie par $\epsilon(x) = \frac{f(x)-f(a)}{x-a} - f'(a) = \frac{f(x)-f(a)-f'(x)(x-a)}{x-a}$. Dans ce cas, en comparant avec la Définition 1.28, la différentielle de la fonction f au point a est l'application linéaire

$$\begin{array}{rccc} df_a : & \mathbb{R} & \longrightarrow & \mathbb{R} \\ & h & \longmapsto & hf'(a) \end{array}$$

Exemple 1.35. Soit $f : \mathbb{R}^2 \to \mathbb{R}, x \mapsto x$ et $g : \mathbb{R}^2 \to \mathbb{R}, y \mapsto y$. La différentielle de f en n'importe quel point (a, b) du plan est donnée par la limite :

$$\lim_{(x,y)\to(a,b)} \frac{\left\| f(x,y) - f(a,b) - \alpha_f \begin{pmatrix} x-a \\ y-b \end{pmatrix} \right\|}{\|(x,y) - (a,b)\|} = 0$$

c'est à dire, en prenant comme direction de convergence le chemin $t \mapsto (a + t, b)$ (par facilité) :

$$\lim_{t\to 0} \frac{\|(a+t) - a - \alpha_f \left(\begin{smallmatrix} t \\ 0 \end{smallmatrix}\right)\|}{\|(a+t, b) - (a, b)\|} = 0$$

ce qui donne :

$$\lim_{t\to 0} \frac{\|t - \alpha_f \left(\begin{smallmatrix} t \\ 0 \end{smallmatrix}\right)\|}{|t|} = \lim_{t\to 0} \|1 - \alpha_f \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)\| = 0$$

Donc on déduit que $\alpha_f = \begin{pmatrix} 1 & ? \end{pmatrix}$. En refaisant le même processus avec le chemin $t \mapsto (a, b + t)$ nous obtenons que la deuxième composante est nulle, et donc $\alpha_f = \begin{pmatrix} 1 & 0 \end{pmatrix}$. De même, on trouve $\alpha_g = \begin{pmatrix} 0 & 1 \end{pmatrix}$. On note habituellement $dx = \alpha_f$ et $dy = \alpha_g$, ce sont des formes différentielles qui, à tout vecteur $\left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right)$ donnent :

$$dx\left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right) = h \qquad \text{et} \qquad dy\left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right) = k$$

La notation n'est pas arbitraire car elles correspondent bien aux incréments infinitésimaux de la physique. Notons que $dx = i^*$ et $dy = j^*$, ce n'est pas un hasard non plus.

Revenons au cas à deux variables, c'est à dire $n = 2$ et $m + 1$ Comme $\mathcal{L}(\mathbb{R}^2, \mathbb{R}) \simeq (\mathbb{R}^2)^*$, la différentielle de f en un point $A = (a, b)$ est une forme linéaire, c'est à dire un élément de $(\mathbb{R}^2)^*$, qui mange un vecteur et renvoie un réel. Nous savons donc qu'elle se décompose sur la base (i^*, j^*) de $(\mathbb{R}^2)^*$ comme tout application linéaire, sous la forme $df_{(a,b)} = ri^* + sj^*$ pour deux composantes r, s à déterminer. Pour ce faire, nous évaluons $df_{(a,b)}$ sur $\vec{i}$ et $\vec{j}$, pour obtenir :

$$df_{(a,b)}(\vec{i}) = D_{\vec{i}} f(a, b) = \frac{\partial f}{\partial x}(a, b) \qquad \text{et} \qquad df_{(a,b)}(\vec{j}) = D_{\vec{j}} f(a, b) = \frac{\partial f}{\partial y}(a, b)$$

Et comme d'autre part nous avons que $df_{(a,b)}(\vec{i}) = r$ et $df_{(a,b)}(\vec{j}) = s$ grâce aux identités (1.2), nous déduisons les composantes de la forme linéaire $df_{(a,b)}$ sur la base (O, i^*, j^*) de $(\mathbb{R}^2)^*$ comme :

$$df_{(a,b)} = \frac{\partial f}{\partial x}(a,b)\, i^* + \frac{\partial f}{\partial y}(a,b)\, j^*$$

c'est à dire sous la forme matricielle ligne 1×2 qui sont les éléments de $\left(\mathbb{R}^2\right)^*$:

$$df_{(a,b)} = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) & \frac{\partial f}{\partial y}(a,b) \end{pmatrix} \tag{1.4}$$

On peut voir i^* et j^* comme les formes différentielles dx et dy de l'Exemple 1.35 ce qui permet d'écrire :

$$df_{(a,b)} = \frac{\partial f}{\partial x}(a,b)\, dx + \frac{\partial f}{\partial y}(a,b)\, dy$$

Ce qui veut dire que pour tout $\vec{v} = (h,k)$, on a la propriété de linéarité suivante :

$$df_{(a,b)}(h,k) = h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b)$$

Attention cette formule n'est vraie que si f est différentiable en a, c'est à dire non seulement ses dérivées directionnelles existent dans toutes les directions, mais aussi elles varient linéairement en fonction du vecteur $\vec{v} = \left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right)$ avec lequel on différentie. L'hypothèse d'avoir un point à l'intérieur du domaine de définition est important car on veut différentier dans toutes les directions. Si on est à la frontière, on ne peut pas le faire. Il arrive la même chose pour les fonctions à une variable.

Interprétation géométrique : Comme dans le cas d'une fonction à une variable, on peut tenter d'interpréter géométriquement les notions définies plus haut. Vous savez que, pour une fonction f à une variable, le nombre dérivé $f'(a)$ représente le coefficient directeur de la tangente à la courbe représentative de la fonction f en son point d'abscisse a. Autrement dit, la tangente étant la droite la plus proche de la courbe, on peut écrire au voisinage de a l'approximation affine suivante : $f(x) \simeq f(a) + f'(a)(x-a)$ (équation de la tangente), ou encore en changeant les notations $f(a+h) \simeq f(a) + hf'(a)$, approximation valable pour des « petites » valeurs de h. En passant $f(a)$ à gauche on obtient la variation $\Delta_a f \simeq hf'(a)$ mais si on note df_a au lieu de $\Delta_a f$ et bien on obtient la formule de la différentielle de f pour une fonction à une variable. C'est-à-dire que l'accroissement de la fonction f (à gauche, qui correspond à $f(a+h) - f(a)$) est proportionnel à l'accroissement de la variable (qui correspond à h), avec pour coefficient de proportionnalité $f'(a)$.

Dans le cas d'une fonction à deux variables, les dérivées partielles en un point (a,b) représentent également des coefficients directeurs de tangentes, en l'occurence des deux tangentes à la surface représentative de f incluses dans les plans verticaux contenant les axes (Ox) et (Oy). La surface admet bien d'autres tangentes (une infinité), mais la connaissance de deux d'entre elles suffit à déterminer le plan tangent à la surface représentative de f, et donc à donner une approximation de $f(a+h, b+k)$ pour des petites valeurs de h et k. C'est ce que fait la différentielle à l'aide de la formule suivante : $f(a+h, b+k) \simeq h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b)$. Autrement dit (et pour parler en termes plus « économiques »), la différentielle exprime l'accroissement marginal de la fonction f au point (a,b) en fonction des accroissements marginaux de chacune des variables.

Exemple 1.36. Calculons la différentielle de la fonction norme carrée : $f : (x,y) \mapsto \|(x,y)\|^2 = x^2 + y^2$ de $\mathbb{R}^2$ dans $\mathbb{R}$, à un point (a,b) fixé. Pour tout $\left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right) \in \mathbb{R}^2$ on cherche à trouver une forme linéaire $\alpha : \mathbb{R}^2 \to \mathbb{R}$ telle que :

$$f(a+h, b+k) = f(a,b) + \alpha\left(\begin{smallmatrix} h \\ k \end{smallmatrix}\right) + \underset{(h,k)\to(0,0)}{o}\left(\|(h,k)\|\right)$$

On développe la partie gauche pour obtenir quelque chose à droite :

$$\begin{aligned} f(a+h,b+k) &= \|(a+h,b+k)\|^2 = (a+h)^2+(b+k)^2 \\ &= a^2+b^2+2ah+2bk+h^2+k^2 \\ &= \|(a,b)\|^2 + 2(ah+bk) + \|(h,k)\|^2 \\ &= f(a,b) + 2\left\langle \binom{a}{b}, \binom{h}{k} \right\rangle + \underset{(h,k)\to(0,0)}{o}\big(\|(h,k)\|\big) \end{aligned}$$

Et donc on observe que la différentielle $\alpha = df_{(a,b)}$ de f au point (a,b) est une fonction satisfaisant la relation suivante :

$$\begin{array}{rccc} df_{(a,b)} : & \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ & \binom{h}{k} & \longmapsto & 2\left\langle \binom{a}{b}, \binom{h}{k} \right\rangle \end{array}$$

La différentielle de la norme euclidienne au point (a,b) est deux fois le produit scalaire avec le vecteur $\binom{a}{b}$! L'écriture de la différentielle sous la forme (1.4) est $df_{(a,b)} = \begin{pmatrix} 2a & 2b \end{pmatrix}$.

Nous savons que si la fonction f est différentiable en un point A alors les dérivées partielles en ce point existent. Inversement, l'existence des dérivées partielles en un point ne dit rien sur la différentiabilité de f en ce point. Par exemple, si on reprend l'Exemple 1.16, on peut voir que la fonction f admet des dérivées directionnelles à l'origine selon toutes les directions, mais la fonction f ne peut pas être différentiable en l'origine car elle n'y est pas continue. On se retrouve un peu comme dans la situation vue pour les fonctions partielles, qui ne nous disent pas si une fonction est continue (c'est d'ailleurs à cette observation que l'Exemple 1.16 a été créé).

L'utilité des dérivées partielles va se voir dans une autre formulation de la différentielle. Nous savons qu'à toute application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^m$ correspond une matrice $m \times n$ (une fois choisie une base des deux espaces). Cela est bien entendu aussi vrai pour la différentielle au point A. Choisissons une base (e_j) de $\mathbb{R}^n$ et une base (e'_i) de $\mathbb{R}^m$. Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. Pour tout $1 \leq i \leq m$, on note $f^i : \mathbb{R}^n \to \mathbb{R}$ la i-ème composante de la fonction f dans la base de $\mathbb{R}^m$.

Définition 1.37. Matrice Jacobienne en dimension finie. *Supposons que la fonction $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ soit différentiable au point $A \in \mathring{D}$. Alors la matrice correspondant à l'application linéaire df_A (dans les bases choisies) est appelée* matrice Jacobienne *de f en A :*

$$J_A(f) = \left(\frac{\partial f^i}{\partial x^j}(A)\right)_{\substack{1\leq i\leq m \\ 1\leq j\leq n}} = \begin{pmatrix} \frac{\partial f^1}{\partial x^1}(A) & \frac{\partial f^1}{\partial x^2}(A) & \dots & \frac{\partial f^1}{\partial x^{n-1}}(A) & \frac{\partial f^1}{\partial x^n}(A) \\ \frac{\partial f^2}{\partial x^1}(A) & & & & \frac{\partial f^2}{\partial x^n}(A) \\ \dots & & \dots & & \dots \\ \frac{\partial f^{m-1}}{\partial x^1}(A) & & & & \frac{\partial f^{m-1}}{\partial x^n}(A) \\ \frac{\partial f^m}{\partial x^1}(A) & \frac{\partial f^m}{\partial x^2}(A) & \dots & \frac{\partial f^m}{\partial x^{n-1}}(A) & \frac{\partial f^m}{\partial x^n}(A) \end{pmatrix}$$

Remarque 1.38. Il faut éviter de confondre les vecteurs lignes et vecteurs colonnes, et à ne pas confondre les fonctions composantes f^i avec les fonctions partielles f_j (ce ne sont pas les mêmes indices) ! ! Dans une matrice, l'input se fait par le haut (donc ici on doit avoir n colonnes, et l'indice des colonnes est j), et l'output se fait par les lignes (donc ici m lignes, et l'indice des lignes est i). La j-ième colonne est pour la dérivée j-ème et la i-ème ligne est pour la i-ème fonction composante. Les coefficients de la matrice Jacobienne au point A dépendent donc du choix des deux bases de $\mathbb{R}^n$ et $\mathbb{R}^m$! L'intérêt de la matrice Jacobienne est qu'on peut écrire la différentielle en termes matriciels. La différentielle est un objet qui ne change pas, mais les coefficients de la matrice peuvent changer selon le choix des bases, mais dans ce cas les coordonnées des vecteurs changent aussi donc au final on retrouve le même résultat.

Remarque 1.39. L'indice i est appelé *contravariant* et l'indice j est appelé *covariant*.

Exemple 1.40. Pour une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ differentiable en $A = (a, b)$, comme la différentielle de f au point (a, b) est une forme linéaire – c'est à dire un élément de $(\mathbb{R}^2)^*$, nous avons vu qu'on peut décomposer la différentielle de f en (a, b) sur la base (O, i^*, j^*) de $(\mathbb{R}^2)^*$ comme dans la formule (1.4) sous forme matricielle :

$$df_{(a,b)} = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) & \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$$

Le membre de droite est la matrice Jacobienne de f en $A = (a, b)$.

Proposition 1.41. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction différentiable en $A \in \mathring{D}$, et soit $\vec{v} \in \mathbb{R}^n$. Alors on a :*

$$df_A(\vec{v}) = J_A(f) \times \vec{v}$$

Remarque 1.42. A gauche on a bien un vecteur de $\mathbb{R}^m$ c'est à dire un vecteur colonne à m composantes, et à droite on a un produit matriciel entre la matrice Jacobienne $m \times n$ et le vecteur colonne $\vec{v}$ à n composantes, c'est à dire qu'on obtient un vecteur colonne à m composantes, comme à gauche.

Exemple 1.43. La fonction $f : \mathbb{R}_+^* \times]-\pi, \pi[\longrightarrow \mathbb{R}^2 \backslash (\mathbb{R}_- \times \{0\}), (r, \theta) \longmapsto (x, y) = (r\cos(\theta), r\sin(\theta))$ est définie sur un ouvert. Calculons sa différentielle :

$$\begin{aligned} f(r+h, \theta+k) &= (r\cos(\theta+k) + h\cos(\theta+k), r\sin(\theta+k) + h\sin(\theta+k)) \\ &= (r\cos(\theta)\cos(k) - r\sin(\theta)\sin(k) + h\cos(\theta)\cos(k) - h\sin(\theta)\sin(k), \\ &\quad r\sin(\theta)\cos(k) + r\cos(\theta)\sin(k) + h\sin(\theta)\cos(k) + h\cos(\theta)\sin(k)) \end{aligned}$$

Donc pour h, k très petits on a notamment $\cos(k) \sim 1$ et $\sin(k) \sim k$, d'où :

$$\begin{aligned} f(r+h, \theta+k) &= (r\cos(\theta), r\sin(\theta)) + (h\cos(\theta) - kr\sin(\theta), h\sin(\theta) + kr\cos(\theta)) + hk(-\sin(\theta), \cos(\theta)) \\ &= f(r, \theta) + \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix} \times \begin{pmatrix} h \\ k \end{pmatrix} + o(||(h, k)||) \end{aligned}$$

du fait que $|hk| \leq \frac{h^2+k^2}{2} = \frac{(||(h,k)||_2)^2}{2}$. La différentielle de f au point (r, θ) peut donc bien être représentée par une matrice 2×2, qui se trouve exactement être la matrice Jacobienne.

La matrice Jacobienne permet d'écrire facilement certains résultats compliqués mais très importants :

Proposition 1.44. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ et $g : \mathbb{R}^m \to \mathbb{R}^p$ deux fonctions telles que f est différentiable au point $A \in \mathbb{R}^n$ et g est différentiable au point $f(A) \in \mathbb{R}^p$. La fonction composée $g \circ f : \mathbb{R}^n \to \mathbb{R}^p$ est différentiable en A et les formules suivantes sont équivalentes :*

1. *Au niveau des dérivées partielles on a la* règle des chaines *:*

$$\text{pour tout } 1 \leq j \leq n \qquad \frac{\partial (g \circ f)}{\partial x^j}(A) = \sum_{i=1}^{m} \frac{\partial f_i}{\partial x^j}(A) \times \frac{\partial g}{\partial y^i}(f(A))$$

2. *Les différentielles de f, g et $g \circ f$ satisfont :*

$$d(g \circ f)_A = dg_{f(A)} \circ df_A$$

3. *Les matrices Jacobiennes de f, g et $g \circ f$ satisfont :*

$$J_A(g \circ f) = J_{f(A)}(g) \times J_A(f)$$

Démonstration. Pour tout $1 \leq i \leq m$, on note $f^i : \mathbb{R}^n \to \mathbb{R}$ la i-ème composante de la fonction f, à ne pas confondre avec les fonctions partielles ! On observe alors qu'en toute généralité, si on prend un vecteur de base $e_j \in \mathbb{R}^n$ et on effectue la dérivée directionnelle de $g \circ f$ par rapport à e_j on a par la règle des chaines :

$$d(g \circ f)_A(e_j) = D_{e_j}(g \circ f)(A) = \frac{\partial(g \circ f)}{\partial x^j}(A) = \sum_{i=1}^{m} \frac{\partial f^i}{\partial x^j}(A) \times \frac{\partial g}{\partial y^i}(f(A)) = J_{f(A)}(g) \times J_A(f) \times e_j$$

Ce qui donne le produit des matrices Jacobiennes. □

L'analogie avec la dimension 1 – les fonctions réelles à une variable – continue : nous savons que si une fonction à une variable réelle admet un maximum ou un minimum, sa dérivée en ce point s'annule. Nous allons avoir le même résulat ici, une fois généralisée de façon tout à fait naturelle la notion de maximum ou minimum :

Définition 1.45. *Soit $f : U \subset \mathbb{R}^n \to \mathbb{R}$ une fonction définie sur un ouvert U et à valeurs dans $\mathbb{R}$, et différentiable en un point $A \in U$. On dit que A est :*

- *un* miminum local *s'il existe une boule ouverte $B(A, \epsilon) \subset U$ de rayon $\epsilon > 0$ telle que $f(A) \leq f(x)$ pour tout $x \in B(A, \epsilon)$;*
- *un* maximum local *s'il existe une boule ouverte $B(A, \epsilon) \subset U$ de rayon $\epsilon > 0$ telle que $f(A) \geq f(x)$ pour tout $x \in B(A, \epsilon)$;*
- *un* extremum local *si c'est un minimum ou un maximum local ;*
- *un* point critique *de f si $J_A(f)$ est la matrice rectangulaire $1 \times n$ nulle, i.e. si $df_A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R})$ est l'application linéaire nulle.*

Remarque 1.46. On parle de minimum et de maximum *global* si la définition est valide pour tout ϵ.

Exemple 1.47. Reprenons les fonctions de l'exemple 1.13. On a $f(x, y) = x^2 + y^2$ pour le paraboloïde elliptique et $g(x, y) = x^2 - y^2$ pour le paraboloïde hyperbolique. On a $df_{(a,b)} = \begin{pmatrix} 2a & 2b \end{pmatrix}$ et $dg_{(a,b)}\begin{pmatrix} 2a & 2b \end{pmatrix}$, donc $(0, 0)$ est un point critique de f et g. C'est un minimum de f mais c'est ce qu'on appelle un *point col* ou *point selle* pour g : selon la direction x l'origine est un minimum, mais c'est un maximum selon la direction y.

Le dernier item de la définition a un parallèle avec les nombres dérivés pour les fonctions à une variable : une fonction $f : \mathbb{R} \to \mathbb{R}$ admet un point critique a si $f'(a) = 0$. Comme le montre les exemples, les extrema locaux sont des points critiques, mais tous les points critiques ne sont pas des extrema, comme le montre par exemple l'origine $a = 0$ pour la fonction $x \mapsto x^3$. Nous n'avons fait que généraliser cette notion aux fonctions à plusieurs variables. Cela nous donne une condition nécessaire similaire au cas une variable pour trouver les extrema locaux des fonctions à plusieurs variables à valeurs dans $\mathbb{R}$ dans notre cas :

Proposition 1.48. *Si A est un extremum local alors $J_f(A) = 0$, i.e. $df_A = 0$.*

Exemple 1.49. La fonction $f(x, y) = x^2 + y^2$ donne comme dérivée partielle $\frac{\partial f}{\partial x} = 2x$ et $\frac{\partial f}{\partial y} = 2y$. La matrice Jacobienne au point (x, y) est $J_{(x,y)}(f) = \begin{pmatrix} 2x & 2y \end{pmatrix}$, et ceci s'annule effectivement à l'origine (le graphe de f y admet un minimum). C'est un minimum car le graphe de f est un paraboloïde elliptique.

Exemple 1.50. Attention tous les points critiques ne sont pas des extrema. Par exemple la fonction $f(x, y) = x^2 - y^2$ donne comme dérivée partielle $\frac{\partial f}{\partial x} = 2x$ et $\frac{\partial f}{\partial y} = -2y$. La matrice Jacobienne au point (x, y) est $J_{(x,y)}(f) = \begin{pmatrix} 2x & -2y \end{pmatrix}$, qui s'annule effectivement à l'origine, mais ce n'est ni un minimum ni un maximum car le graphe de f est un Pringle/col/selle de cheval (on l'appelle paraboloïde hyperbolique).

Définition 1.51. *Nous avons deux cas extrêmes, lorsque $m = 1$ ou $m = n$, que nous donnons ici :*

1. *Lorsque $m = n$, la matrice Jacobienne de $f : \mathbb{R}^n \to \mathbb{R}^n$ en A est carrée, et sa trace est appelée* divergence de la fonction f en A *: $\mathrm{div}_A(f) = \sum_{i=1}^n \frac{\partial f_i}{\partial x^i}(A)$.*
2. *Lorsque $m = 1$, la transposée de la matrice Jacobienne de $f : \mathbb{R}^n \to \mathbb{R}$ en A est un n-vecteur appelé le* gradient de f en A*, et noté $\mathrm{grad}_A(f)$ ou $\nabla f(A)$:*

$$\nabla f(A) = \left(\frac{\partial f}{\partial x^i}(A)\right)_{1\leq i\leq n} = \begin{pmatrix} \frac{\partial f}{\partial x^1}(A) \\ \frac{\partial f}{\partial x^2}(A) \\ \dots \\ \frac{\partial f}{\partial x^n}(A) \end{pmatrix} = (J_A(f))^t$$

Exemple 1.52. Dernière notion utile en économie et abordée un peu plus haut d'un point de vue mathématique : les lignes de niveau de la fonction P sont appelés *isoquants* de la fonction de production. Ils représentent des lignes sur lesquelles la production ne varie pas, et on peut donc affirmer que, le long d'un isoquant, la différentielle dP s'annule. Autrement dit, si le vecteur $t \mapsto (h_t, k_t)$ est un vecteur tangent à un isoquant, on a alors $h_t \frac{\partial P}{\partial K}(K_0, L_0) + k_t \frac{\partial P}{\partial L}(K_0, L_0) = 0$. Les rapports entre les deux dérivées partielles en un point d'un isoquant sont appelés *coefficients d'elasticité* : ils représentent la facilité à échanger du capital contre du travail (ou vice-versa) pour garder une production constante. Ainsi, si on a $\frac{\partial P}{\partial K}(K_0, L_0) = -3\frac{\partial P}{\partial L}(K_0, L_0)$, cela signifie que, si on diminue le capital de 1% en partant de la situation (K_0, L_0), il faudra augmenter le travail de 3% pour garder le même niveau de production.

Proposition 1.53. *Supposons que f est différentiable sur l'intérieur de son ensemble de définition D. Alors nous avons la formule suivante :*

$$\forall\, A \in \mathring{D},\ \forall\, \vec{v} \in \mathbb{R}^n \qquad df_A(\vec{v}) = \langle \nabla f(A), \vec{v} \rangle$$

Démonstration. On le fait grâce au calcul matriciel. Soit $e_1, \dots, e_n$ une base de $\mathbb{R}^n$. Supposons que $\vec{v} = v^1 e_1 + \dots + v^n e_n$, alors on a d'après la définition de la matrice Jacobienne et la Proposition 1.41 :

$$df_A(\vec{v}) = \begin{pmatrix} \frac{\partial f}{\partial x^1}(A) & \dots & \frac{\partial f}{\partial x^n}(A) \end{pmatrix} \times \begin{pmatrix} v^1 \\ \dots \\ v^n \end{pmatrix} = \left\langle \begin{pmatrix} \frac{\partial f}{\partial x^1}(A) \\ \dots \\ \frac{\partial f}{\partial x^n}(A) \end{pmatrix}, \begin{pmatrix} v^1 \\ \dots \\ v^n \end{pmatrix} \right\rangle = \langle \nabla f(A), \vec{v} \rangle$$

C'est bien le résultat demandé. □

Remarque 1.54. Notons qu'en général, le terme $\langle \nabla f(A), \vec{v} \rangle$ peut être écrit $\nabla_{\vec{v}} f(A)$, cette notation permettant de faire le lien avec les connections en géométrie différentielle. Autrement dit, nous avons $\nabla_{\vec{v}} f(A) = df_{As}(\vec{v})$.

1.3 Gradient et lignes de niveaux

Définition 1.55. *Si f est différentiable en tout point de l'intérieur de son domaine de définition D, on dit que f est* différentiable *sur $\mathring{D}$.*

Pour toute fonction à une variable qui est dérivable sur tout son ensemble de définition (qu'on suppose ouvert), on peut associer une fonction dérivée en associant, à tout point a de l'ensemble de définition de f, sa dérivée $f'(a)$. On peut procéder de même avec une fonction

différentiable en tout point de l'intérieur de son ensemble de définition, en associant, à tout point $A = (a, b) \in \mathring{D}$, la *différentielle de f au point A* :

$$\begin{aligned} df : \mathring{D} \subset \mathbb{R}^2 &\longrightarrow \left(\mathbb{R}^2\right)^* \\ (a,b) &\longmapsto df_{(a,b)} : \begin{pmatrix} h \\ k \end{pmatrix} \mapsto h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b) \end{aligned}$$

La fonction différentielle $df : \mathring{D} \to \left(\mathbb{R}^2\right)^*$ associe, à tout point de $\mathring{D}$ un élément du dual de $\mathbb{R}^2$, c'est à dire une forme linéaire (covecteur) sur $\mathbb{R}^2$. C'est ce qu'on appelle un *champ de covecteurs* ou *codistribution de rang 1*, ou *forme différentielle*. Si la fonction $df : \mathring{D} \to \left(\mathbb{R}^2\right)^*$ est continue dans le sens des fonctions à plusieurs variables, alors on dit que f *est de classe* $\mathcal{C}^1$.Pour définir la continuité de ce type de fonction, on observe que $\left(\mathbb{R}^2\right)^* \simeq \mathbb{R}^2$, et donc qu'on peut identifier tout covecteur avec un vecteur en prenant sa transposée. On peut donc prendre comme norme sur le dual $\left(\mathbb{R}^2\right)^*$ la norme euclidienne classique sur $\mathbb{R}^2$. C'est à dire que pour tout covecteur $(\,r\ s\,)$, on a :

$$\left\|(\,r\ s\,) - df_{(a,b)}\right\| = \sqrt{\left(r - \frac{\partial f}{\partial x}(a,b)\right)^2 + \left(s - \frac{\partial f}{\partial y}(a,b)\right)^2}$$

Cela nous permet donc de faire sens de la phrase : la différentielle de f est continue si :

$$\forall\, \epsilon > 0,\ \exists\, \delta > 0 \text{ tel que } \forall\, (x,y) \in D \cap B((a,b),\epsilon),\ df_{(x,y)} \in B(df_{(a,b)}, \epsilon)$$

Définition 1.56. Différentielle en dimension finie. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction. Si f est différentiable en tout point $A \in \mathring{D}$ de l'intérieur de son domaine de définition, alors on dit que f est* différentiable sur $\mathring{D}$*, et la fonction à plusieurs variables :*

$$\begin{aligned} df : \ \mathring{D} &\longrightarrow \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \\ A &\longmapsto df_A : v \mapsto D_v f(A) \end{aligned}$$

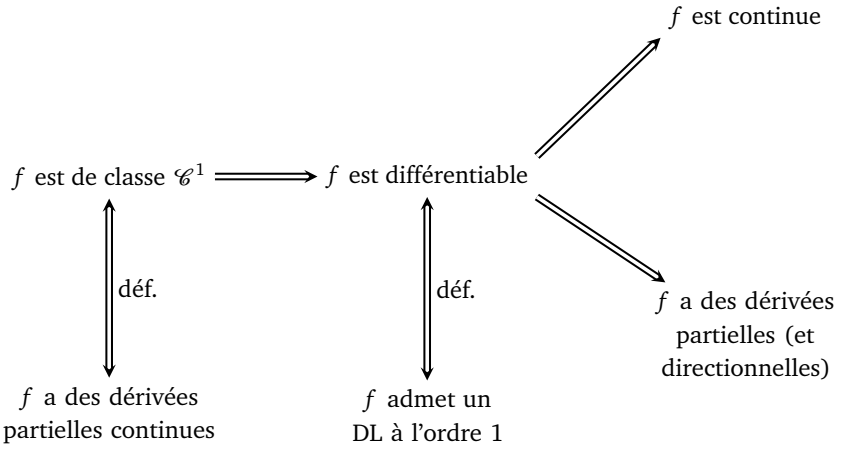
est appelée la différentielle de f*. Dans ce cas f est nécessairement continue sur $\mathring{D}$. Si de plus la fonction différentielle $df : \mathring{D} \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ est continue, alors on dit que f est* continûment différentiable *ou* de classe $\mathcal{C}^1$*.*

Exemple 1.57. Un type de fonction à deux variables souvent utilisé en économie est la fonction de Cobb-Douglas, qui modélise la production P en fonction du capital K et du travail L via la formule $P(K, L) = cK^\alpha L^\beta$. Pour plus de simplicité, on prend souvent $c = 1$, et $\beta = 1 - \alpha$ (avec $\alpha \in [0, 1]$), donc $P(K, L) = K^\alpha L^{1-\alpha}$. Ceci a également pour avantage de rendre la fonction homogène, c'est-à-dire que $P(\lambda K, \lambda L) = \lambda P(K, L)$ (autrement dit, si vous multipliez le capital et le travail simultanément par un même facteur, la production subira la même augmentation).

Les dérivées partielles de cette fonction sont appelés en économie *rendements marginaux*. En utilisant la notation différentielle, ces rendements marginaux donnent les coefficients d'augmentation marginale de la production quand on augmente marginalement le travail ou le capital. Ainsi, si $\frac{\partial P}{\partial K}(K_0, L_0) = 3$, cela signifie qu'en augmentant de 1% le capital en partant d'une situation où le capital était de K_0 et le travail de L_0, la production augmentera d'environ 3%. Avec l'équation donnée plus haut, on constate que $\frac{\partial P}{\partial K}(K, L) = \alpha K^{\alpha-1} L^{1-\alpha} = \alpha \left(\frac{L}{K}\right)^{1-\alpha}$, et $\frac{\partial P}{\partial L}(K, L) = (1-\alpha) K^\alpha L^{-\alpha} = (1-\alpha) \left(\frac{K}{L}\right)^\alpha$. Si on calcule désormais les dérivées secondes, on obtient notamment $\frac{\partial^2 P}{\partial K^2}(K, L) = \alpha(\alpha - 1) K^{\alpha-2} L^{1-\alpha}$ et $\frac{\partial^2 P}{\partial L^2}(K, L) = \alpha(\alpha - 1) K^\alpha L^{-1-\alpha}$. Ces deux expressions sont négatives, c'est un résultat qu'on connait en économie sous le nom de *principe des rendements décroissants* : plus la production augmente, plus les rendements marginaux sont faibles.

Malheureusement il peut être parfois pas très pratique de manipuler la différentielle et on préfère utiliser les dérivées partielles. En particulier, si on évalue la différentielle de f au point A et en le vecteur de base e_i, on obtient la dérivée partielle : $df_A(e_i) = \frac{\partial f}{\partial x^i}(A)$. Ce qui nous donne le lien suivant entre différentielle et dérivées partielles :

Proposition 1.58. *La fonction $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ est de classe $\mathcal{C}^1$ si et seulement si ses dérivées partielles existent et sont continues.*



Remarque 1.59. Soulignons ici que le caractère $\mathcal{C}^1$ ne dépend pas de la base par rapport à laquelle les dérivées partielles sont définies. Cela est fondamental, et sera généralisé aux ordre supérieurs plus tard.

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables et soit $(a, b) \in \mathring{D}$. Rappelons qu'on peut décomposer la différentielle de f en un point (a, b) sur la base (O, i^*, j^*) de $\left(\mathbb{R}^2\right)^*$ comme :

$$df_{(a,b)} = \frac{\partial f}{\partial x}(a, b)\, i^* + \frac{\partial f}{\partial y}(a, b)\, j^*$$

Ce qui veut dire que sous forme matricielle, nous avons :

$$df_{(a,b)} = \begin{pmatrix} \frac{\partial f}{\partial x}(a, b) & \frac{\partial f}{\partial y}(a, b) \end{pmatrix}$$

En lui appliquant la transposée qui est un isomorphisme entre vecteurs de $\mathbb{R}^2$ et covecteurs de $(\mathbb{R}^2)^*$, on obtient donc un vecteur de $\mathbb{R}^2$, qui se nomme le *gradient* de f en $A = (a, b)$:

$$\nabla f(a, b) = \left(df_{(a,b)}\right)^t = \begin{pmatrix} \frac{\partial f}{\partial x}(a, b) \\ \frac{\partial f}{\partial y}(a, b) \end{pmatrix}$$

Le gradient de f possède donc les mêmes informations que la différentielle de f. Si f est différentiable en tout point de $\mathring{D}$, alors le gradient est défini en tout point aussi. Cela induit une application duale de la différentielle, en prenant la transposée :

$$\begin{aligned} \nabla f = (df)^t : \quad \mathring{D} &\longrightarrow \mathbb{R}^2 \\ (a, b) &\longmapsto \nabla f(a, b) \end{aligned}$$

Là où la fonction différentielle définissait un champ de covecteurs, le gradient de f définit ce qu'on appelle plus généralement un *champ de vecteurs*. C'est un objet qui associe un vecteur à tout point de $\mathbb{R}^2$:

Définition 1.60. *Soit $D \subset \mathbb{R}^n$ un domaine. Un champ de vecteurs sur D est une application $\vec{X} : D \longrightarrow \mathbb{R}^n$ (aussi notée sans la flèche). Si cette application vectorielle est continue (resp. de classe $\mathcal{C}^1$), on dit que le champ de vecteurs est continue (resp. de classe $\mathcal{C}^1$).*

En général dans ce cours on considère des champs de vecteurs continus ou de classe $\mathcal{C}^1$. On peut représenter un champ de vecteur vu de dessus, en regardant le plan, et en traçant quelques flèches qui désignent la *tendance* qu' a le champ de vecteur dans cette région. La plupart des champs de vecteurs que l'on étudie seront continus (lisses en vrai), c'est à dire que la longueur et la direction des flèches varient continument lorsqu'on parcourt le plan $\mathbb{R}^2$.

Exemple 1.61. Regardons les deux champs de vecteurs suivants :

$$\vec{F} : (x, y) \longmapsto 3\vec{i} - 2\vec{j} \qquad \text{et} \qquad \vec{G} : (x, y) \longmapsto y\vec{i} - x\vec{j}$$

Le premier est constant, tandis que le deuxième est un champ de vecteur qui tourne autour de l'origine dans le sens horaire.

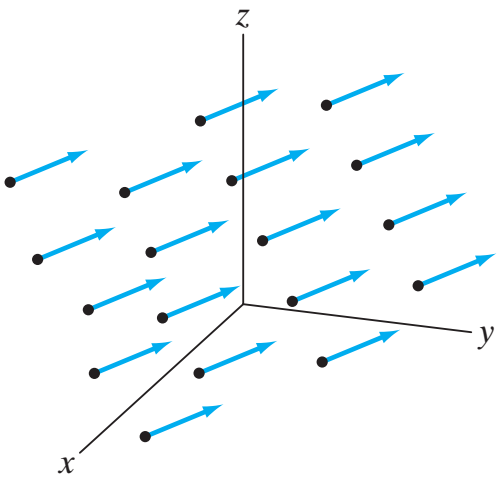


Figure 3.40 The constant vector field $\mathbf{F} = \mathbf{a}$.

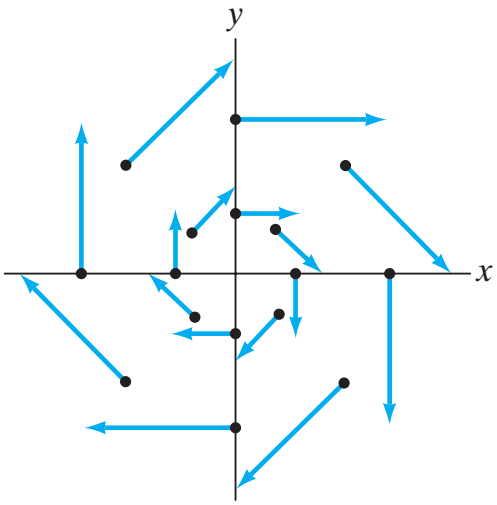
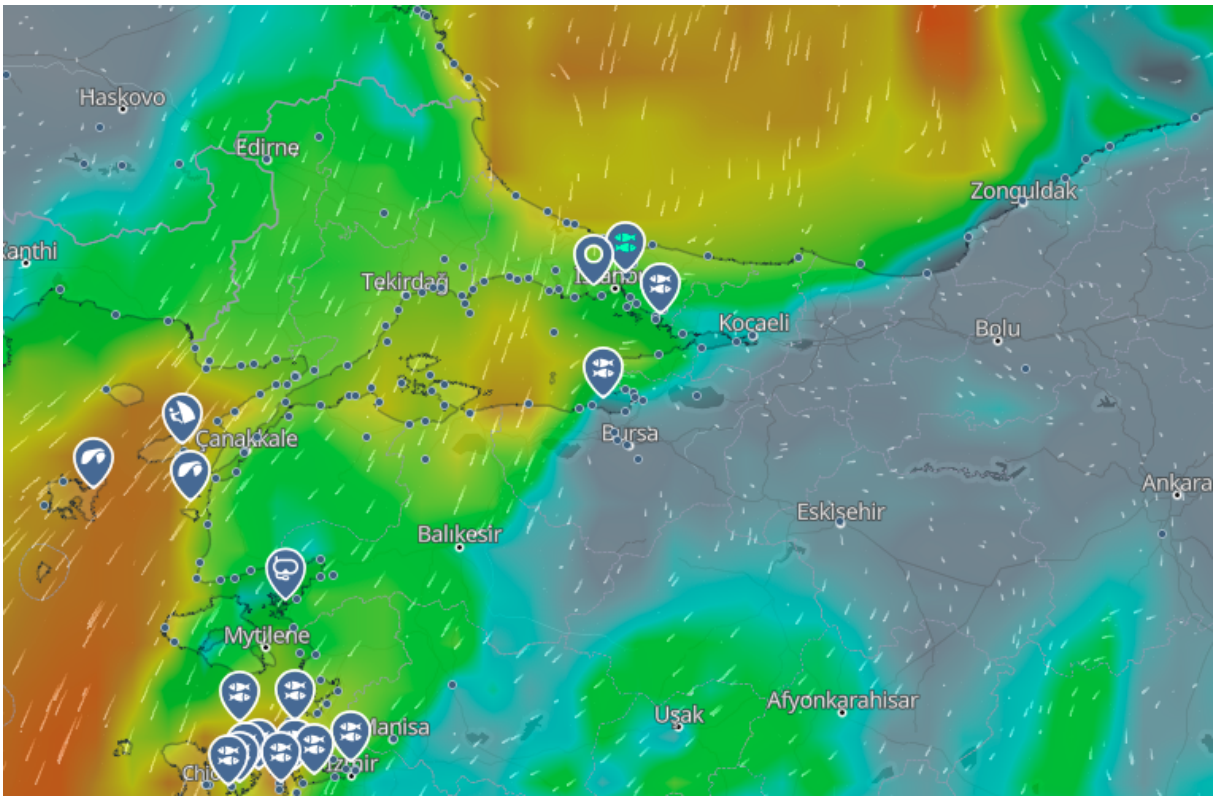


Figure 3.41 The vector field $\mathbf{G} = y\mathbf{i} - x\mathbf{j}$ resembles the velocity field of a rotating fluid.

Exemple 1.62. La carte des vents de surface sur Terre forme une représentation d'un champ de vecteurs "vent". Ce champ de vecteurs est une application de $\mathbb{R}^2$ (car la tête est localement plate) dans $\mathbb{R}^2$ (l'espace de la vitesse des vents). Le théorème du point fixe de Brouwer, dit aussi "de la Boule Chevelue" ou de la "noix de coco", nous dit qu'il doit nécessairement exister à tout temps t un point de la Terre en lequel le vent est nul. Ce théorème n'est pas prouvable physiquement, mais est un résultat puissant de topologie (toute application continue de la sphère dans elle même admet au moins un point fixe).



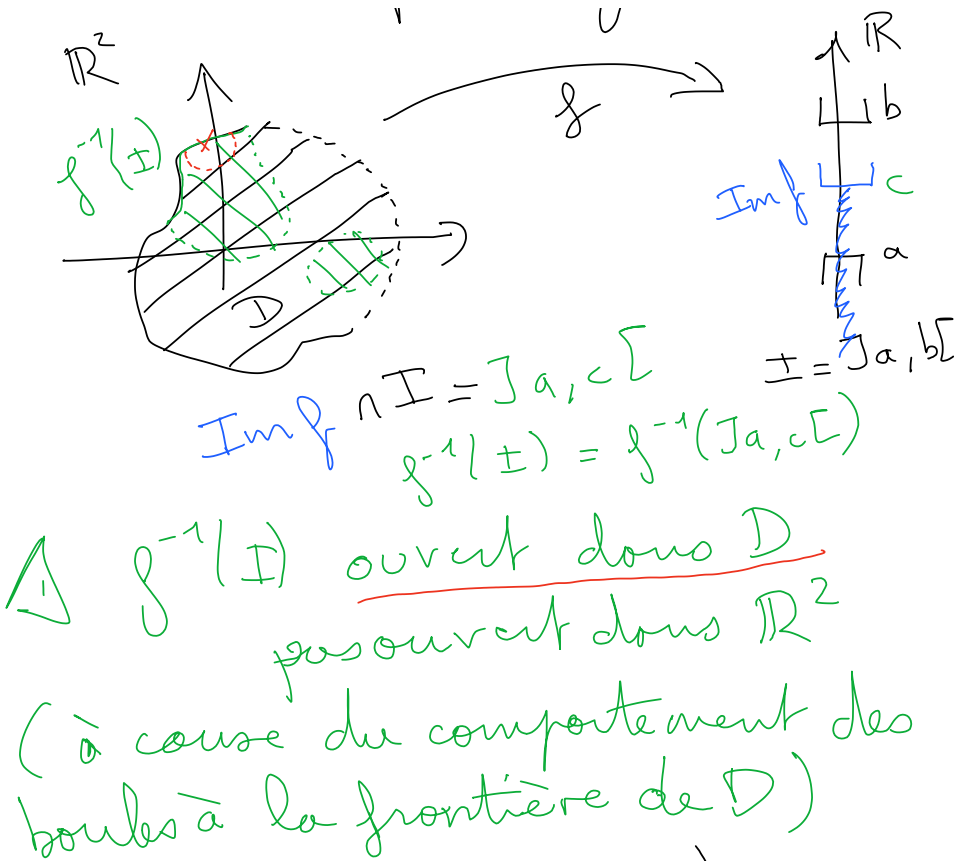
Nous allons voir maintenant une propriété intéressante possédée par le gradient, en tant que champ de vecteur. Nous devons pour cela introduire deux notions importantes, l'une tracée sur le domaine de définition, l'autre tracée sur le graphe. Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables. Soit $P \subset \mathbb{R}$ un sous-ensemble de $\mathbb{R}$, on note la préimage de P par f le sous-ensemble de $D \subset \mathbb{R}^2$ suivant :

$$f^{-1}(P) = \Big\{(x, y) \in D \subset \mathbb{R}^2 \text{ tel que } f(x, y) \in P\Big\}$$

Attention, cerains points de $\mathrm{Im}(f)$ ne sont pas forcément dans P, et certains points de P ne sont pas forcément dans l'image de f, mais cela n'a pas d'importance, on ne regarde que ceux qui sont dans $\mathrm{Im}(f) \cap P$ pour construire $f^{-1}(P)$. Nous avons alors la propriété suivante (admise) :

Proposition 1.63. *Les trois conditions suivantes sont équivalentes :*

1. *la fonction $f : D \to \mathbb{R}$ est continue ;*
2. *pour tout intervalle ouvert $I \subset \mathbb{R}$, la préimage $f^{-1}(I)$ est une partie ouverte de D ;*
3. *pour toute partie fermée $F \subset \mathbb{R}$, la préimage $f^{-1}(F)$ est une partie fermée de D.*



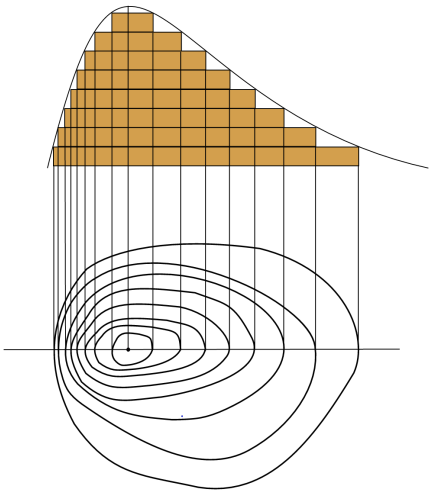
Exemple 1.64. Pour la fonction $f(x, y) = \ln(1 + x + y)$, la fonction est continue. En effet pour tout intervalle ouvert de la forme $]a, b[$, la préimage de cet intervalle est la partie de D qui est strictement comprise entre les deux droites d'équations $y = -x - 1 + e^a$ et $y = -x - 1 + e^b$.

Définition 1.65. *On appelle* ligne de niveau $k \in \mathbb{R}$ *le sous-ensemble de* D *défini par :*

$$L_k = f^{-1}(k) = \{(x, y) \in D \quad tel\ que \quad f(x, y) = k\} \subset D$$

C'est à dire que tous les points de la ligne de niveau k sont envoyés vers la valeur $k \in \mathbb{R}$. En particulier, si $k \notin \mathrm{Im}(f)$, alors $L_k = \varnothing$ (ensemble vide), et nous avons la proposition suivante simple à démontrer :

$$L_k \neq L_{k'} \qquad \text{si et seulement si} \qquad k \neq k'$$



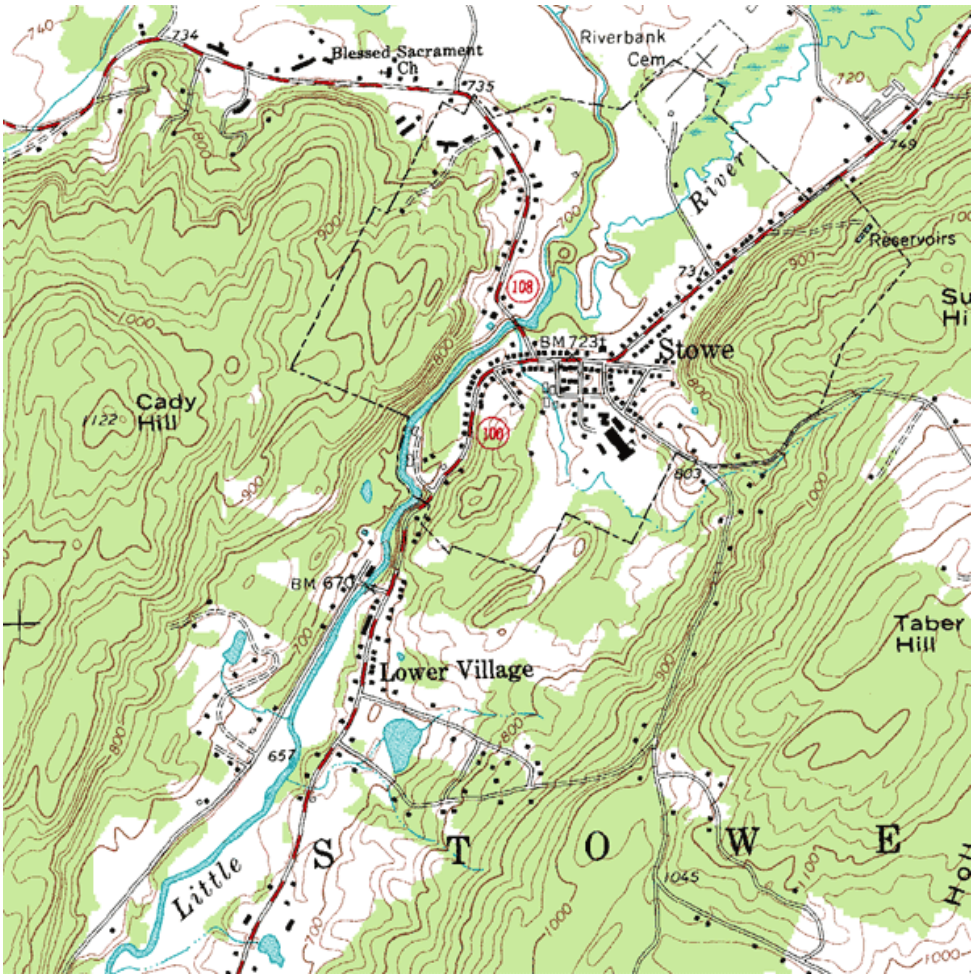
Une ligne de niveau est une courbe algébrique dans le plan $\mathbb{R}^2$, donc peut être un point, une courbe ou une union de courbes (cas plus compliqués), voire un plan (si la fonction est constante). Si la fonction f est continue alors toutes ses lignes de niveau sont fermées dans D, puisque le point $\{k\}$ est fermé dans $\mathbb{R}$ donc sa préimage par une fonction continue est fermée (voir la Proposition 1.63).

Proposition 1.66. *Les lignes de niveau de la fonction* $f : D \to \mathbb{R}$ *forment une partition de* D*, c'est à dire :*

$$\bigcup_{k \in \mathbb{R}} L_k = D \qquad et \qquad L_k = L_{k'} \ si\ et\ seulement\ si\ k = k'$$

Démonstration. Cela vient du fait que les lignes de niveau de la fonction f sont des classes d'équivalence c'est à dire $L_k \neq L_{k'}$ si et seulement si $k \neq k'$. □

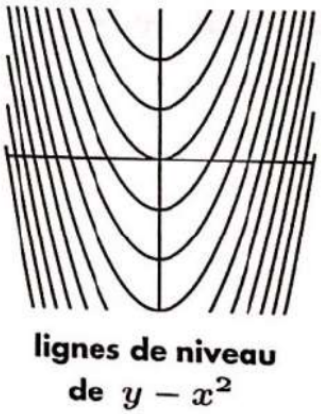
Remarque 1.67. Le nom "ligne de niveau" vient de la notion éponyme sur les cartes de randonnée. On voit tracées les lignes de niveau de la fonction Altitude qui donne, pour tout point (x, y) de la Terre, son altitude.

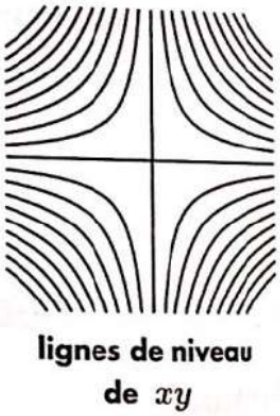


Exemple 1.68. La fonction $f : (x, y) \mapsto x^2 + y^2$ est à valeurs dans $\mathbb{R}_+$. Soit $k \in \mathbb{R}$. La ligne de niveau k est le sous-ensemble de $\mathbb{R}^2$ défini par :

$$L_k = \{(x, y) \in \mathbb{R}^2 \text{ tel que } x^2 + y^2 = k\}$$

Si $k < 0$, $l_k = \emptyset$. Si $k = 0$, $L_k = \{(0, 0)\}$, c'est un point. Si $k > 0$, L_k est le cercle de centre l'origine et de rayon $\sqrt{k}$.





Exemple 1.69. Prenons la fonction $f(x, y) = \ln(1 + x + y)$ de l'exemple 1.11, dont on a vu le domaine de définition, et dont on sait qu'elle est continue. Alors soit $k \in \mathbb{R}$, l'équation $\ln(1 + x + y) = k$ se résout par $1 + x + y = e^k$ c'est à dire la droite d'équation $y = -x + e^k - 1$. Les lignes de niveau de la fonction f sont donc les droites parallèle à la droite $y = -x - 1$, translatées positivement par e^k. Elles sont toutes incluses dans le domaine de définition de la fonction f, défini par l'ensemble des points (x, y) tels que $y > -x - 1$.

Exemple 1.70. Prenons la fonction de l'exemple 1.12, $g(x, y) = \exp\left(\frac{x+y}{x^2-y}\right)$, les lignes de niveau sont définies par le choix d'un $k \in \mathbb{R}$. Si $k \leq 0$ alors les lignes de niveau sont vides. Si $k > 0$, on on a :

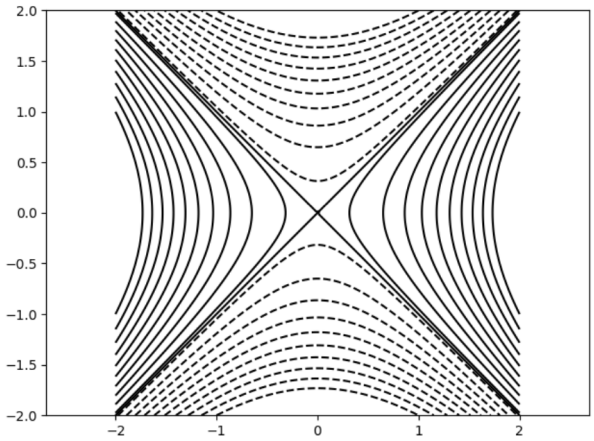
$$\exp\left(\frac{x+y}{x^2-y}\right) = k \qquad \Longleftrightarrow \qquad x + y = \ln(k)(x^2 - y) \text{ et } x^2 \neq y$$

C'est une équation algébrique à résoudre (compliqué), cela donne une courbe algébrique. On peut la ré-écrire sous la forme :

$$Ax^2 + 2Dx + 2Ey = 0 \qquad \text{avec } A = \ln(k), D = -\frac{1}{2}, E = -\frac{\ln(k) + 1}{2}$$

Alors, la technique pour résoudre cette équation nous dit que si $k = 1$, la ligne de niveau est une droite d'équation $y = -x$, si $k = 1/e$ alors la ligne de niveau est l'union de deux droites parallèles $x = 1$ et $x = 0$. En dehors de ces deux cas, la ligne de niveau est une parabole d'équation $y = \frac{\ln(k)}{\ln(k)+1}x^2$.

Exemple 1.71. Pour le paraboloïde elliptique, les lignes de niveau sont des cercles de rayon $\sqrt{k}$; en particulier, si $k < 0$, la ligne de niveau est vide. Pour le paraboloïdes hyperboliques, si $k \neq 0$, la ligne de niveau est une hyperbole (l'union de fonctions de type $y = 1/x$), et si $k = 0$, la ligne de niveau est l'union de deux droites sécantes d'équation $y = \pm x$ (les lignes en pointillé sont les lignes de niveau négatives).



Exemple 1.72. Nous prenons la fonction suivante :

$$\begin{aligned} f :]0, +\infty[\times\mathbb{R} &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \ln(x) + y^2 \end{aligned}$$

Le domaine de définition est le demi-plan droit ouvert. Les lignes de niveaux sont les courbes de ce demi-plan qui satisfont, pour tout $k \in \mathbb{R}$:

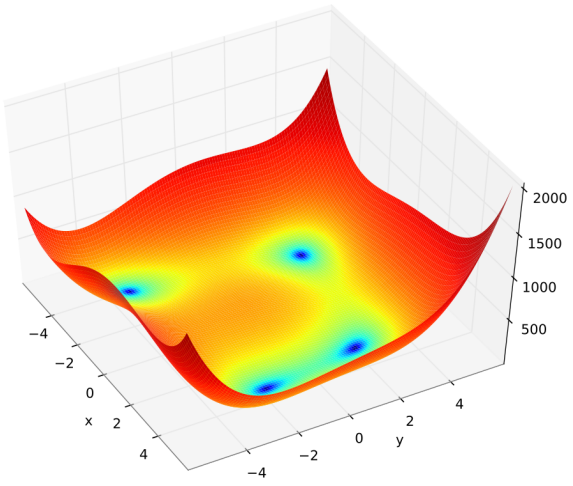
$$L_k = f^{-1}(k) = \left\{(x, y) \in D \text{ tel que } \ln(x) + y^2 = k\right\} = \left\{(x, y) \in D \text{ tel que } x = e^{k-y^2}\right\}$$

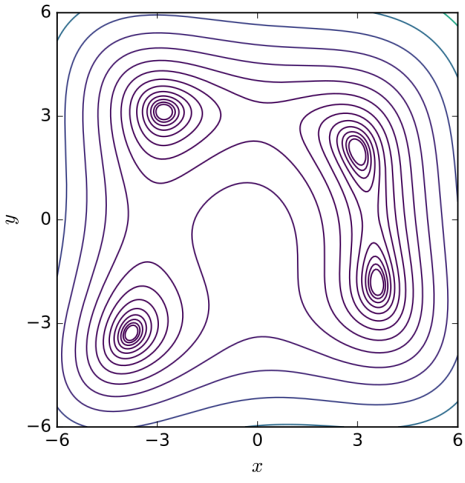
La fonction $y \mapsto e^{-y^2}$ est bien connue. Le graphe de cette courbe est celui de $x \mapsto e^{-x^2}$, pivoté de $-\frac{\pi}{2}$. Les courbes de niveau de la fonction f sont alors des homothéties de cette courbe, d'un ratio e^k.

Exemple 1.73. La fonction de Himmelblau (1924-2011) est utilisée pour tester l'optimization d'algorithmes. Elle est donnée par :

$$f(x, y) = (x^2 + y - 11)^2 + (x + y^2 - 7)^2$$

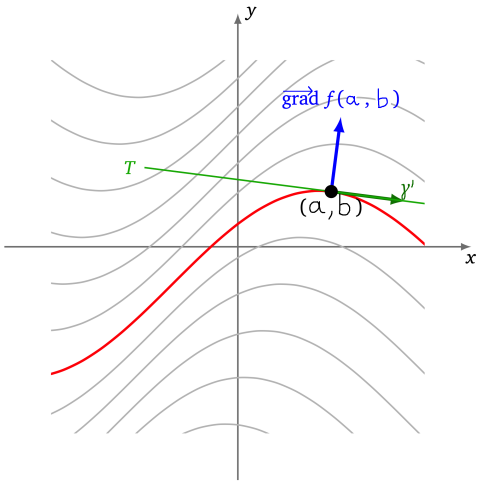
On peut donner les représentations suivantes en 3 dimensions et ses lignes de niveau dans un graphe avec échelle logarithmique (les images viennent de Wikipedia) :

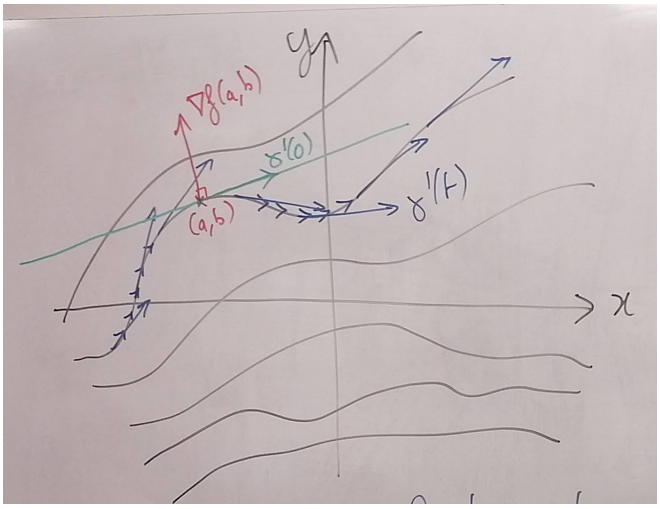




Remarque 1.74. On rappelle qu'un *point critique* de la fonction $f : D \to \mathbb{R}$ est un point où la différentielle en ce point est nulle, c'est à dire : $(a, b) \in \mathring{D}$ est un point critique si $df_{(a,b)} : \mathbb{R}^2 \to \mathbb{R}$ est l'application nulle. Il est équivalent de dire que le vecteur gradient $\nabla f(a, b)$ est nul.

Proposition 1.75. *Soit $A = (a, b)$ un point du plan qui n'est pas un point critique de f. Alors le vecteur gradient en ce point $\nabla f(a, b)$ est orthogonal à la ligne de niveau passant par ce point. Autrement dit, le champ de vecteurs gradient de f est orthogonal aux lignes de niveau partout où $\nabla f \neq \binom{0}{0}$.*





Expliquons cette proposition. Choisissons un point (a, b) qui n'est pas un point critique de f, donc en lequel le vecteur gradient n'est pas zéro. Posons $k = f(a, b)$ et désignons par L_k la ligne de niveau k, qui passe donc par (a, b). L'idée de la preuve de la Proposition 1.75 est que nous devons obtenir l'équation de la tangente à la ligne de niveau L_k au point (a, b), et vérifier que cette droite est orthogonale au vecteur gradient au point (a, b). Pour cela nous devons trouver une paramétrisation locale de la ligne de niveau, autour du point (a, b). Supposons que nous ayons une telle paramétrisation. Notons

$$\begin{array}{rcl} \gamma : \ ]-\epsilon, \epsilon[& \longrightarrow & D \\ t & \longmapsto & \gamma(t) = (u(t), v(t)) \end{array}$$

le chemin paramétrisant cette courbe dans un voisinage de (a, b) et tel que $\gamma(0) = (a, b)$. Les fonctions u et v sont les composantes du chemin γ et sont des fonctions à une variable, dérivables. La tangente à la courbe paramétrée γ en (a, b) est la droite passant par (a, b) et de vecteur directeur $\binom{u'(0)}{v'(0)}$. C'est le vecteur vitesse du chemin $t \longmapsto \gamma(t)$ au temps $t = 0$. Un vecteur $\vec{v} = \binom{h}{k}$ est *orthogonal* à la courbe au point $(a, b) = \gamma(0)$ si :

$$\left\langle \begin{pmatrix} u'(0) \\ v'(0) \end{pmatrix}, \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle = 0$$

C'est cette équation que nous voudrons montrer pour prouver la Proposition 1.75. Maintenant, pour trouver une paramétrisation de la courbe de niveau L_k dans un voisinage du point (a, b), la meilleure façon de procéder c'est d'arriver à décrire localement L_k comme le graphe d'une fonction, soit de la variable x (comme en analyse classique), soit de la variable y (c'est nouveau !). Le chapitre suivant est dédié à expliquer les conditions sous lesquelles nous pouvons faire cela.

Exemple 1.76. Si nous prenons la fonction classique $f : (x, y) \mapsto x^2 + y^2$, les lignes de niveaux sont des cercles. Prenons $(a, b) \neq (0, 0)$ (seul point critique de la fonction). Nous avons que $\nabla f(a, b) = \binom{2a}{2b}$. Soit $k = a^2 + b^2$. La ligne de niveau k est le cercle de centre l'origine et de rayon $\sqrt{k}$. Le point (a, b) appartient à la ligne de niveau k. On peut paramétrer ce cercle grâce aux coordonnées polaires :

$$\begin{array}{rcl} \gamma : \ [-\pi, \pi] & \longrightarrow & \mathbb{R}^2 \\ t & \longmapsto & \gamma(t) = \left(\sqrt{k}\cos(t), \sqrt{k}\sin(t)\right) \end{array}$$

Le point (a, b) est sur ce cercle donc il existe t_0 tel que $a = \sqrt{k}\cos(t_0)$ et $b = \sqrt{k}\sin(t_0)$. Le vecteur vitesse du chemin γ au temps t_0 est donné par :

$$\overrightarrow{\gamma'(t_0)} = \begin{pmatrix} -\sqrt{k}\sin(t_0) \\ \sqrt{k}\cos(t_0) \end{pmatrix} = \begin{pmatrix} -b \\ a \end{pmatrix}$$

Nous avons donc :

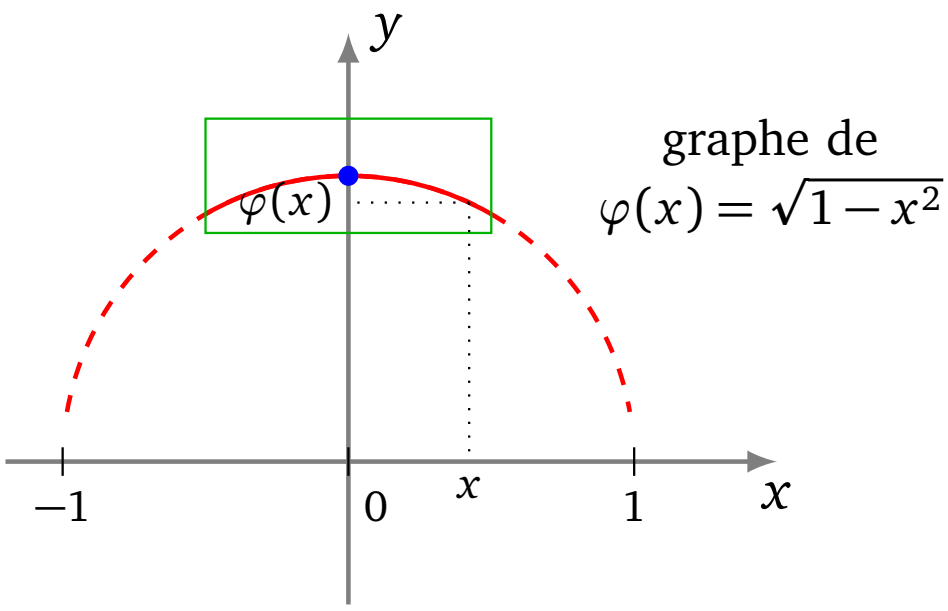
$$\left\langle \nabla f(a,b), \overrightarrow{\gamma'(t_0)} \right\rangle = \left\langle \begin{pmatrix} 2a \\ 2b \end{pmatrix}, \begin{pmatrix} -b \\ a \end{pmatrix} \right\rangle = -2ab + 2ab = 0$$

On retrouve bien que le vecteur $\nabla f(a, b)$ est orthogonal au vecteur $\overrightarrow{\gamma'(t_0)}$ (le vecteur tangent à la ligne de niveau passant en (a, b)).

1.4 Le théorème des fonctions implicites

Le théorème des fonctions implicites est un théorème compliqué à formuler, mais dont les idées sous-jacentes sont importantes. L'idée est de voir les lignes de niveau d'une fonction à deux variables x et y comme des courbes paramétrées localement, par x ou y. Prenons la ligne de niveau 1 de la fonction $f(x, y) = x^2 + y^2$, c'est le cercle de centre $(0, 0)$ et de rayon 1, noté S^1. Autour du point $(0, 1)$, une portion d'arc de cercle S^1 peut être vu comme le graphe d'une fonction (dérivable) de la variable x :

$$\begin{aligned} \varphi : \ ]-r, r[&\longrightarrow \mathbb{R} \\ x &\longmapsto \sqrt{1-x^2} \end{aligned}$$



On peut paramétrer d'autres arcs de cercle par rapport à la variable x, mais la condition nécessaire est que l'arc de cercle en question (ouvert aux extrémités) doit être vu comme le graphe d'une fonction (dérivable) de la variable x. En particulier, ce graphe ne peut pas avoir de tangente verticale (ça n'arrive jamais pour une fonction de la variable x). On peut donc paramétrer des arcs de cercles par la variable x que si ces arcs de cercle n'intersectent pas l'axe

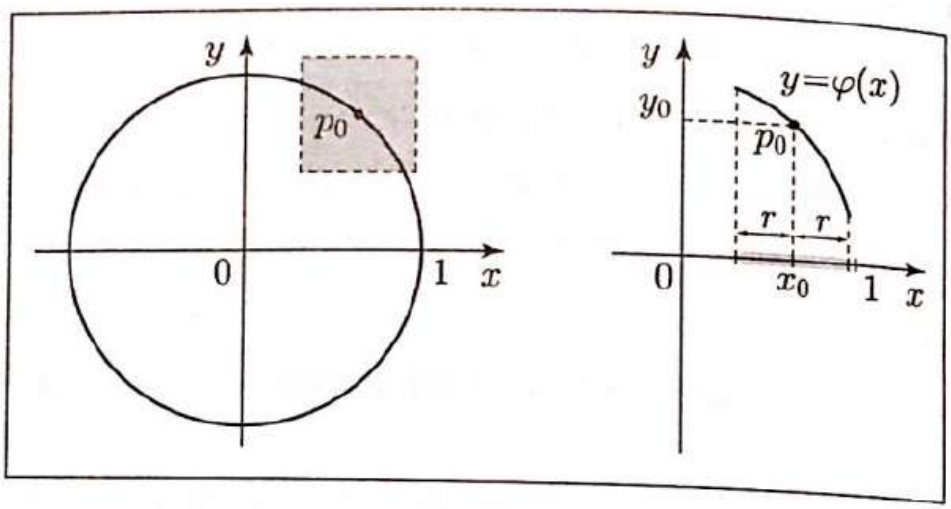
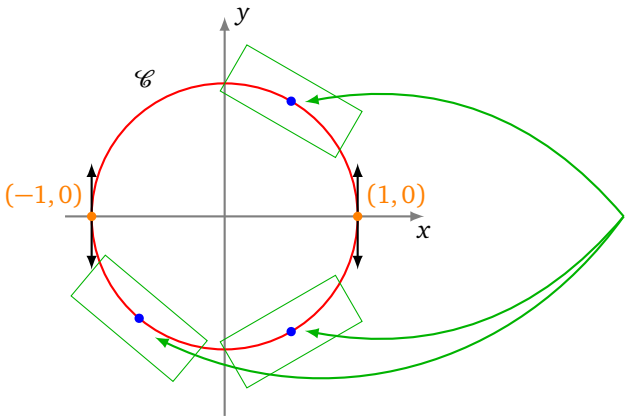


FIGURE 3 – On peut paramétrer un arc de cercle dans le voisinage de p_0 par la fonction $x \mapsto \sqrt{1-x^2}$. C'est la même fonction pour tout le demi-cercle supérieur (ouvert aux extrémités).

des abscisses. Autour des deux points $(1,0)$ et $(-1,0)$, aucun arc de cercle (ouvert aux bords) ne peut s'écrire comme le graphe d'une fonction de x. L'explication, c'est qu'en ces points il y a une tangente verticale, et que le graphe d'une fonction de x dérivable n'a jamais de tangente verticale.

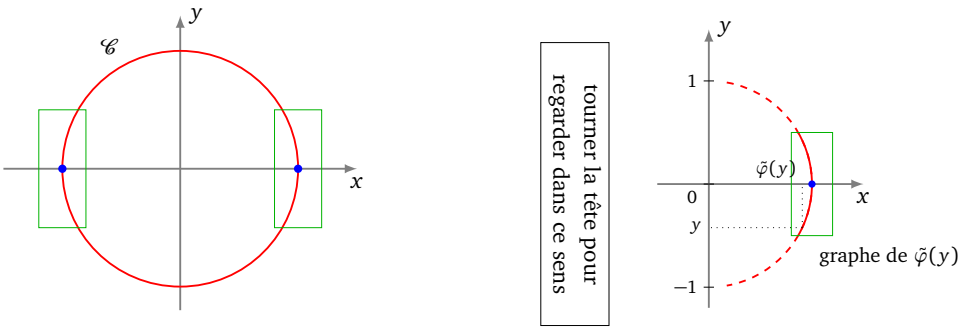


exemples de points de $\mathscr{C}$ autour desquels y s'exprime comme une fonction de x

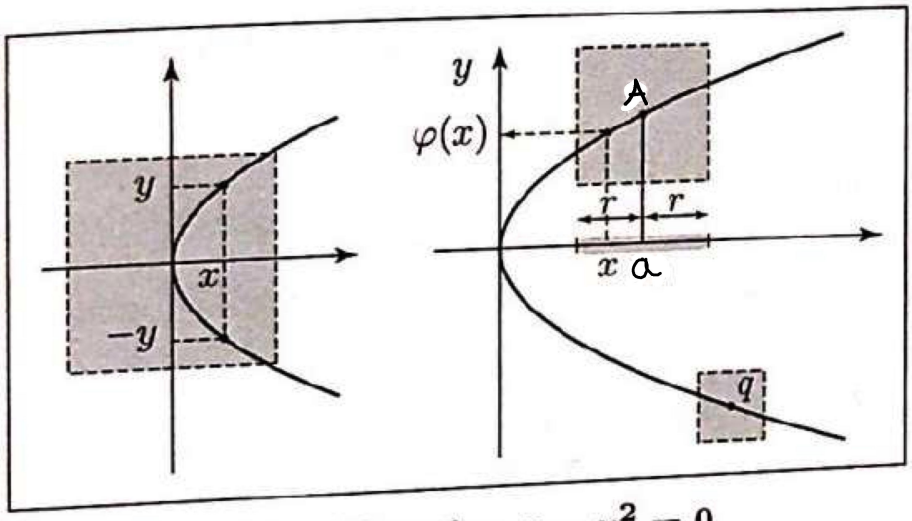
Dans tous les points où on peut théoriquement exprimer l'arc de cercle comme le graphe d'une fonction de x, il n'est parfois pas très simple d'exprimer cette fonction explicitement. Cependant l'existence d'une telle fonction est garanti, c'est pourquoi on parle de théorème des fonctions *implicites*. Dans le cas du cercle, par chance, la fonction est assez simple donc nous pouvons toujours connaitre la fonction φ : le demi-cercle supérieur (ouvert aux extrémités) est le graphe de la fonction $x \mapsto \sqrt{1-x^2}$ tandis que le demi-cercle inférieur (ouvert aux extrémités) est le graphe de la fonction $x \mapsto -\sqrt{1-x^2}$. Notons que dans les deux cas, la fonction est de classe $\mathcal{C}^1$ car on exclut les point $(1,0)$ et $(-1,0)$.

Maintenant que se passe-t-il aux points qui coupent l'axe des abscisses ? On peut tout simplement considérer la portion de cercle autour de $(1,0)$ comme le graphe d'une fonction de la variable y, c'est à dire $y \mapsto \widetilde{\varphi}(y)$. C'est un peu déroutant car les rôles de x et de y sont inversés : au lieu que y soit fonction de x, ici x est fonction de y. C'est comme si on avait tourné la feuille

de 90°. Et en fait, on peut exprimer les arcs de cercles comme des fonctions de y partout...sauf aux point où la tangente au cercle est horizontale, c'est à dire en $(0,1)$ et $(0,-1)$.

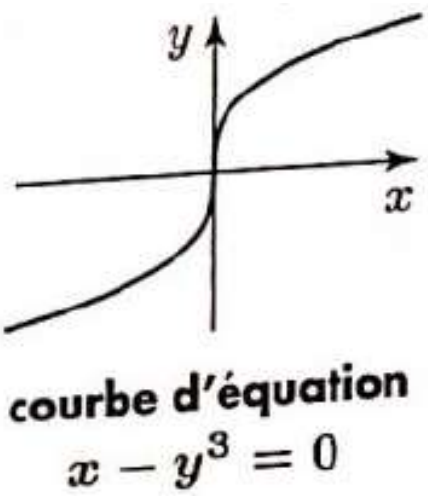


Exemple 1.77. Autre exemple : on regarde la courbe d'équation $y^2 = x$ c'est à dire la parabole horizontale vers la droite, d'axe Ox et de sommet l'origine. Prenons un point (a,b) avec $b > 0$ qui se trouve sur la courbe ; cela veut dire que $a = b^2$. Choisissons un nombre $r > 0$ tel que $0 < r < a$. Pour tout $x \in]a-r, a+r[$, la branche dans le demi plan supérieur de la parabole peut s'écrire comme $y = \varphi(x) = \sqrt{x}$, une fonction de classe $\mathcal{C}^1$ (c'est à dire dont chaque dérivée partielle est une fonction continue) dans le segment $]a-r, a+r[$. On dit que l'équation $x - y^2 = 0$ se résout en x sur l'intervalle $]a-r, a+r[$.



Maintenant, si on prend $A = (0,0)$ l'origine, on voit que n'importe quelle partie de la courbe centrée en l'origine est doublée en haut et en bas. On ne peut pas résoudre en x, mais on peut résoudre l'équation $x - y^2 = 0$ en y. C'est à dire qu'on peut écrire $x = \tilde{\varphi}(y) = y^2$. C'est une paramétrisation de la courbe parabolique à l'origine, de classe $\mathcal{C}^1$.

Exemple 1.78. Prenons maintenant un dernier exemple, la courbe algébrique d'équation $x - y^3 = 0$. Alors dans ce cas, même si la dérivée le long de la courbe est verticale à l'origine, on peut quand même paramétrer la courbe par x ou par y, indifféremment, par contre la paramétrisation par x n'est pas de classe $\mathcal{C}^1$. En effet, on peut écrire y en fonction de x par la formule $y = \varphi(x) = \sqrt[3]{x}$, mais cette fonction n'est pas dérivable en 0. Par contre la fonction $x = \widetilde{\varphi}(y) = y^3$ est de classe $\mathcal{C}^1$ en 0. La différence vient du fait que la dérivée partielle de la fonction $(x, y) \mapsto x - y^3$ par rapport à y s'annule en $y = 0$, mais pas la dérivée partielle par rapport à x.



Théorème des Fonctions Implicites. *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$ (c'est à dire dont chaque dérivée partielle est une fonction continue). Soit $A = (a, b)$ un point de la ligne de niveau k de f (c'est à dire $f(a, b) = k$) qui se trouve à l'intérieur $\mathring{D}$, et tel que :*

$$\frac{\partial f}{\partial y}(a, b) \neq 0 \tag{1.5}$$

Alors il existe un intervalle ouvert I contenant a, un intervalle ouvert J contenant b, et une unique fonction de classe $\mathcal{C}^1$ dénotée $\varphi : I \to J$ telle que :

$$\varphi(a) = b \qquad et \qquad \forall\, (x, y) \in I \times J, \quad f(x, y) = k \iff y = \varphi(x)$$

ce qui veut dire que, pour tout $x \in I$, on a $f(x, \varphi(x)) = k$.

Remarque 1.79. La condition (1.5) signifie que la tangente à la ligne de niveau au point (a, b) n'est *pas* verticale. Si elle est verticale en ce point, un énoncé symétrique du théorème existe pour la dérivée de f par rapport à x. Donc si jamais la dérivée de f par rapport à y est nulle en (a, b), c'est à dire si la tangente à la ligne de niveau au point (a, b) est verticale, on regarde la dérivée de f par rapport à x et on applique le théorème.

Théorème des Fonctions Implicites (2nde variable). *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe $\mathcal{C}^1$. Soit $A = (a, b)$ un point de la ligne de niveau k de f (c'est à dire $f(a, b) = k$) qui se trouve à l'intérieur $\mathring{D}$, et tel que :*

$$\frac{\partial f}{\partial x}(a, b) \neq 0$$

Alors il existe un intervalle ouvert I contenant a, un intervalle ouvert J contenant b, et une unique fonction de classe $\mathcal{C}^1$ dénotée $\widetilde{\varphi} : I \to J$ telle que :

$$\widetilde{\varphi}(b) = a \qquad et \qquad \forall\, (x, y) \in I \times J, \quad f(x, y) = k \iff x = \widetilde{\varphi}(y)$$

ce qui veut dire que, pour tout $y \in I$, on a $f(\widetilde{\varphi}(y), y) = k$.

Exemple 1.80. Dans l'exemple de la sphère, on pose $f(x, y) = x^2 + y^2$, et on regarde la ligne de niveau 1 (cercle de rayon 1). Dans ce cas on avait bien $\frac{\partial f}{\partial y}(a, b) \neq 0$ partout dans le demi-plan supérieur et le demi-plan inférieur, mais pas aux deux points $(1, 0)$ et $(-1, 0)$, où les tangentes sont verticales. En ces points, et plus généralement dans tout le demi-plan droit et le demi-plan gauche, on a au contraire $\frac{\partial f}{\partial x}(a, b) \neq 0$. Conclusion : pour les points du demi-plan supérieur (resp. inférieur), le théorème nous dit qu'on aura une fonction de classe $\mathcal{C}^1$ $\varphi : x \mapsto \sqrt{1 - x^2}$ (resp. $\varphi : x \mapsto -\sqrt{1 - x^2}$) telle que $y = \varphi(x)$, tandis que pour les point du demi-plan droit (resp. gauche), on aura une fonction $\widetilde{\varphi} : y \mapsto \sqrt{1 - y^2}$ (resp. $\widetilde{\varphi} : y \mapsto -\sqrt{1 - y^2}$) telle que $x = \widetilde{\varphi}(y)$.

Exemple 1.81. Dans l'exemple de la parabole horizontale 1.77, on pose $f(x,y) = x - y^2$ et on regarde la ligne de niveau 0. Alors la dérivée par rapport à x de la fonction f ne s'annule jamais, cela veut dire qu'on peut toujours exprimer x en fonction de y (oui c'est la définition de la fonction même). Par contre, la dérivée de f par rapport à y s'annule en $(0,0)$ et donc cela veut dire qu'on peut exprimer y en fonction de x grâce à une fonction de classe $\mathcal{C}^1$ que pour $x \neq 0$.

Exemple 1.82. Dans l'Exemple 1.78, on pose comme fonction $f(x,y) = x - y^3$ et on regarde la ligne de niveau 0. La dérivée par rapport à y s'annule en $(0,0)$ donc, bien qu'on puisse exprimer y en fonction de x partout sur $\mathbb{R}$, on sait que la fonction $y = \sqrt[3]{x}$ n'est pas de classe $\mathcal{C}^1$ en 0. Au contraire, comme la dérivée de f par rapport à x n'est jamais nulle, on peut exprimer x comme une fonction de classe $\mathcal{C}^1$ de la variable y, c'est précisément f.

Démonstration. En redéfinissant la fonction $f : (x,y) \mapsto f(x,y)$ en retranchant k – c'est à dire en étudiant la fonction $g : (x,y) \mapsto f(x,y) - k$, la ligne de niveau k de la fonction f est la ligne de niveau 0 de la fonction g. On peut donc toujours supposer que $k = 0$. Prenons le point (a,b) sur cette ligne de niveau 0. En outre, on suppose que $\frac{\partial f}{\partial y}(a,b) > 0$ car, si jamais $\frac{\partial f}{\partial y}(a,b) < 0$, on peut se ramener au premier cas en étudiant $-f$ plutôt que f, puisque si $\frac{\partial f}{\partial y}(a,b) < 0$, alors $\frac{\partial(-f)}{\partial y}(a,b) > 0$.

L'idée de la preuve est de dire que si on zoome assez proche du point (a,b), il existe un rectangle $I \times J$ centré en (a,b) tel que, pour tout point de la frontière inférieure du rectangle, $f(x,y) < 0$, et pour tout point de la frontière supérieure du rectangle, $f(x,y) > 0$ (voir la Figure 4). Alors, si on fixe une valeur de x, par le théorème des valeurs intermédiaires (et stricte croissance de f lorsqu'on augmente y), il existe un unique y (qui dépend de x) tel que $(x,y) \in I \times J$ et $f(x,y) = 0$. La correspondance injective $x \mapsto y$ définit une fonction $\varphi : I \to J$ qui a pour propriété de paramétrer la ligne de niveau 0 (par sa définition même). On peut donc écrire que $(x,y) \in L_0$ si et seulement si $y = \varphi(x)$. C'est le résultat du théorème.

Passons à la preuve rigoureuse. Rappelons que $\frac{\partial f}{\partial y}(a,b) > 0$. Comme la fonction f est de classe $\mathcal{C}^1$, la fonction dérivée $(x,y) \mapsto \frac{\partial f}{\partial y}(x,y)$ est continue en les deux variables. Donc il existe un nombre réel $\rho > 0$ tel que $\frac{\partial f}{\partial y} > 0$ dans un voisinage de (a,b). Or un voisinage d'un point dans $\mathbb{R}^2$ se caractérise par le fait de contenir une boule ouverte centrée en (a,b). Selon la norme choisie, la boule ouverte est un disque ouvert (norme euclidienne), ou bien un rectangle ouvert (nome infinie), etc. Toutes les normes son équivalentes en dimension finies donc nous choisissons la norme la plus pratique, ici la norme infinie. Ainsi, nous avons qu'il existe $\rho > 0$ tel que (c'est le rectangle vert sur la Figure 4) :

$$\forall\, (x,y) \in [a-\rho, a+\rho] \times [b-\rho, b+\rho] \qquad \frac{\partial f}{\partial y}(x,y) > 0$$

De cette dérivée sur la seconde variable strictement positive, on en déduit que pour tout $x \in [a-\rho, a+\rho]$ la fonction à partielle $f_x : y \mapsto f(x,y)$ est strictement croissante sur le segment vertical $[b-\rho, b+\rho]$.

Remarque 1.83. Autrement dit, la restriction au rectangle $[a-\rho, a+\rho] \times [b-\rho, b+\rho]$ permet de se restreindre à un voisinage du point (a,b) dans lequel la ligne de niveau 0 ne fait pas de boucle (comme les méandres du Mississipi). Par contre, nous ne savons pas encore si la ligne de niveau de sort pas du rectangle *avant* d'atteindre le côté droit du rectangle, à l'abscisse $a+\rho$ (c'est ce qui est illustré sur la Figure 4).

Résolvons le problème soulevé par la précédente remarque. Comme on a $f(a,b) = 0$, et que $f_x : y \mapsto f(x,y)$ est strictement croissante sur $[b-\rho, b+\rho]$, on en déduit que $f(a, b-\rho) < 0$ et $f(a, b+\rho) > 0$. Comme la fonction f est continue, la propriété d'être strictement supérieur (ou

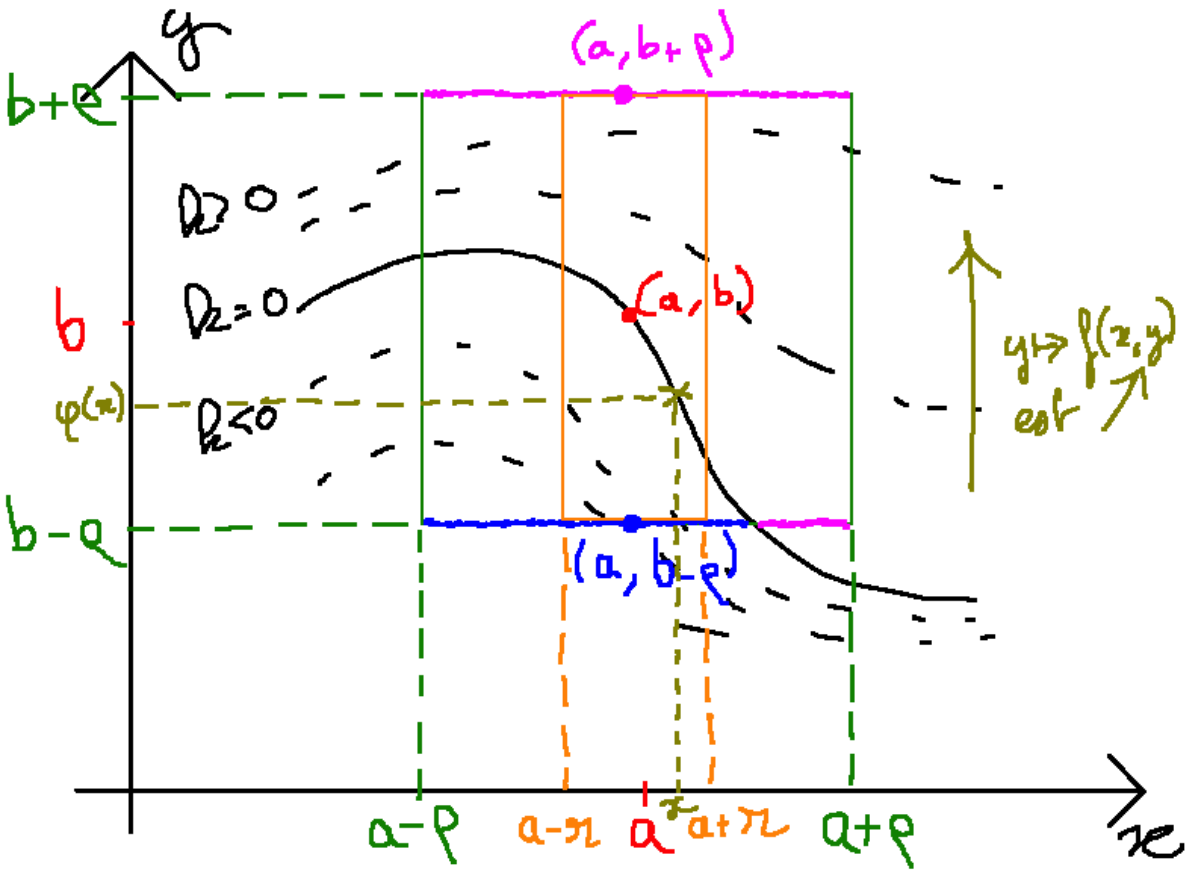


FIGURE 4 – Illustration de la démarche du théorème des fonctions implicites. Le premier choix de $\rho > 0$ se base sur le fait que $\frac{\partial f}{\partial y}(a,b) > 0$. Cependant ce ρ peut encore être trop grand car on veut que tout le bord inférieur du rectangle satisfasse $f(\text{bord inférieur}) < 0$ et tout le bord supérieur satisfasse $f(\text{bord supérieur}) > 0$. On choisit alors $r > 0$ assez petit de façon à avoir cette condition (rectangle orange). Ensuite il suffit d'utiliser le théorème des valeurs intermédiaires : à chaque $x \in]a-r, a+r[$ correspond un unique $y \in]b-\rho, b+\rho[$ tel que $f(x,y) = 0$. Pour chaque x, ce y est noté $\varphi(x)$.

inférieur) à 0 est une propriété locale. Ainsi, il existe un voisinage (dans $\mathbb{R}$) du point a tel que pour tout x de ce voisinage, $f(x, b-\rho) < 0$ (en bleu sur la Figure 4). Idem pour $f(x, b+\rho) > 0$ (en mauve en haut du rectangle sur la Figure 4). Formellement, cela établit qu'il existe $0 < r \leq \rho$ tel que, pour tout $x \in [a-r, a+r]$, $f(x, b-\rho) < 0$ et $f(x, b+\rho) > 0$ (le bord inférieur et supérieur du rectangle orange sur la Figure 4).

Fixons $x \in [a-r, a+r]$. Comme f_x est strictement croissante et $f(x, b-\rho) < 0$ et $f(x, b+\rho) > 0$, d'après le théorème des valeurs intermédiaires, il existe un unique nombre $y \in]b-\rho, b+\rho[$ tel que $f(x,y) = 0$. On note ce y particulier, associé à chaque x, $\varphi(x)$. On a donc une correspondance injective entre $x \in]a-r, a+r[$ et $y = \varphi(x) \in]b-\rho, b+\rho[$, telle que $f(x, \varphi(x)) = 0$. En particulier, pour $x = a$, comme $f(a,b) = 0$ et que la correspondance est injective, on en déduit que $\varphi(a) = b$. Cette correspondance injective définit une application injective

$$\begin{aligned} \varphi : \ I &\longrightarrow J \\ x &\longmapsto \varphi(x) \end{aligned}$$

où on a noté $I =]a-r, a+r[$ et $J =]b-\rho, b+\rho[$. On peut montrer que c'est une fonction de

classe $\mathcal{C}^1$ (compliqué). C'est l'application qu'on recherche. □

Remarque 1.84. Le théorème s'appelle Théorème des Fonctions *Implicites* car on ne connait pas la fonction φ explicitement. On a prouvé qu'elle existe, mais on n'a pas donné de formule précise. Ce n'est pas un problème car c'est sa dérivée qui nous intéresse, et celle-ci nous la connaissons.

Corollaire 1.85. *Soit $\varphi : I \to J$ la fonction définie dans le Théorème des Fonctions Implicites. Alors on a pour tout $x \in I$:*

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))} \tag{1.6}$$

Démonstration. Pour retrouver facilement la formule (1.6), il suffit de dériver la fonction $x \mapsto f(x, \varphi(x))$ par rapport à la variable x, et de noter que pour tout $x \in I$, $f(x, \varphi(x)) = 0$ donc la dérivée est constante, nulle :

$$0 = \frac{df}{dx}(x, \varphi(x)) = \frac{\partial f}{\partial x}(x, \varphi(x)) + \varphi'(x)\frac{\partial f}{\partial y}(x, \varphi(x))$$

Ensuite, on sait qu'au point (a, b) au moins, $\frac{\partial f}{\partial y}(a, b) \neq 0$. Et comme f est de classe $\mathcal{C}^1$, la condition (1.5) reste vraie pour tout $(x, \varphi(x))$ dans un voisinage de (a, b), et donc cela garantit qu'on peut diviser par $\frac{\partial f}{\partial y}(x, \varphi(x))$ pour obtenir l'équation (1.6). □

Corollaire 1.86. *Soit (a, b) un point du plan qui n'est pas un point critique de f et $k = f(a, b)$. Alors l'équation de la droite tangente à la ligne de niveau k au point (a, b) est donnée par :*

$$(y - b)\frac{\partial f}{\partial y}(a, b) + (x - a)\frac{\partial f}{\partial x}(a, b) = 0 \tag{1.7}$$

Remarque 1.87. Autrement dit, pour écrire la formule (1.7) sous une forme mieux connue, si on suppose que $\frac{\partial f}{\partial y}(a, b) \neq 0$, et si on note $\varphi : I \to J$ la fonction donnée par le Théorème des Fonctions Implicites, l'équation de la droite se lit :

$$y = \varphi'(a)x + b - a\varphi'(a)$$

Notons que les coordonnées dépendent du choix de la base de $\mathbb{R}^2$ choisie, mais la ligne de niveau et la droite tangent en un point sont des objets géométriques donc ils ne changent pas par changement de base (seules leur description par coordonnées changent). L'idée du théorème des fonctions implicites est d'avoir un bon choix de coordonnées pour décrire la ligne de niveau dans le voisinage d'un point (a, b) comme le graphe d'une fonction, qu'on peut donc dériver et obtenir la tangente.

Démonstration. Si $\frac{\partial f}{\partial y}(a, b) = 0$ alors la tangente à la ligne de niveau k au point (a, b) est verticale (le point n'est pas un point critique donc la tangente existe bien). L'équation de la droite tangente est donc $x = a$, ce qui correspond bien à l'équation (1.7) pour $\frac{\partial f}{\partial y}(a, b) = 0$ et $\frac{\partial f}{\partial x}(a, b) \neq 0$.

Si par contre $\frac{\partial f}{\partial y}(a, b) \neq 0$, alors par le Théorème des Fonctions Implicites, il existe I (resp. J) intervalle ouvert contentant x (resp. y) et une unique fonction $\varphi : I \to J$ telle que $f(x, \varphi(x)) = k$ pour tout $x \in I$. Alors la droite tangente au graphe de la fonction φ au point $x = a$ est précisément la droite tangente à la ligne de niveau k au point (a, b). L'équation de la droite tangente au graphe de la fonction φ au point $x = a$ est donnée par :

$$y = (x - a)\varphi'(a) + b$$

De là, en rappelant la formule (1.6) pour la dérivée de φ, et en l'évaluant en $x = a$ tout en sachant que $\varphi(a) = b$, on obtient :

$$\varphi'(a) = -\frac{\frac{\partial f}{\partial x}(a,b)}{\frac{\partial f}{\partial y}(a,b)}$$

Cela nous donne donc :

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

Ce qui donne bien l'équation (1.7). □

Définition 1.88. *Un vecteur directeur d'une droite donnée est n'importe quel vecteur non-nul $\vec{X}$ dont la direction est la même que la droite. Un vecteur orthogonal (ou normal) à cette droite est un vecteur non-nul $\vec{Y}$ tel que $\left\langle \vec{X}, \vec{Y} \right\rangle = 0$, pour n'importe quel vecteur directeur $\vec{X}$.*

Pour une droite verticale, donc d'équation $x =$ constante, alors un vecteur directeur est colinéaire au vecteur vertical $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$, tandis qu'un vecteur normal est colinéaire au vecteur horizontal $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$. Dans les deux cas il en existe une infinité. Pour une droite non-verticale, c'est à dire d'équation $y = \alpha x + \beta$, un vecteur directeur est colinéaire au vecteur $\left(\begin{smallmatrix} 1 \\ \alpha \end{smallmatrix}\right)$ tandis qu'un vecteur normal est colinéaire au vecteur $\left(\begin{smallmatrix} -\alpha \\ 1 \end{smallmatrix}\right)$. On peut en effet vérifier que le produit scalaire de ces deux vecteurs est nul. Dans les deux cas il y en a une infinité. Plus généralement, on a un résultat plus synthétique qui unifie ces deux cas, en observant que toute droite (non verticale et verticale) peut être décrite par un seul type d'équation algébrique, totalement générale :

$$Ax + By + C = 0$$

On retrouve pour $B = 0$, le cas des droites verticales d'équation $x = -\frac{C}{A} =$ constante, tandis que pour $B \neq 0$ – les droites non-verticales – on a $y = -\frac{A}{B}x - \frac{C}{B}$ et donc, en posant $\alpha = -\frac{A}{B}$ et $\beta = -\frac{C}{B}$, on obtient l'équation classique $y = \alpha x + \beta$. Toutes les droites peuvent être décrites par l'équation ci dessus.

Proposition 1.89. *La droite d'équation $Ax + By + C = 0$ admet comme vecteur directeur :*

$$\vec{T} = \begin{pmatrix} B \\ -A \end{pmatrix}$$

et comme vecteur normal :

$$\vec{N} = \begin{pmatrix} A \\ B \end{pmatrix}$$

Dans le cas d'une droite verticale – c'est à dire pour $B = 0$ – le vecteur directeur $\vec{T} = \left(\begin{smallmatrix} 0 \\ -A \end{smallmatrix}\right)$ est colinéaire à $\left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$ qu'on avait retrouvé plus haut, et le vecteur normal $\vec{N} = \left(\begin{smallmatrix} A \\ 0 \end{smallmatrix}\right)$ est colinéaire à $\left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$ comme vu avant. Pour une droite non verticale – c'est à dire pour $B \neq 0$ – la droite d'équation $Ax + By + C = 0$ peut aussi s'écrire $y = \alpha x + \beta$ avec $\alpha = -\frac{A}{B}$ et $\beta = -\frac{C}{B}$. On peut factoriser par B dans le vecteur directeur $\vec{T}$, et on obtient $\vec{T} = B\left(\begin{smallmatrix} 1 \\ -\frac{A}{B} \end{smallmatrix}\right)$, qui est bien colinéaire à $\left(\begin{smallmatrix} 1 \\ \alpha \end{smallmatrix}\right)$, qui est bien le vecteur directeur de la droite d'équation $y = \alpha x + \beta$ introduit plus haut. De même on peut factoriser par B dans le vecteur normal $\vec{N}$ et on obtient $\vec{N} = B\left(\begin{smallmatrix} \frac{A}{B} \\ 1 \end{smallmatrix}\right)$, qui est colinéaire à $\left(\begin{smallmatrix} -\alpha \\ 1 \end{smallmatrix}\right)$ qu'on avait trouvé plus haut. Tout est cohérent. Nous pouvons maintenant démontrer la proposition vue en fin de chapitre précédent :

Proposition 1.90. *Soit (a, b) un point du plan qui n'est pas un point critique de f. Alors le vecteur gradient en ce point $\nabla f(a, b)$ est orthogonal à la ligne de niveau passant par ce point.*

Démonstration. Soit (a,b) un point qui n'est pas un point critique de f et supposons que $f(a,b) = k$. Si $\frac{\partial f}{\partial y}(a,b) = 0$ alors, comme (a,b) n'est pas un point critique de f, d'une part la tangente à la ligne de niveau k au point (a,b) est verticale, et d'autre part nécessairement $\frac{\partial f}{\partial x}(a,b) \neq 0$. Dans ce cas le gradient de f en (a,b) – qui est $\nabla f(a,b) = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) \\ 0 \end{pmatrix}$ – est horizontal. On a donc bien que le gradient est orthogonal à la tangente (verticale) à la ligne de niveau au point (a,b).

Si $\frac{\partial f}{\partial y}(a,b) \neq 0$, alors par le Théorème des Fonctions Implicites, il existe I (resp. J) un intervalle ouvert contentant a (resp. b) et une unique fonction $\varphi : I \to J$ telle que $f(x, \varphi(x)) = k$ pour tout $x \in I$. Autrement dit, la ligne de niveau k dans le voisinage de (a,b), est le graphe de la fonction φ. L'équation de la tangente à ce graphe au point $(a, \varphi(a))$ est donc donné classiquement par :

$$y = \varphi'(a)(x-a) + b$$

D'après la discussion précédant la Proposition 1.89, un vecteur directeur de cette tangente est :

$$\vec{X} = \begin{pmatrix} 1 \\ \varphi'(a) \end{pmatrix} = \begin{pmatrix} 1 \\ -\frac{\frac{\partial f}{\partial x}(a,b)}{\frac{\partial f}{\partial y}(a,b)} \end{pmatrix}$$

Le produit scalaire entre le gradient de f en (a,b) et le vecteur tangent à la ligne de niveau k en (a,b) est donc :

$$\left\langle \nabla f(a,b), \vec{X}_{(a,b)} \right\rangle = df_{(a,b)}\left(\vec{X}_{(a,b)}\right) = 1 \times \frac{\partial f}{\partial x}(a,b) - \frac{\frac{\partial f}{\partial x}(a,b)}{\frac{\partial f}{\partial y}(a,b)} \times \frac{\partial f}{\partial y}(a,b) = 0$$

On a donc bien l'orthogonalité entre le gradient et le vecteur tangent. □

Remarque 1.91. On aurait pu prouver cette Proposition de façon beaucoup plus directe en rappelant que l'équation de la tangente à la ligne de niveau k au point (a,b) est donnée par la formule (1.7). Mettons là sous la forme $Ax + By + C = 0$:

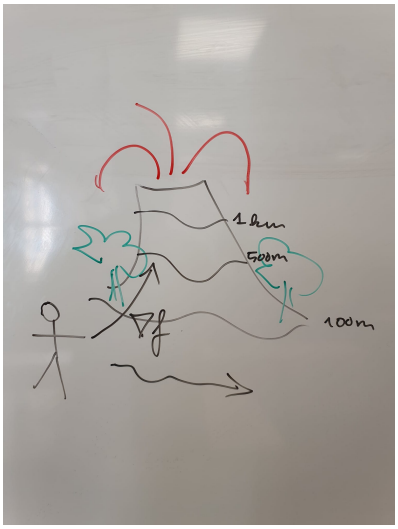
$$\underbrace{\frac{\partial f}{\partial x}(a,b)}_{=A} x + \underbrace{\frac{\partial f}{\partial y}(a,b)}_{=B} y \underbrace{- \left(a\frac{\partial f}{\partial x}(a,b) + b\frac{\partial f}{\partial y}(a,b) \right)}_{=C} = 0$$

On en déduit par la Proposition 1.89 qu'un vecteur orthogonal à cette droite est le vecteur $\vec{N} = \begin{pmatrix} \frac{\partial f}{\partial x}(a,b) \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$. C'est précisément le gradient de f au point (a,b).

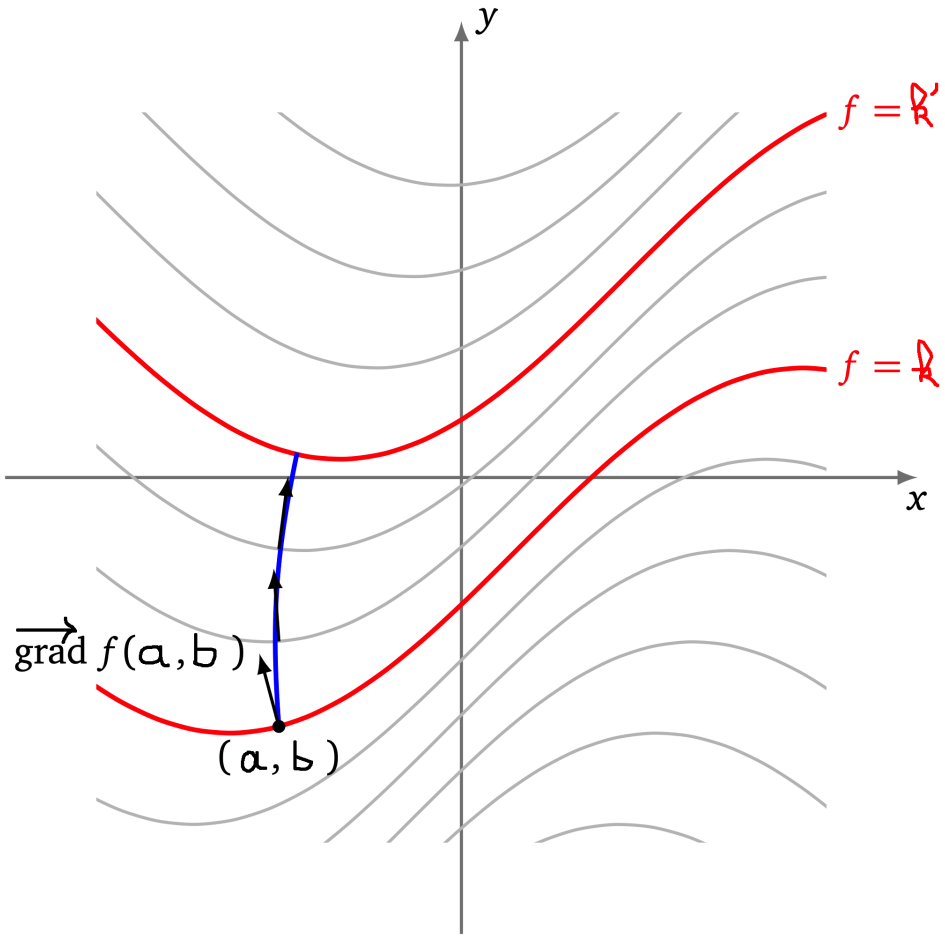
Exemple 1.92. Prenons la fonction $f(x,y) = x^2 + y^2$. Le seul point critique est l'origine $(0,0)$ et la ligne de niveau k est le cercle de rayon $\sqrt{k}$. Le gradient est champ de vecteur radial (donc orthogonal aux cercles), pointant vers l'extérieur et défini par $\nabla f(x,y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$.

Nous savons que le gradient est orthogonal aux lignes de niveaux, mais nous ne savons pas dans quelle direction il pointe. La proposition suivante nous donne une information importante.

Proposition 1.93. *Soit (a, b) un point du plan qui n'est pas un point critique de f (qu'on suppose de classe $\mathcal{C}^1$). Le vecteur gradient $\nabla f(a, b)$ pointe dans la direction de la plus forte pente vers les lignes de niveaux croissants.*



Remarque 1.94. Autrement dit, si l'on veut passer le plus vite possible de la ligne de niveau k à la ligne de niveau $k' > k$, à partir du point (a, b), de niveau $f(a, b) = k$, alors il faut démarrer en suivant la direction du gradient $\nabla f(a, b)$ et en suivant le champ de vecteurs gradients jusqu'à la ligne de niveau k'. En anglais on dit *steepest ascent*. Comme métaphore, un skieur voulant aller vite choisit la plus forte pente descendante en un point de la montagne : c'est la direction opposée au gradient de la fonction altitude, qui à un point de latitude et longitude (x, y) donne l'altitude en ce point.



Démonstration. La dérivée directionnelle de la fonction f suivant le vecteur non nul $\vec{v}$ au point (a, b) décrit la variation de f autour de ce point lorsqu'on se déplace dans la direction $\vec{v}$. Par la suite on suppose que la norme du vecteur $\vec{v}$ est 1, puisqu'on s'intéresse à la direction suivant laquelle la fonction f augmente le plus. Par définition de la différentielle et du gradient, nous avons :

$$D_{\vec{v}} f(a, b) = df_{(a,b)}(\vec{v}) = \langle \nabla f(a, b), \vec{v} \rangle$$

Or nous savons que le produit scalaire dans $\mathbb{R}^2$ satisfait :

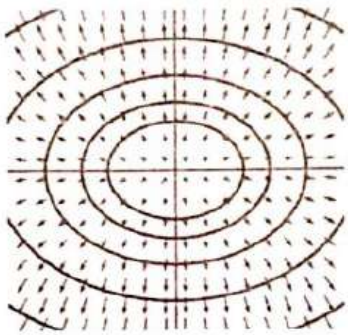
$$\langle \vec{u}, \vec{v} \rangle = \|\vec{u}\|\|\vec{v}\| \cos(\theta(\vec{u}, \vec{v}))$$

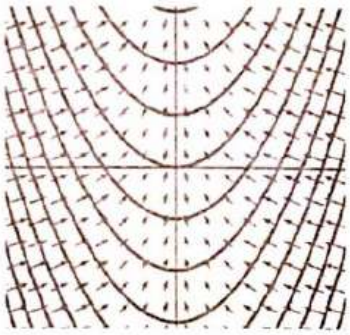
où $\theta(\vec{u}, \vec{v})$ est l'angle entre les vecteur $\vec{u}$ et $\vec{v}$. Nous avons donc, en rappelant que $\|\vec{v}\| = 1$:

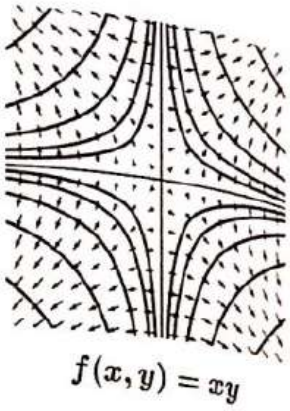
$$D_{\vec{v}} f(a, b) = \|\nabla f(a, b)\| \cos(\theta(\nabla f(a, b), \vec{v}))$$

Donc $D_{\vec{v}} f(a, b)$ est maximale lorsque l'angle θ est nul, c'est-à-dire lorsque $\vec{v}$ pointe dans la même direction que $\nabla f(a, b)$. Autrement dit la variation de f est maximale lorsqu'on se déplace dans la direction du gradient. Cela nous dit que le gradient indique la croissance maximale de la fonction f. □

Exemple 1.95. La fonction $f : (x, y) \mapsto xy$ a pour lignes de niveau $k \neq 0$ les hyperboles d'équation $y = \frac{k}{x}$. Pour $k = 0$, les lignes de niveau sont l'union de l'axe horizontal et de l'axe vertical, ce n'est pas une courbe, c'est un sous-ensemble (non régulier) de $\mathbb{R}^2$







Nous allons finir ce chapitre sur une conséquence du théorème des fonctions implicites, ou plutôt sur un théorème équivalent au théorème des fonctions implicites, et on admet la preuve :

Théorème d'inversion locale. *Soit I, J deux intervalles ouverts de $\mathbb{R}$, soit $\varphi : I \to J$ une fonction de classe $\mathcal{C}^1$, et soit $a \in I$. Si $\varphi'(a) \neq 0$, alors φ est localement inversible autour de a, c'est à dire qu'il existe un intervalle ouvert $I_a \subset I$ contenant a et un intervalle ouvert $J_b \subset J$ contenant $b = \varphi(a)$, tels que*

1. *la restriction de φ à I_a – notée ψ – est une bijection entre I_a et J_b, et*
2. *la fonction réciproque $\psi^{-1} : J_b \to I_a$ est de classe $\mathcal{C}^1$.*

Définition 1.96. *Une fonction ψ qui est bijective, de classe $\mathcal{C}^1$ et dont la réciproque est de classe $\mathcal{C}^1$ est appelée* $\mathcal{C}^1$-difféomorphisme.

1.5 Surfaces d'équation $z = f(x, y)$

Une ligne de niveau d'une fonction à deux variables est dans la plupart des cas une courbe du plan $\mathbb{R}^2$ qu'on peut localement paramétrer grâce au théorème des fonctions implicites. Inversement, il se trouve que toute courbe paramétrée du plan peut s'écrire comme la ligne de niveau 0 d'une fonction lisse, c'est à dire de classe $\mathcal{C}^\infty$. De même, dans $\mathbb{R}^3$, une surface de dimension 2 peut donc s'écrire comme l'ensemble de niveau 0 d'une fonction lisse (différentiable une inifnité

de fois) de trois variables $g(x,y,z)$. L'ensemble de niveau $k \in \mathbb{R}$ de la fonction g est défini comme la préimage du nombre k :

$$S_k = g^{-1}(k) = \Big\{(x,y,z) \in \mathbb{R}^3 \ \text{ tel que } \ g(x,y,z) = k\Big\}$$

Comme on suppose que g est continue, c'est un ensemble fermé. Plus généralement, tout sous-ensemble fermé de $\mathbb{R}^n$ peut s'écrire que l'ensemble de niveau zéro d'une fonction lisse [1]. Par le Théorème des Fonctions Implicites, toute courbe du plan peut être vue localement comme le graphe d'une fonction $y = \varphi(x)$ ou $x = \widetilde{\varphi}(y)$. Ainsi, pour connaitre le comportement local d'une courbe, il suffit de bien savoir étudier les fonctions à une variable et la géométrie de leur graphe.

Comme pour les courbes dans $\mathbb{R}^2$, il existe un théorème des fonctions implicites pour les surfaces qui nous dit que, si la condition $\frac{\partial g}{\partial z}(a,b,c) \neq 0$ est satisfaite, alors on peut trouver une unique fonction de classe $\mathcal{C}^1$ à deux variables $\psi : D \to I$ avec D un domaine ouvert de $\mathbb{R}^2$ contenant (a,b), I un intervalle ouvert contenant c, et telle que la surface S_0 dans un voisinage de (a,b,c) peut être vue comme le graphe de la fonction ψ, c'est à dire que S_0 dans ce voisinage consiste en tous les points (x,y,z) avec $z = \psi(x,y)$. Comme pour les courbes, il nous suffit donc de bien comprendre le fonctionnement des fonctions à deux variables et la géométrie de leur graphe pour tout savoir sur l'étude locale de n'importe quelle surface de dimension 2 ! Dans cette section, nous allons donc majoritairement nous limiter à l'étude des surfaces du type $z = f(x,y)$ pour une fonction $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ de classe $\mathcal{C}^1$. L'ensemble :

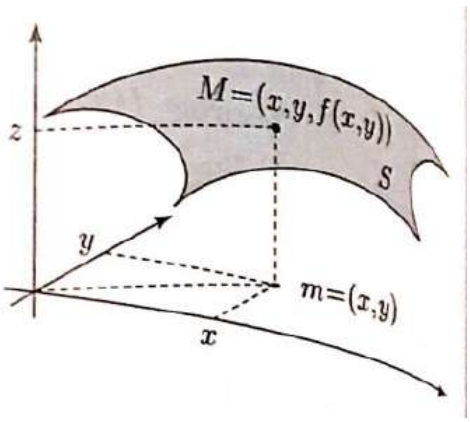
$$S = \Big\{(x,y,z) \in D \times \mathbb{R} \subset \mathbb{R}^3 \quad \text{tel que} \quad z = f(x,y)\Big\} = \mathrm{Gr}(f)$$

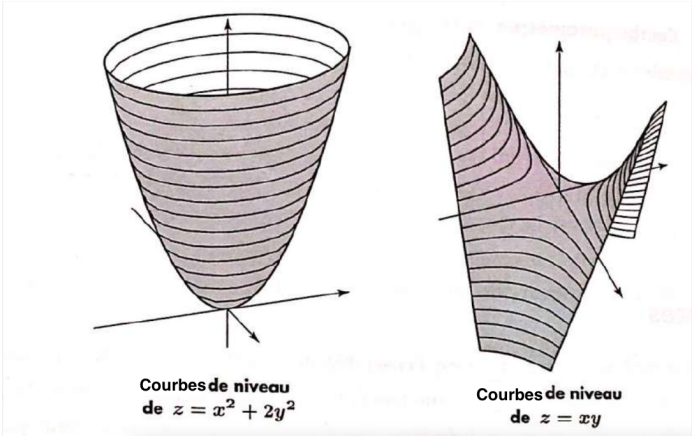
est appelée *surface d'équation* $z = f(x,y)$ ou *graphe de la fonction* f. Pour tout point $m = (x,y)$ appartenant au domaine de définition D, correspond un unique point $M = (x,y,f(x,y))$ de la surface S d'équation $z = f(x,y)$. Le point M se projette en m sur le plan xOy. La surface S est par définition le graphe de la fonction de deux variables f.

Si on relève une ligne de niveau k sur le graphe de la fonction f, on obtient ce qu'on appelle une *courbe de niveau* k. Elle est incluse dans le graphe de f. Plus précisément la courbe de niveau $k \in \mathbb{R}$ de la fonction f est le sous-ensemble de $\mathrm{Gr}(f)$ défini par :

$$C_k = \{(x,y,z = f(x,y)) \in \mathrm{Gr}(f) \ \text{ tel que } \ z = k\}$$

Les courbes de niveau sont bien entendu disjointes si $k \neq k'$. La ligne de niveau k est une courbe du plan $\mathbb{R}^2$ (dans le domaine de définition), la courbe de niveau k est une courbe de l'espace $\mathbb{R}^3$ (sur le graphe de f), située *au dessus* de la ligne de niveau k. On obtient la courbe de niveau k en partant de la ligne de niveau k et en remontant à l'altitude k. On observe donc que la surface d'équation $z = f(x,y)$ est précisément l'union des courbes de niveau, en "empilant" les courbes de niveau.





1. https://math.stackexchange.com/questions/791248/every-closed-subset-e-subseteq-mathbbrn-is-the-zero-point-set-of-a-smooth

Rappelons que dans le plan $\mathbb{R}^2$, les droites ont pour équation $Ax + By + C = 0$. C'est une équation linéaire en les deux coordonnées x et y, c'est une caractéristique des droites. Dans un espace à deux dimensions, une équation algébrique en les deux variables définit une courbe algébrique. Lorsque cette équation algébrique est linéaire en les variables, alors ces courbe est une droite. Si $C = 0$ alors l'équation devient $Ax + By = 0$ et cette droite passe par l'origine donc c'est un sous-espace vectoriel (ou droite vectorielle) de $\mathbb{R}^2$. Si $C \neq 0$, on dit que la droite est un sous-espace affine (ou droite affine) de $\mathbb{R}^2$. Le mot *affine* veut dire que ce sous-ensemble se comporte presque comme un espace vectoriel, mis à part qu'il ne contient pas l'élément nul. Dans l'espace à deux dimensions, pour définir une courbe en une dimension, il faut une équation. Si on a un système de deux équations algébriques à résoudre, par exemple le système :

$$\begin{cases} 3x + 2y - 4 = 0 \\ -2x + 2y + 3 = 0 \end{cases}$$

Ce système d'équations représente deux droites. Résoudre ce système revient à trouver l'ensemble des points de coordonnées (x, y) satisfaisant les deux équations, c'est à dire les points qui appartiennent aux deux droites. Dans le cas présent, les deux droites se coupent en un unique point $\left(\frac{7}{5}, -\frac{1}{10}\right)$. C'est la solution du système d'équations. Le système suivant :

$$\begin{cases} 3x + 2y - 4 = 0 \\ -6x - 4y + 1 = 0 \end{cases}$$

consiste en deux droites parallèles, donc l'ensemble des points qui appartiennent aux deux droites est vide. Cela représente un cas qui nous intéresse peu. Dans tous les cas, nous voyons que pour définir un point dans $\mathbb{R}^2$, on peut le définir comme l'intersection de deux droites. C'est à dire pour définir un sous-ensemble de dimension 0 dans un espace en deux dimension, il faut et il suffit de deux équations algébriques linéaires.

L'idée générale c'est qu'une équation algébrique dans un espace de dimension n définit une surface de dimension $n - 1$. Un système de deux équations algébriques (si elles sont indépendantes) définissent une surface de dimension $n - 2$ (ou un ensemble vide). Au dessus ça devient un peu plus compliqué car ça dépend des cas mais l'idée générale si tout se passe bien c'est qu'un un système de k équations algébriques indépendantes définit une surface de dimension $n - k$ (il y a des exceptions!!!!). Ainsi, dans l'espace $\mathbb{R}^3$, une équation algébrique des variables x, y, z définit une surface. Deux équations définissent une courbe (ou une surface si les deux équations sont reliées, ou bien l'ensemble vide), et trois équations indépendantes définissent un point de l'espace. En particulier, les équations algébriques linéaires de x, y, z de type $Ax+By+Cz+D = 0$ définissent un plan vectoriel ou *affine* de l'espace. Affine ici veut dire que le plan ne passe pas par l'origine. Un système de deux équations linéaires consiste géométriquement entre l'intersection de deux plans (s'ils ne sont pas parallèles) et donc l'ensemble des points solutions est une droite de l'espace.

L'intersection de deux plans affines de $\mathbb{R}^3$, lorsqu'ils ne sont pas confondus ou parallèles, forme une droite de l'espace. Une droite dans $\mathbb{R}^3$ est donc définie par deux équations de plans qui sont indépendantes algébriquement. Une équation de plan affine de $\mathbb{R}^3$ est $ax+by+cz+d = 0$. Une équation algébrique de droite est le système de deux équations :

$$\begin{cases} a_1x + b_1y + c_1z + d_1 = 0 \\ a_2x + b_2y + c_2z + d_2 = 0 \end{cases}$$

de telle façon que le triplet (a_1, b_1, c_1) n'est pas proportionnel (colinéaire) à (a_2, b_2, c_2). En résumé : une droite dans le plan $\mathbb{R}^2$ est représentée par une équation algébrique linéaire $Ax +$

$By + C = 0$. Dans l'espace $\mathbb{R}^3$, un plan est représenté par une équation algébrique linéaire $Ax+By+Cz+D = 0$ tandis qu'une droite de l'espace est définie par l'intersection de deux plans, donc par un système de DEUX équations algébriques linéaires (c'est un peu contradictoire). L'idée à retenir c'est qu'une équation algébrique en n variables définit une surface de dimension $n-1$ dans $\mathbb{R}^n$.

Dans le plan $\mathbb{R}^2$ comme dans l'espace $\mathbb{R}^3$ nous pouvons décrire les droites à l'aide d'un paramètre. Il suffit d'un point et d'un vecteur directeur. La droite passant par le point (a,b,c) et de vecteur directeur $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$ est le sous-ensemble de $\mathbb{R}^3$ donné par :

$$\left\{(x,y,z) \in \mathbb{R}^3 \text{ tel que } (x,y,z) = (a,b,c) + t\begin{pmatrix} u \\ v \\ w \end{pmatrix}, \forall\, t \in \mathbb{R}\right\}$$

C'est un exemple simple de courbe paramétrée ! !

Définition 1.97. *Un* chemin *ou* courbe paramétrée *de classe $\mathbb{C}^k$ (dans $\mathbb{R}^3$) est une fonction $\gamma : I \to \mathbb{R}^3, t \mapsto (u(t), v(t), w(t))$, où I est un intervalle réel, et tel que les fonctions coordonnées $u, v, w : I \to \mathbb{R}$ sont de classe $\mathbb{C}^k$. L'image de γ est appelée* support *de la courbe paramétrée. Soit $\gamma : I \to \mathbb{R}^3, t \mapsto (u(t), v(t), w(t))$ un chemin de classe $\mathcal{C}^1$. On définit le* vecteur vitesse *au temps $t_0 \in \mathring{I}$ par le vecteur de $\mathbb{R}^3$ suivant :*

$$\overrightarrow{\gamma'}(t_0) = \begin{pmatrix} u'(t_0) \\ v'(t_0) \\ w'(t_0) \end{pmatrix}$$

On dit que le chemin est régulier *en t_0 si $\overrightarrow{\gamma'}(t_0) \neq \vec{0}$.*

Exemple 1.98. La droite passant par le point (a,b,c) et de vecteur directeur $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$ peut être vue comme une courbe paramétrée si on écrit :

$$\begin{aligned} \gamma : \mathbb{R} &\longrightarrow \mathbb{R}^3 \\ t &\longmapsto (a,b,c) + t\begin{pmatrix} u \\ v \\ w \end{pmatrix} = (a+tu, b+tv, c+tw) \end{aligned}$$

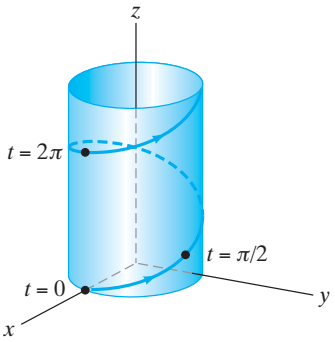
Le point (a,b,c) correspond au point $\gamma(0)$, et le vecteur directeur est le vecteur vitesse de γ.

Exemple 1.99. Soit $r > 0$, la fonction $\gamma : \mathbb{R} \to \mathbb{R}^3$ définie par :

$$\gamma(t) = (r\cos(t), r\sin(t), t)$$

est appelée *hélice circulaire*. L'hélice vit sur le cylindre d'équation $x^2+y^2 = r^2$. L'hélice admet pour vecteur vitesse le vecteur suivant :

$$\overrightarrow{\gamma'}(t) = \begin{pmatrix} -r\sin(t) \\ r\cos(t) \\ 1 \end{pmatrix}$$



La courbe est régulière en tout temps t.

La notion de courbe paramétrée est très utilisée en mécanique, où la fonction γ représente la trajectoire d'un objet physique dans l'espace et la dérivée de la fonction γ est très naturellement le vecteur vitesse. Dans la suite on suppose donc que la fonction $\gamma : I \to \mathbb{R}^3$ est de classe $\mathcal{C}^1$. Un

chemin $\gamma : I \to \mathbb{R}^3$ peut se décomposer en trois composantes $t \mapsto (u(t), v(t), w(t))$. Chacune des composantes u, v, w est une fonction $I \to \mathbb{R}$. Cela permet de connaitre le vecteur vitesse avec précision.

Soit $\gamma : I \to \mathbb{R}^3$ un chemin de classe $\mathcal{C}^1$ de fonctions coordonnées $u, v, w : I \to \mathbb{R}$, et soit $t_0 \in \mathring{I}$. Le vecteur vitesse $\overrightarrow{\gamma'}(t_0)$ (s'il n'est pas nul) est tangent au chemin γ au temps t_0. C'est donc un vecteur directeur de la droite tangente à la courbe passant par $\gamma(t_0)$. La connaissance de se vecteur permet de donner une équation paramétrique de la droite, c'est à dire de la voir comme une courbe paramétrée rectiligne. L'équation paramétrique de cette droite tangente est donnée par le chemin :

$$\begin{aligned} \sigma : \ \mathbb{R} &\longrightarrow \mathbb{R}^3 \\ s &\longmapsto \gamma(t_0) + s\overrightarrow{\gamma'}(t_0) \end{aligned}$$

Notez bien que $\gamma(t_0) = \begin{pmatrix} u(t_0) \\ v(t_0) \\ w(t_0) \end{pmatrix}$ et $\overrightarrow{\gamma'}(t_0) = \begin{pmatrix} u'(t_0) \\ v'(t_0) \\ w'(t_0) \end{pmatrix}$ sont invariants ici, seule la variable s varie. Autrement dit on peut écrire la droite tangente au support de la courbe γ au point $\gamma(t_0)$ comme le système paramétrique suivant :

$$\begin{cases} x = u(t_0) + su'(t_0) \\ y = v(t_0) + sv'(t_0) \\ z = w(t_0) + sw'(t_0) \end{cases} \qquad s \in \mathbb{R}$$

Comme nous avons supposé que le vecteur vitesse $\overrightarrow{\gamma'}(t_0)$ est non nul, une des trois composantes $u(t_0), v(t_0)$ ou $w(t_0)$ est non nulle, par exemple $u(t_0)$ (ça peut être l'une des autres). On peut alors exprimer s en fonction de x :

$$s = \frac{x - u(t_0)}{u'(t_0)}$$

et remplacer dans les deux équations du dessous. La droite tangente au support de la courbe γ au point $\gamma(t_0)$ peut alors se récrire comme l'intersection de deux plans :

$$\begin{cases} y = v(t_0) + \frac{v'(t_0)}{u'(t_0)}(x - u(t_0)) \\ z = w(t_0) + \frac{w'(t_0)}{u'(t_0)}(x - u(t_0)) \end{cases}$$

C'est donc bel et bien une droite de l'espace.

Exemple 1.100. Soit $\gamma : \mathbb{R} \to \mathbb{R}^3$ le chemin de classe $\mathcal{C}^1$ défini par :

$$\gamma(t) = (3t + 2, t^2 - 7, t - t^2)$$

On étudie le point $\gamma(1) = (5, -6, 0)$. Le vecteur vitesse de ce chemin au temps t est :

$$\overrightarrow{\gamma'}(t) = \begin{pmatrix} 3 \\ 2t \\ 1 - 2t \end{pmatrix}$$

Et donc $\overrightarrow{\gamma'}(1) = \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix}$. L'équation paramétrique de la droite tangente à la courbe paramétrée au point $\gamma(1)$ est donc :

$$\sigma(s) = (5, -6, 0) + s \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} \quad \text{c'est à dire les points } (x, y, z) \text{ de coordonnées } \quad \begin{cases} x = 5 + 3s \\ y = -6 + 2s \\ z = -s \end{cases}$$

pour tout $s \in \mathbb{R}$. En observant que $s = -z$, on peut remplacer s dans les deux premières équations, et on obtient les équations de plan suivantes :

$$\begin{cases} x = 5 - 3z \\ y = -6 - 2z \end{cases}$$

C'est donc bien une droite de l'espace.

Remarque 1.101. La recette pour passer d'une équation paramétrique d'une droite à deux équations algébriques de deux plans qui définissent la même droite se fait comme suit : 1. on a un système de trois équations qui dépendent d'un paramètre s : $\begin{cases} x(s) = 0 \\ y(s) = 0 \\ z(s) = 0 \end{cases}$; 2. on choisit une équation pour exprimer s en fonction d'une des trois variables x, y ou z ; 3. on remplace s par cette variable dans les deux autres équations. Cela donne deux équations algébriques de plan, donc une droite de l'espace.

Soit maintenant une fonction à deux variables $f : D \to \mathbb{R}$ de classe au moins $\mathcal{C}^1$ et prenons un point $(a, b) \in \mathring{D}$. On pose $c = f(a, b)$ et on regarde la surface S d'équation $z = f(x, y)$. Soit $t \mapsto u(t)$ une fonction de classe $\mathcal{C}^1$ telle que $u(0) = a$ et soit $t \mapsto v(t)$ une autre fonction de classe $\mathcal{C}^1$ telle que $v(0) = b$. Ces fonctions définissent une courbe paramétrée du plan de classe $\mathcal{C}^1$

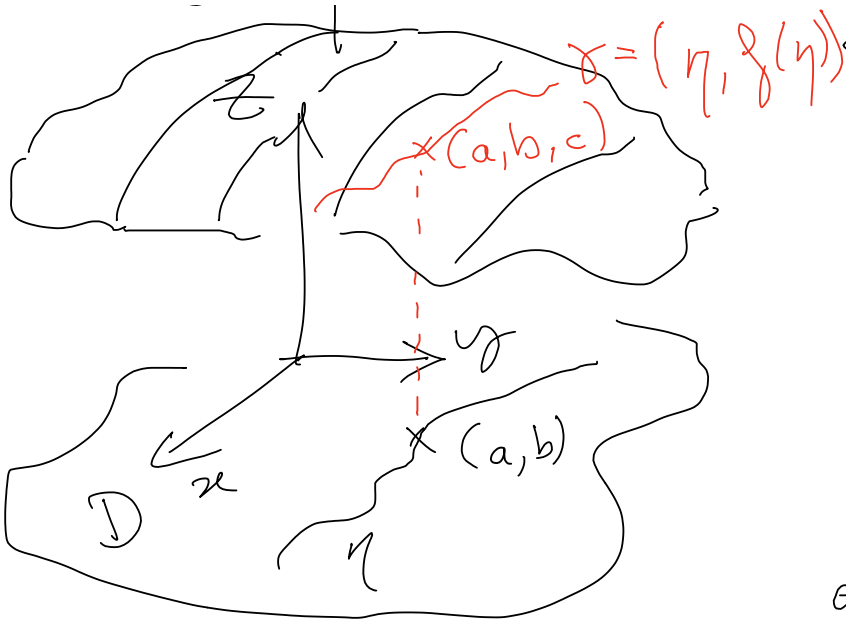
$$\begin{array}{rccc} \eta : & \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ & t & \longmapsto & (u(t), v(t)) \end{array}$$

qui passe par le point (a, b) au temps $t = 0$. On suppose qu'elle est régulière en $t = 0$, c'est à dire que soit $u'(0) \neq 0$ soit $v'(0) \neq 0$, ce qui donne $\overrightarrow{\eta'}(0) \neq \vec{0}$. On définit la fonction $t \mapsto w(t) = f(u(t), v(t))$ qui est de classe $\mathcal{C}^1$ par composition, et qui passe par $c = f(a, b)$ au temps $t = 0$. On définit le chemin suivant de classe $\mathcal{C}^1$

$$\begin{array}{rccc} \gamma : & \mathbb{R} & \longrightarrow & \mathbb{R}^3 \\ & t & \longmapsto & (u(t), v(t), w(t)) \end{array}$$

Il est tel que $\gamma(0) = (a, b, c)$.

Remarque 1.102. On voit que si la courbe paramétrée η suit la ligne de niveau passant par (a, b) alors $w'(t) = 0$ pour tout t (car le gradient est orthogonal aux lignes de niveau). Cela veut dire que dans ce cas γ reste à la même altitude, car on suit la *courbe de niveau* passant par (a, b, c).



Cette courbe paramétrée reste dans la surface d'équation $z = f(x, y)$ pour tout temps t, et passe par le point (a, b, c) au temps $t = 0$. Par la règle des chaines la dérivée de la fonction w vaut :

$$w'(t) = u'(t)\frac{\partial f}{\partial x}(u(t), v(t)) + v'(t)\frac{\partial f}{\partial y}(u(t), v(t))$$

Autrement dit, en notant $\overrightarrow{\eta'}(t) = \begin{pmatrix} u'(t) \\ v'(t) \end{pmatrix}$ le vecteur tangent à la courbe η au point $\eta(t) = (u(t), v(t))$, nous avons :

$$w'(t) = df_{\eta(t)}\left(\overrightarrow{\eta'}(t)\right) = \left\langle \nabla f(\eta(t)), \overrightarrow{\eta'}(t) \right\rangle$$

En particulier, nous avons au temps $t = 0$ que $w'(0) = \left\langle \nabla f(a, b), \overrightarrow{\eta'}(0) \right\rangle$. Donc au temps $t = 0$ nous avons comme vecteur tangent à la courbe γ le vecteur suivant (non-nul car $\overrightarrow{\eta'}(0)$) :

$$\overrightarrow{\gamma'}(0) = \begin{pmatrix} \overrightarrow{\eta'}(0) = \begin{pmatrix} u'(0) \\ v'(0) \end{pmatrix} \\ \left\langle \nabla f(a, b), \overrightarrow{\eta'}(0) \right\rangle \end{pmatrix} \tag{1.8}$$

On voit donc que la vitesse de la courbe paramétrée γ au temps $t = 0$ a une composante horizontale – la vitesse du chemin η – et une composante verticale $df_{(a,b)}\left(\overrightarrow{\eta'}(0)\right) = \left\langle \nabla f(a, b), \overrightarrow{\eta'}(0) \right\rangle$ qui mesure l'accroissement de la fonction f dans la direction du vecteur $\overrightarrow{\eta'}(0)$. Cela nous donne l'équation paramétrique de la droite tangente à la courbe γ au temps $t = 0$, c'est à dire en (a, b, c) :

$$\begin{aligned} \sigma : \ \mathbb{R} &\longrightarrow \mathbb{R}^3 \\ s &\longmapsto (a, b, c) + s \begin{pmatrix} \overrightarrow{\eta'}(0) = \begin{pmatrix} u'(0) \\ v'(0) \end{pmatrix} \\ \left\langle \nabla f(a, b), \overrightarrow{\eta'}(0) \right\rangle \end{pmatrix} \end{aligned}$$

La courbe γ étant incluse dans le graphe de f, nous en déduisons que la droite tangente est tangente au graphe de f au point (a, b, c). Un autre choix de chemin η tel que $\eta(0) = (a, b)$ et

régulier en $t = 0$ donne un autre chemin γ tel que $\gamma(0) = (a, b, c)$ et régulier en $t = 0$, et nous aurait donc donné une autre droite tangente au graphe de f au point (a, b, c). Pour chaque choix de chemin η, on obtient ainsi une droite tangente au graphe de f au point (a, b, c).

Remarque 1.103. L'union des droites tangentes ainsi définies forme un sous-ensemble de $\mathbb{R}^3$, mais nous devons approfondir l'analyse pour savoir quelle type de sous-ensemble (c'est un plan qui passe par (a, b, c) et tangent à la surface d'équation $z = f(x, y)$).

Maintenant, en notant $\overline{\langle -, - \rangle}$ le produit scalaire canonique sur $\mathbb{R}^3$ pour le distinguer de celui sur $\mathbb{R}^2$ on a :

$$\overline{\left\langle \begin{pmatrix} u \\ v \\ w \end{pmatrix}, \begin{pmatrix} h \\ k \\ l \end{pmatrix} \right\rangle} = uh + vk + wl$$

Alors, quel que soit le choix du chemin η, le vecteur tangent $\overrightarrow{\gamma'}(0)$ défini dans l'Equation (1.8) est orthogonal dans $\mathbb{R}^3$ au vecteur :

$$\vec{N} = \begin{pmatrix} -\nabla f(a,b) = \begin{pmatrix} -\frac{\partial f}{\partial x}(a,b) \\ -\frac{\partial f}{\partial y}(a,b) \end{pmatrix} \\ 1 \end{pmatrix}$$

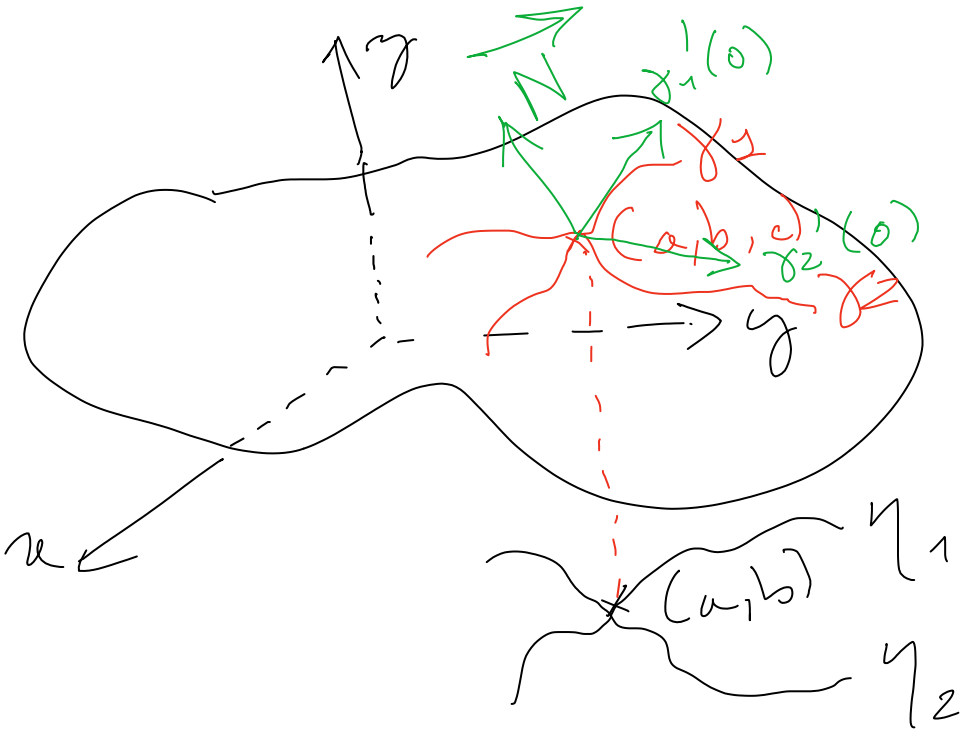
En effet on a :

$$\overline{\left\langle \overrightarrow{\gamma'}(0), \vec{N} \right\rangle} = \underbrace{-u'(0) \times \frac{\partial f}{\partial x}(a,b) - v'(0) \times \frac{\partial f}{\partial x}(a,b)}_{= -\left\langle \overrightarrow{\eta'}(0), \nabla f(a,b) \right\rangle} + \underbrace{w'(0) \times 1}_{= \left\langle \nabla f(a,b), \overrightarrow{\eta'}(0) \right\rangle} = 0$$

Nous observons que :

1. $\vec{N}$ ne dépend que du point (a, b), et pas du choix de chemin η qui passe en (a, b) au temps $t = 0$ (c'est à dire du vecteur vitesse $\overrightarrow{\eta'}(0)$), et que
2. $\vec{N}$ est orthogonal au chemin γ au temps $t = 0$ (c'est à dire au point $\gamma(0) = (a, b, c)$), et ce quel que soit le choix du chemin η. Autrement dit $\vec{N}$ est orthogonal à toutes les droites tangentes au graphe de f en (a, b, c).

Nous allons voir que cette dernière propriété du vecteur $\vec{N}$ définit précisément le plan tangent au graphe de f au point (a, b, c). En effet nous savons qu'un vecteur de $\mathbb{R}^3$ définit de façon unique un plan vectoriel de $\mathbb{R}^3$. Cela passe par un résultat qui généralise directement la proposition 1.89.



Proposition 1.104. *Les composantes d'un vecteur normal à un plan de l'espace d'équation* $Ax + By + Cz + D = 0$ *sont*

$$\vec{N} = \begin{pmatrix} A \\ B \\ C \end{pmatrix}$$

.

Démonstration. Soit $M_1 = (x_1, y_1, z_1)$ et $M_2 = (x_2, y_2, z_2)$ deux points du plan d'équation $Ax + By + Cz + D = 0$. Alors le vecteur $\overrightarrow{M_1M_2} = \begin{pmatrix} x_2 - x_1 \\ y_2 - y_1 \\ z_2 - z_1 \end{pmatrix}$ satisfait comme équation :

$$A(x_2 - x_1) + B(y_2 - y_1) + C(z_2 - z_1) = 0$$

ce qui correspond à avoir $\left\langle \vec{N}, \overrightarrow{M_1M_2} \right\rangle = 0$. Le vecteur $\overrightarrow{M_1M_2}$ est donc bien orthogonal au vecteur $\vec{N}$. □

Remarque 1.105. Tous les autres vecteurs normaux au plan sont colinéaires à celui-ci.

Prenons le vecteur $\vec{N} = \begin{pmatrix} -\frac{\partial f}{\partial x}(a,b) \\ -\frac{\partial f}{\partial y}(a,b) \\ 1 \end{pmatrix}$ qu'on a déterminé plus haut. Un plan orthogonal au vecteur $\vec{N}$ a pour équation :

$$-\frac{\partial f}{\partial x}(a,b)x - \frac{\partial f}{\partial y}(a,b)y + z + D = 0$$

avec un coefficient D quelconque. Pour $x = a, y = b, z = c$ on peut fixer D de façon unique, pour obtenir l'équation du plan normal au vecteur $\vec{N}$ passant par le point (a, b, c) :

$$D = \frac{\partial f}{\partial x}(a,b)a + \frac{\partial f}{\partial y}(a,b)b - c$$

D'autre part rappelons que, quel que soit le choix du chemin η passant par le point (a,b) au temps $t=0$, la droite tangente au chemin γ au point (a,b,c) est orthogonale au vecteur $\vec{N}$. Cela veut dire que donc elle est une droite dans le plan défini avec le coefficient ci-dessus.

Définition 1.106. *Soit $f : D \to \mathbb{R}$ une fonction à deux variables de classe au moins $\mathcal{C}^1$, et $(a,b) \in \mathring{D}$. On pose $c = f(a,b)$. L'ensemble formé par l'union de toutes les droites tangentes au graphe de f au point (a,b,c) forme un plan affine de l'espace contenant le point (a,b,c) appelé* plan tangent à la surface d'équation $z = f(x,y)$ au point (a,b,c) *d'équation :*

$$(x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b) = z - c$$

Remarque 1.107. Le slogan qu'il faut retenir, c'est que le plan tangent au graphe de f au point (a,b,c) est l'endroit où vivent les vecteurs tangents à la surface d'équation $z = f(x,y)$ au point (a,b,c).

Exemple 1.108. Soit la fonction $f(x,y) = x^2 + 2y^2$. L'équation du plan tangent à la surface d'équation $z = f(x,y)$ en un point (a,b) est donc :

$$2a(x-a) + 4b(y-b) = z - (a^2 + 2b^2)$$

C'est à dire, de façon plus claire :

$$2ax + 4by - z = a^2 + 2b^2$$

En particulier, au point $(a,b) = (0,0)$ l'équation du plan tangent à la surface est $z = 0$, c'est à dire le plan horizontal ! D'autre part on a que $f(x,y) \geq 0$ pour tout (x,y), donc le graphe de f est localisé au dessus du plan horizontal $z = 0$.

Nous pouvons voir la définition du plan tangent d'une autre façon. Soit $(a,b) \in \mathring{D}$, et $\epsilon > 0$, et soit les deux chemins de $\mathbb{R}^2$ définis par :

$$\begin{array}{ccc} \eta_x :]-\epsilon,\epsilon[& \longrightarrow & D \\ t & \longmapsto & (a+t,b) \end{array} \qquad \begin{array}{ccc} \eta_y :]-\epsilon,\epsilon[& \longrightarrow & D \\ t & \longmapsto & (a,b+t) \end{array}$$

Ces deux courbes paramétrées définissent deux chemins dans le graphe de f – respectivement $\gamma_x : t \mapsto (\eta_x(t), f(\eta_x(t)))$ et $\gamma_y : t \mapsto (\eta_y(t), f(\eta_y(t)))$ – dont on peut calculer les vecteurs tangents au temps $t = 0$, qui sont donnés par l'Equation (1.8). Comme $\overrightarrow{\eta_x'}(0) = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ et $\overrightarrow{\eta_y'}(0) = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, nous avons

$$\overrightarrow{\gamma_x'}(0) = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(a,b) \end{pmatrix} \qquad \text{et} \qquad \overrightarrow{\gamma_y'}(0) = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$$

Ce sont précisément les deux vecteurs bleus que l'on voit dans la Figure 1.1 ! Le premier vecteur est le vecteur directeur de la tangente au graphe de f en $(a,b,f(a,b))$ dans la direction x, tandis que le deuxième vecteur est le vecteur directeur de la tangente au graphe de f en $(a,b,f(a,b))$ dans la direction y. Par convention, on les note $\vec{T}_x$ et $\vec{T}_y$. Par construction ils sont normaux au vecteur $\vec{N}$ et nous avons les faits géométriques suivants :

Proposition 1.109. *Soit $f : D \to \mathbb{R}$ une fonction à deux variables de classe au moins $\mathcal{C}^1$, et $(a,b) \in \mathring{D}$. On pose $c = f(a,b)$. Le plan tangent au graphe de f au point (a,b,c) est le plan de l'espace engendré par les deux vecteurs :*

$$\vec{T}_x = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial x}(a,b) \end{pmatrix} \qquad \textit{et} \qquad \vec{T}_y = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial y}(a,b) \end{pmatrix}$$

Démonstration. Tout vecteur tangent $\overrightarrow{\gamma'}(0)$ est donné par l'Equation (1.8), et donc comme $\overrightarrow{\eta_x'}(0) = \binom{1}{0}$ et $\overrightarrow{\eta_y'}(0) = \binom{0}{1}$, on obtient que $\overrightarrow{\gamma'}(0) = u'(0)\vec{T}_x + v'(0)\vec{T}_y$. Autrement dit tout vecteur tangent au graphe de f au point (a, b, c) se décompose sur $\vec{T}_x$ et $\vec{T}_y$, qui forme donc une base du plan tangent. □

Définition 1.110. *Le* produit vectoriel *entre deux vecteurs tridimensionnels* $\begin{pmatrix} u \\ v \\ w \end{pmatrix}$ *et* $\begin{pmatrix} h \\ k \\ l \end{pmatrix}$ *est le vecteur de* $\mathbb{R}^3$ *défini par :*

$$\begin{pmatrix} u \\ v \\ w \end{pmatrix} \wedge \begin{pmatrix} h \\ k \\ l \end{pmatrix} = \begin{pmatrix} vl - wk \\ -ul + wh \\ uk - vh \end{pmatrix}$$

Proposition 1.111. *Le vecteur normal canonique* $\vec{N} = \begin{pmatrix} -\nabla f(a,b) \\ 1 \end{pmatrix}$ *est le produit vectoriel de* $\vec{T}_x$ *et* $\vec{T}_y$*, c'est à dire :* $\vec{N} = \vec{T}_x \wedge \vec{T}_y$.

Démonstration. Immédiat. □

Nous nous intéressons maintenant à étudier les maxima et les minima d'une fonction à deux variables. Autrement dit, étudions le cas particulier où (a, b) est un point critique de la fonction f c'est à dire un extremum local ou un point selle, ce qu'on peut traduire par un point du graphe de f où $\nabla f(a, b) = \binom{0}{0}$. Dans ce cas, le vecteur $\vec{N}$ est purement vertical et vaut $\vec{N} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, il pointe vers le bas. Le plan tangent est donc un plan affine horizontal d'équation $z = f(a, b)$.

Si le graphe de f est au dessus (resp. en dessous) du plan dans un voisinage de (a, b), c'est à dire si $f(x, y) \geq f(a, b)$ (resp. $f(x, y) \leq f(a, b)$) localement autour de (a, b), alors on a un mimimum (resp. maximum) local. Il peut arriver que ce soit ni l'un ni l'autre, dans ce cas on parle de *point selle* ou *point col*. Une façon de savoir rapidement si on a un extremum local ou un point selle, c'est de calculer la matrice Hessienne de f au point (a, b).

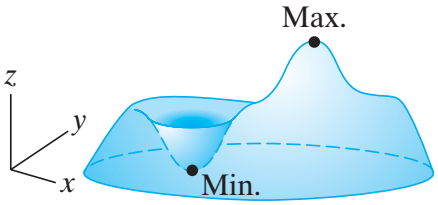


Figure 4.12 The graph of $z = f(x, y)$.

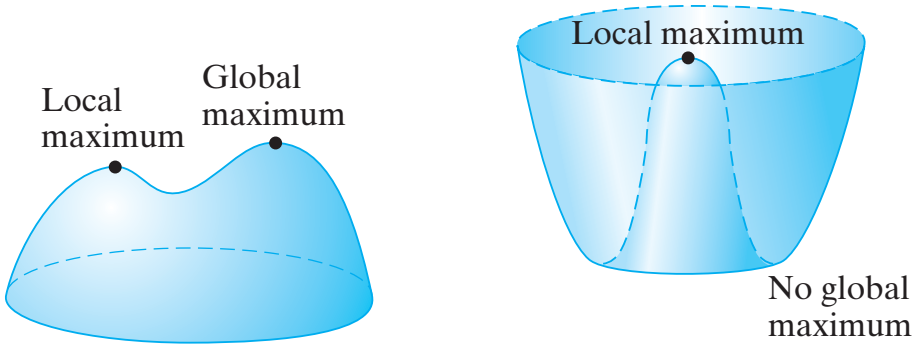


Figure 4.13 Examples of local and global maxima.

Soit $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ une fonction entre deux espaces vectoriels de dimension finie (on suppose l'ensemble de définition D ouvert), qu'on suppose de classe au moins $\mathcal{C}^1$. On pose $e_1, \ldots, e_n$ la base de $\mathbb{R}^n$ qu'on a choisie, et si x est un point du domaine D, on note ses coordonnées $(x^1, \ldots, x^n)$ dans cette base. On peut dériver de manière récursive la fonction f en dérivant les fonctions partielles (voir Definition 1.14) comme dans la Définition 1.21 :

1. à l'ordre 1, on a la définition usuelle : la j-ème dérivée partielle de f est donnée par la dérivée de la j-ème fonction partielle :

$$\forall\, x \in D \qquad \frac{\partial f}{\partial x^j}(x) = \lim_{t\to 0} \frac{f(x^1,\dots,x^{j-1},x^j+t,x^{j+1},\dots,x^n) - f(x)}{t}$$

   Pour tout j, cela définit une fonction $\frac{\partial f}{\partial x^j} : D \to \mathbb{R}^m, x \mapsto \frac{\partial f}{\partial x^j}(x)$, qu'on peut donc dériver si cela est possible.
2. à l'ordre 2, on peut dériver dans la direction e_i la j-ième fonction dérivée partielle $\frac{\partial f}{\partial x^j}$:

$$\forall\, x \in D \qquad \frac{\partial \frac{\partial f}{\partial x^j}}{\partial x^i}(x) = \lim_{t\to 0} \frac{\frac{\partial f}{\partial x^j}(x^1,\dots,x^{i-1},x^i+t,x^{i+1},\dots,x^n) - \frac{\partial f}{\partial x^j}(x)}{t}$$

   Cela définit une fonction qu'on note $\frac{\partial^2 f}{\partial x^i \partial x^j} : D \mapsto \mathbb{R}^m$. Attention, pour l'instant l'ordre de la dérivée partielle n'est pas interchangeable ! On dérive d'abord par rapport à la dérivée la plus à droite, plus celle à gauche. En particulier on ne sait pas encore que $\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$ (ce qu'on verra dans le Théorème de Schwarz).
3. On suppose la fonction f dérivée jusqu'à l'ordre p, selon un ordre donné $\frac{\partial^p f}{\partial x^{i_p}\dots\partial x^{i_1}}$ (c'est à dire d'abord selon e_{i_1}, puis e_{i_2}, etc. puis en dernier e_{i_p}), et on dérive à l'ordre $p+1$ selon une direction choisie $e_{i_{p+1}}$ en utilisant la formule ci-dessus. Cela définit une fonction dérivée partielle à l'ordre $p+1$, notée $\frac{\partial^{p+1} f}{\partial x^{i_{p+1}}\partial x^{i_p}\dots\partial x^{i_1}}$.

Exemple 1.112. On prend la fonction de l'Exemple 1.3. Les deux dérivées partielles d'ordre 1 sont bien définies, et il y a a priori 4 dérivées partielles secondes (d'ordre 2) :

$$\frac{\partial^2 f}{\partial x \partial x}, \quad \frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x}, \text{ et } \frac{\partial^2 f}{\partial y \partial y}.$$

Elles existent pour tout $(x,y) \neq (0,0)$, par contre au point $(0,0)$, la dérivée partielle seconde $\frac{\partial^2 f}{\partial y \partial x}(0,0)$ n'est pas définie car la fonction dérivée partielle $\frac{\partial f}{\partial x}$ n'est pas dérivable dans la direction y en $(0,0)$. Par contre elle est dérivable dans la direction x donc $\frac{\partial^2 f}{\partial x \partial x}(0,0)$ existe (et vaut 0).

Remarque 1.113. En général, lorsqu'on répète deux fois la même variable dans une dérivée d'ordre $p \geq 2$, on fusionne ensemble les variables identiques et on met leur nombre en exposant. Par exemple $\frac{\partial^2 f}{\partial x \partial x}$ s'écrit $\frac{\partial^2 f}{\partial x^2}$ et $\frac{\partial^2 f}{\partial y \partial y}$ s'écrit $\frac{\partial^2 f}{\partial y^2}$.

Définition 1.114. *La fonction $f : D \subset \mathbb{R}^n \to \mathbb{R}^m$ est dite* de classe $\mathcal{C}^p$ *si et seulement si ses dérivées partielles d'ordre p existent et sont continues. Elle est dite de classe $\mathcal{C}^0$ si elle est (au moins) continue, et elle est dite de classe $\mathcal{C}^\infty$ si elle est de classe $\mathcal{C}^p$ pour tout $p \in \mathbb{N}$.*

Remarque 1.115. Soulignons ici que le caractère $\mathcal{C}^p$ ne dépend pas de la base par rapport à laquelle les dérivées partielles sont définies. Cela se prouve très facilement par récurrence, car la matrice de changement de base est insensible à la dérivation.

Theorème de Schwarz. *Soit $f : D \subset \mathbb{R}^n \to \mathbb{R}$ une fonction de classe $\mathcal{C}^2$. Alors ses dérivées partielles à l'ordre 2 sont interchangeables :*

$$\frac{\partial^2 f}{\partial x^i \partial x^j} = \frac{\partial^2 f}{\partial x^j \partial x^i}$$

Démonstration. La preuve est intéressante et se trouve pages 176-178 du Liret-Martinais. □

Remarque 1.116. On ne saurait sous-estimer l'importance du Théorème de Schwarz. Il permet de permuter l'ordre des dérivées sans s'inquiéter car le résultat reste le même. On peut le vérifier sur des fonctions à deux variables, comme $f(x,y) = \sin(x\ln(y))$.

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables et soit $(a,b) \in \mathring{D}$. On suppose que f est différentiable dans un voisinage ouvert $U \subset D$ contenant (a,b). La différentielle de f sur ce voisinage satisfait la forme matricielle suivante :

$$\begin{aligned} df : \ U &\longrightarrow (\mathbb{R}^2)^* \\ (x,y) &\longmapsto df_{(x,y)} = \begin{pmatrix} \frac{\partial f}{\partial x}(x,y) & \frac{\partial f}{\partial y}(x,y) \end{pmatrix} \end{aligned}$$

La différentielle est à valeurs dans $(\mathbb{R}^2)^*$, et nous rappelons que le gradient $\nabla f = (df)^t$ est la transposée de la différentielle. Le gradient est un champ de vecteurs $\nabla f : U \to \mathbb{R}^2$. On suppose en outre que la fonction différentielle $df : U \mapsto (\mathbb{R}^2)^*$ est différentiable au point (a,b) (cela équivaut à supposer que le champ de vecteurs gradient est différentiable en (a,b)). Les composantes du champ de vecteurs gradient $\nabla f = \begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}$ sont donc dérivables par rapport à x et à y, au point (a,b). La matrice Jacobienne du champ de vecteurs gradient $\nabla f = (df)^t : U \to \mathbb{R}^2$ au point (a,b) est alors donnée par :

$$J_{(a,b)}(\nabla f) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x}(a,b) & \frac{\partial^2 f}{\partial y \partial x}(a,b) \\ \frac{\partial^2 f}{\partial x \partial y}(a,b) & \frac{\partial^2 f}{\partial y \partial y}(a,b) \end{pmatrix}$$

En résumé on dit qu'une fonction $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ est *deux fois différentiable* en un point $(a,b) \in \mathring{D}$ si :

1. la fonction $f : D \to \mathbb{R}$ est différentiable dans un voisinage ouvert $U \subset D$ contenant (a,b),
2. le champ de vecteurs gradient $\nabla f : U \mapsto \mathbb{R}^2$ est différentiable au point (a,b).

et dans ce cas on peut définir la différentielle (matrice Jacobienne) du gradient de f au point (a,b). Notons qu'il est *important* que f est non seulement différentiable en (a,b), mais aussi dans une voisinage ouvert de (a,b).

Définition 1.117. *On appelle* matrice Hessienne *de la fonction f au point (a,b) la transposée de la matrice Jacobienne du champ de vecteurs gradient ∇f au point (a,b) :*

$$H_{(a,b)}(f) = \Big(J_{(a,b)}(\nabla f)\Big)^t = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x}(a,b) & \frac{\partial^2 f}{\partial x \partial y}(a,b) \\ \frac{\partial^2 f}{\partial y \partial x}(a,b) & \frac{\partial^2 f}{\partial y \partial y}(a,b) \end{pmatrix}$$

Remarque 1.118. Bien entendu, si $f : D \to \mathbb{R}$ est de classe $\mathcal{C}^2$, alors on peut définir une fonction continue à valeurs dans les matrices 2×2 par $H(f) : D \to \mathcal{M}_{2\times 2}(\mathbb{R}), (x,y) \mapsto H_{(x,y)}(f)$.

Tout comme la matrice Jacobienne de f en (a,b) représente la différentielle df en (a,b), la matrice Hessienne représente la différentielle seconde de f en (a,b). En effet, la matrice Jacobienne du champ de vecteurs gradient ∇f en (a,b) est la matrice associée à la différentielle de $\nabla f = (df)^t$ en (a,b). La différentielle de ∇f au point (a,b) (voir la définition 1.28) est alors une application linéaire de $\mathbb{R}^2$ dans $\mathbb{R}^2$ qu'on définit comme suit :

$$\begin{aligned} d(\nabla f)_{(a,b)} : \quad \mathbb{R}^2 &\longrightarrow \mathbb{R}^2 \\ \vec{v} = \begin{pmatrix} h \\ k \end{pmatrix} &\longmapsto d(\nabla f)_{(a,b)}(\vec{v}) : (\mathbb{R}^2)^* \longrightarrow \mathbb{R}, \\ &\qquad \alpha = \begin{pmatrix} s & t \end{pmatrix} \mapsto hs\frac{\partial^2 f}{\partial x^2}(a,b) + ht\frac{\partial^2 f}{\partial x \partial y}(a,b) \\ &\qquad\qquad + ks\frac{\partial^2 f}{\partial y \partial x}(a,b) + kt\frac{\partial^2 f}{\partial y^2}(a,b) \end{aligned}$$

La formule à droite peut être obtenue par le produit matriciel suivant :

$$d(\nabla f)_{(a,b)}(\vec{v})(\alpha) = \alpha \times H_{(a,b)}(f) \times \vec{v}$$

Dans ce cas, si f est différentiable deux fois en (a,b), nous définissons sa différentielle seconde en (a,b) par la formule suivante :

$$\begin{aligned} d^2 f_{(a,b)} : \ \mathbb{R}^2 &\longrightarrow \left(\mathbb{R}^2\right)^* \\ \vec{v} &\longmapsto d^2 f_{(a,b)}(\vec{v}) : \mathbb{R}^2 \longrightarrow \mathbb{R}, \\ &\qquad \vec{w} \longmapsto (\vec{v})^t \times H_{(a,b)}(f) \times \vec{w} \end{aligned}$$

En particulier, si $\vec{v} = \binom{h}{k}, \vec{w} = \binom{s}{t}$ sont deux vecteurs de $\mathbb{R}^2$, alors on a :

$$d^2 f_{(a,b)}(\vec{v},\vec{w}) = hs\frac{\partial^2 f}{\partial x^2}(a,b) + ht\frac{\partial^2 f}{\partial x \partial y}(a,b) + ks\frac{\partial^2 f}{\partial y \partial x}(a,b) + kt\frac{\partial^2 f}{\partial y^2}(a,b) \tag{1.9}$$

En particulier, nous voyons que si $\vec{v} = e_i$ et $\vec{w} = e_j$ avec $1 \leq i,j \leq 2$, où (e_1,e_2) est la base canonique de $\mathbb{R}^2$ c'est à dire $e_1 = \vec{i}$ et $e_2 = \vec{j}$, nous avons :

$$d^2 f_{(a,b)}(e_i,e_j) = \frac{\partial^2 f}{\partial x^i \partial x^j}(a,b)$$

où $x^1 = x$ et $x^2 = y$. L'ordre i,j est préservé dans le membre de gauche et dans le membre de droite ! D'autre part, par le Theorème de Schwarz, la formule (1.9) est symétrique en $\vec{v}$ et $\vec{w}$, c'est à dire que l'ordre dans lequel on dérive deux fois la fonction f n'a pas d'importance :

Corollaire 1.119. *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction deux fois différentiable au point $(a,b) \in \mathring{D}$. Alors nous avons les résultats suivants :*

1. *sa matrice Hessienne $H_{(a,b)}(f)$ est symmétrique ;*
2. *sa différentielle seconde en (a,b) est une application bilinéaire* symmétrique, *c'est à dire :*

$$d^2 f_{(a,b)}(\vec{v},\vec{w}) = d^2 f_{(a,b)}(\vec{w},\vec{v})$$

Remarque 1.120. Ce théorème apparaît pour la première fois dans un cours de calcul différentiel donné par Weierstrass en 1861 auquel assistait alors Hermann Schwarz à Berlin. Il est très important car il nous dit qu'on peut permuter l'ordre de dérivation partiel, donc qu'il n'a pas d'importance.

La matrice Hessienne nous permet de savoir si un point critique de la matrice Jacobienne est un maximum ou un minimum, ou autre. Elle jour le rôle de la dérivée seconde pour les fonctions à plusieurs variables.

Proposition 1.121. *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables deux fois différentiable en $(a,b) \in \mathring{D}$. Supposons en outre que $(a,b) \in \mathring{D}$ est un point critique de f. Alors :*

— *si* $\det(H_{(a,b)}(f)) > 0$, *nous avons deux cas :*
  1. *si* $\mathrm{tr}(H_{(a,b)}(f)) > 0$, *le point (a,b) est un* minimum local *de f (le graphe de la fonction est localement au dessus du plan tangent en (a,b)) ;*
  2. *si* $\mathrm{tr}(H_{(a,b)}(f)) < 0$, *le point (a,b) est un* maximum local *de f (le graphe de la fonction est localement au dessous du plan tangent en (a,b)).*

— *si* $\det(H_{(a,b)}(f)) < 0$, *le point (a,b) est un* point selle *(le graphe de la fonction traverse le plan tangent en (a,b)).*

— *si* $\det(H_{(a,b)}(f)) = 0$, *on ne peut pas conclure.*

Remarque 1.122. On rappelle que le déterminant d'une matrice $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ vaut $\det(A) = ad - bc$ et sa trace est $\mathrm{tr}(A) = a + d$.

Exemple 1.123. Prenons la fonction définie par $f(x, y) = x^2 + 2y^2$ (paraboloïde elliptique) vue précédemment. La matrice Hessienne en un point quelconque (a, b) est :

$$H_{(a,b)}(f) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$$

La matrice est constante de détemirnant et de trace positifs. Le seul point critique de la fonction f est l'origine $(0, 0)$. C'est donc un minimum (global). On retrouve le fait que le graphe de la fonction est au dessus du plan tangent en $(0, 0)$.

Exemple 1.124. Exemple de mon professeur de 2007. Soit $\alpha > 0$, on définit la fonction suivante :

$$\begin{aligned} f : \ (\mathbb{R}_+^*)^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto x + y + \frac{\alpha}{xy} \end{aligned}$$

Il y a un unique point critique défini par l'équation $\nabla f(a, b) = \vec{0}$, c'est $(a, b) = (\sqrt[3]{\alpha}, \sqrt[3]{\alpha})$. La Hessienne en ce point est donnée par

$$H_{(\sqrt[3]{\alpha}, \sqrt[3]{\alpha})}(f) = \frac{1}{\sqrt[3]{\alpha}} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

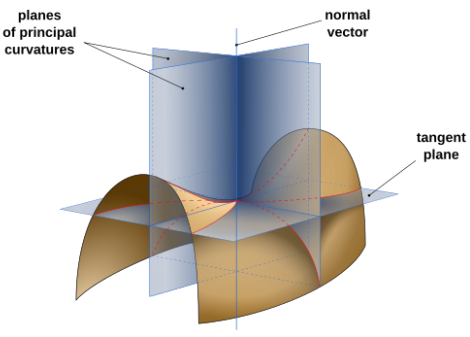
Le déterminant et la trace sont positifs donc le point $(\sqrt[3]{\alpha}, \sqrt[3]{\alpha})$ est un minimum.

Rappelons que la matrice Hessienne est la généralisation à plusieurs variables de la dérivée seconde. On le voit dans la Proposition 1.121. En effet, si on a une fonction à une variable $g : \mathbb{R} \to \mathbb{R}$, la matrice Hessienne de g au point $a \in \mathbb{R}$ est juste la dérivée seconde $g''(a)$. Du signe de $g''(a)$ on déduit si la fonction est convexe, concave ou si c'est un point d'inflexion en ce point. La dérivée première $g'(a)$ nous donne l'équation de la tangente au graphe de g au point a :

$$y = g(a) + (x - a)g'(a)$$

C'est la meilleure approximation linéaire du graphe au voisinage de a, c'est à dire la droite qui minimise la distance au graphe de g dans ce voisinage. La dérivée seconde donne la convexité/concavité de la courbe au point a. En particulier, on a que si $g''(a) > 0$ alors la fonction est convexe, tandis que si $g''(a) < 0$ alors la fonction est concave. On retrouve cela dans la trace de la Hessienne. En dimension 1, la dérivée seconde de la fonction g est la courbure du graphe de la fonction g au point a. En dimension 2, les valeurs propres de la Hessienne de f en (a, b) sont en réalité les courbures principales du graphe de la fonction f au point (a, b). La Hessienne est donc la généralisation directe de la dérivée seconde. On peut aussi interprèter la Proposition 1.121 à partir de la notion de diagonalisation.

En effet, en tout point $(a, b, f(a, b))$ d'une surface $z = f(x, y)$, il y a deux direction orthogonales du plan qui sont privilégiées, qu'on notee $\vec{v}_1$ et $\vec{v}_2$, et qui définissent deux chemins du plan $\eta_1 : t \mapsto (a, b) + t\vec{v_1}$ et $\eta_2 : t \mapsto (a, b) + t\vec{v_2}$ qui ont la propriété suivante : tout chemin γ qui passe par $(a, b, f(a, b))$ a une courbure en $(a, b, f(a, b))$ qui est comprise entre la courbure du chemin $\gamma_1 : t \mapsto (\eta_1(t), f(\eta_1(t)))$ et celle du chemin $\gamma_2 : t \mapsto (\eta_2(t), f(\eta_2(t)))$. Autrement dit, au point $(a, b, f(a, b))$, le chemin γ_1 a courbure minimale k_1 et le chemin γ_2 a courbure maximale k_2, et tout chemin γ qui passe par $(a, b, f(a, b))$ a une courbure $k_1 \leq k \leq k_2$ en ce point.



Définition 1.125. *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables deux fois différentiable en $(a,b) \in \mathring{D}$. Supposons en outre que $(a,b) \in \mathring{D}$ est un point critique de f. Les valeurs propres de la matrice Hessienne sont appelées* courbures principales *de la surface $z = f(x,y)$ au point $(a,b,f(a,b))$ et sont notées k_1 et k_2. Nous avons en outre les deux notions suivantes :*

— *la* courbure de Gauss *de la surface $z = f(x,y)$ en ce point est donnée par :*

$$K = k_1 k_2 = \det\left(H_{(a,b)}(f)\right)$$

— *la* courbure moyenne *de la surface $z = f(x,y)$ en ce point est donnée par :*

$$H = \frac{k_1 + k_2}{2} = \frac{1}{2}\mathrm{tr}\left(H_{(a,b)}(f)\right)$$

Remarque 1.126. La courbure moyenne et la courbure de Gauss peuvent être définies en tout point de la surface grâce aux courbures principales. Il n'est pas nécessaire d'être dans un point critique (voir Wikipedia), par contre les formules sont plus compliquées. Une surface est dite *minimale* si sa courbure moyenne est nulle en tout point. Cela veut dire que les courbures principales k_1 et k_2 sont opposées l'une de l'autre en tout point, et dans ce cas la courbure de Gauss est négative. Ce sont donc nécessairement des surfaces hyperboliques.

On peut donc récrire la Proposition 1.121 à l'aide des courbures :

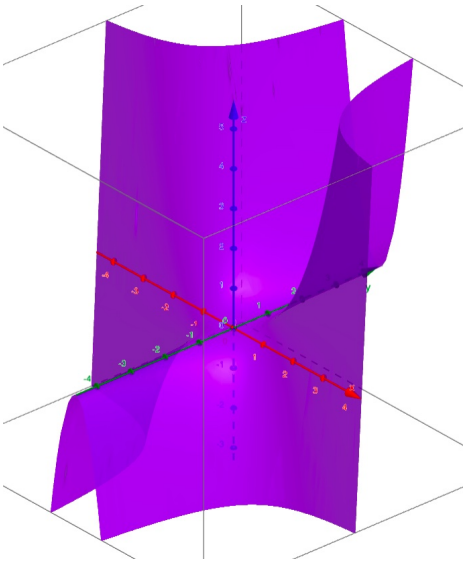
Proposition 1.127. *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction à deux variables deux fois différentiable en $(a,b) \in \mathring{D}$. Supposons en outre que $(a,b) \in \mathring{D}$ est un point critique de f. Alors :*

— *si $K > 0$, nous avons deux cas :*
  1. *si $H > 0$, le point (a,b) est un* minimum local *de f ;*
  2. *si $H < 0$, le point (a,b) est un* maximum local *de f.*

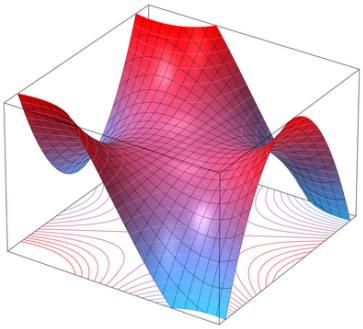
— *si $K < 0$, le point (a,b) est un* point selle.

— *si $K = 0$, on ne peut pas conclure.*

Exemple 1.128. Regardons un cas où $K = 0$. C'est par exemple le cas de la fonction $f(x,y) = yx^2$ dont le graphe est l'union de paraboles d'équations $z = yx^2$. On peut donc voir y comme un paramètre qui modifie les paraboles. Comme $\nabla f(x,y) = \begin{pmatrix} 2xy \\ x^2 \end{pmatrix}$, tout l'axe des y est un ensemble de points critiques, c'est à dire lorsque $x = 0$. Si on marche le long de l'axe y – ou le long de toute droite parallèle à cet axe – on reste à altitude constante. C'est cette direction qui induit la valeur propre nulle. La matrice Hessienne est donnée par $H_{(x,y)}(f) = \begin{pmatrix} 2y & 2x \\ 2x & 0 \end{pmatrix}$. Sur l'axe des y c'est à dire pour $x = 0$, on a $H_{(0,y)}(f) = \begin{pmatrix} 2y & 0 \\ 0 & 0 \end{pmatrix}$. Et on retrouve bien que $K = 0$ mais que $H = 2y$. Dans la direction des y on reste à altitude constante, mais dans la direction des x (à y constant) on monte ou descend une parabole d'équation $z = yx^2$.



Exemple 1.129. Un autre exemple sympathique est la surface appelée "selle de singe" et définie par l'équation $z = x^3 - 3xy^2$. A l'origine $(0,0)$ la Hessienne est nulle, donc elle ne nous dit rien sur la surface à part que la courbure de Gauss et la courbure moyenne sont nulles en ce point. L'origine est ce qu'on appelle un *point ombilic* (voir Wikipedia pour la définition). Le nom de la surface vient du fait qu'une selle pour un singe nécessite trois creux : deux pour les jambes et un pour la queue.



Définition 1.130. Définition de la Hessienne en dimension finie. *On dit qu'une fonction $f : D \subset \mathbb{R}^n \to \mathbb{R}$ est* deux fois différentiable *en un point $A \in \mathring{D}$ si :*

1. *la fonction $f : D \to \mathbb{R}$ est différentiable dans un voisinage ouvert $U \subset D$ contenant A,*
2. *le champ de vecteurs gradient $\nabla f : U \mapsto \mathbb{R}^n$ est différentiable au point A.*

On appelle matrice Hessienne *de la fonction f au point A la transposée de la matrice Jacobienne du champ de vecteurs gradient ∇f au point A :*

$$H_A(f) = (J_A(\nabla f))^t = \begin{pmatrix} \frac{\partial^2 f}{\partial x^1 \partial x^1}(A) & \frac{\partial^2 f}{\partial x^1 \partial x^2}(A) & \dots & \frac{\partial^2 f}{\partial x^1 \partial x^{n-1}}(A) & \frac{\partial^2 f}{\partial x^1 \partial x^n}(A) \\ \frac{\partial^2 f}{\partial x^2 \partial x^1}(A) & & & & \frac{\partial^2 f}{\partial x^2 \partial x^n}(A) \\ \dots & & \dots & & \dots \\ \frac{\partial^2 f}{\partial x^{n-1} \partial x^1}(A) & & & & \frac{\partial^2 f}{\partial x^{n-1} \partial x^n}(A) \\ \frac{\partial^2 f}{\partial x^n \partial x^1}(A) & \frac{\partial^2 f}{\partial x^n \partial x^2}(A) & \dots & \frac{\partial^2 f}{\partial x^n \partial x^{n-1}}(A) & \frac{\partial^2 f}{\partial x^n \partial x^n}(A) \end{pmatrix}$$

Nous définissons alors la différentielle seconde de $f : \mathbb{R}^n \to \mathbb{R}$ en A par la formule suivante :

$$\begin{aligned} d^2 f_A : \ \mathbb{R}^n &\longrightarrow (\mathbb{R}^n)^* \\ v &\longmapsto d^2 f_A(v) : \mathbb{R}^n \longrightarrow \mathbb{R}, \\ &\qquad w \longmapsto v^t \times H_A(f) \times w \end{aligned}$$

Remarque 1.131. Notons bien que la Hessienne d'une fonction n'est définie que si la fonction est à valeurs dans $\mathbb{R}$. D'autre part nous voyons d'après la définition que si $(e_1, \dots, e_n)$ est la base canonique de $\mathbb{R}^n$, et que $v = e_i$ et $w = e_j$ avec $1 \leq i, j \leq n$, alors :

$$d^2 f_A(e_i, e_j) = \frac{\partial^2 f}{\partial x^i \partial x^j}(A)$$

Notons que l'ordre des indices i, j est le même à gauche et à droite. En outre, le Théorème de Schwarz et la Proposition 1.121 sont valides pour les fonctions $f : \mathbb{R}^n \to \mathbb{R}$, où $n \geq 3$.

Définition 1.132. *Soit $f : D \subset \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe au moins $\mathcal{C}^2$, alors elle admet un développement de Taylor-Young à l'ordre 2, dont les formules suivantes sont équivalentes :*

$$\begin{aligned} f(a+h, b+k) &= f(a,b) + df_{(a,b)}\begin{pmatrix} x-a \\ y-b \end{pmatrix} + \frac{1}{2} d^2 f_{(a,b)}\left(\begin{pmatrix} x-a \\ y-b \end{pmatrix}, \begin{pmatrix} x-a \\ y-b \end{pmatrix}\right) + \underset{(x,y)\to(a,b)}{o}\left(\left\|\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right\|^2\right) \\ &= f(a,b) + \left\langle \nabla f(a,b), \begin{pmatrix} x-a \\ y-b \end{pmatrix} \right\rangle + \frac{1}{2}\begin{pmatrix} x-a & y-b \end{pmatrix} \times H_{(a,b)}(f) \times \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \underset{(x,y)\to(a,b)}{o}\left(\left\|\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right\|^2\right) \\ &= f(a,b) + \frac{\partial f}{\partial x}(a,b)(x-a) + \frac{\partial f}{\partial y}(a,b)(y-b) \\ &+ \frac{\partial^2 f}{\partial x^2}(a,b)\frac{(x-a)^2}{2} + \frac{\partial^2 f}{\partial x \partial y}(a,b)(x-a)(y-b) + \frac{\partial^2 f}{\partial y^2}(a,b)\frac{(y-b)^2}{2} + \underset{(x,y)\to(a,b)}{o}\left(\left\|\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right\|^2\right) \end{aligned}$$

En posant $h = x - a$ et $k = y - b$ on a l'écriture équivalente :

$$f(a+h,b+k) = f(a,b) + h\frac{\partial f}{\partial x}(a,b) + k\frac{\partial f}{\partial y}(a,b) \\ + \frac{h^2}{2}\frac{\partial^2 f}{\partial x^2}(a,b) + hk\frac{\partial^2 f}{\partial x \partial y}(a,b) + \frac{k^2}{2}\frac{\partial^2 f}{\partial y^2}(a,b) + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right)$$

Remarque 1.133. On rappelle que la norme euclidienne sur $\mathbb{R}^2$ vaut $\left\|\binom{h}{k}\right\| = \sqrt{h^2+k^2}$.

Le développement de Tayor à l'ordre 1 correspond à l'équation du plan tangent au graphe de f en (a,b) :

$$z = f(a,b) + (x-a)\frac{\partial f}{\partial x}(a,b) + (y-b)\frac{\partial f}{\partial y}(a,b)$$

La différence entre cette altitude z et la valeur de $f(x,y)$ est un petit o de $\sqrt{(x-a)^2+(y-b)^2}$. On peut donc voir le plan tangent comme une approximation linéaire du graphe de f en (a,b). D'autre part, en posant le développement de Taylor à l'ordre 2 peut se voir comme l'équation d'un paraboloïde (ou quadrique), qui approxime le mieux le graphe de f dans un voisinage du point (a,b).

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

La différence entre cette altitude z et la valeur de $f(x,y)$ est un petit o de $(x-a)^2+(y-b)^2$.

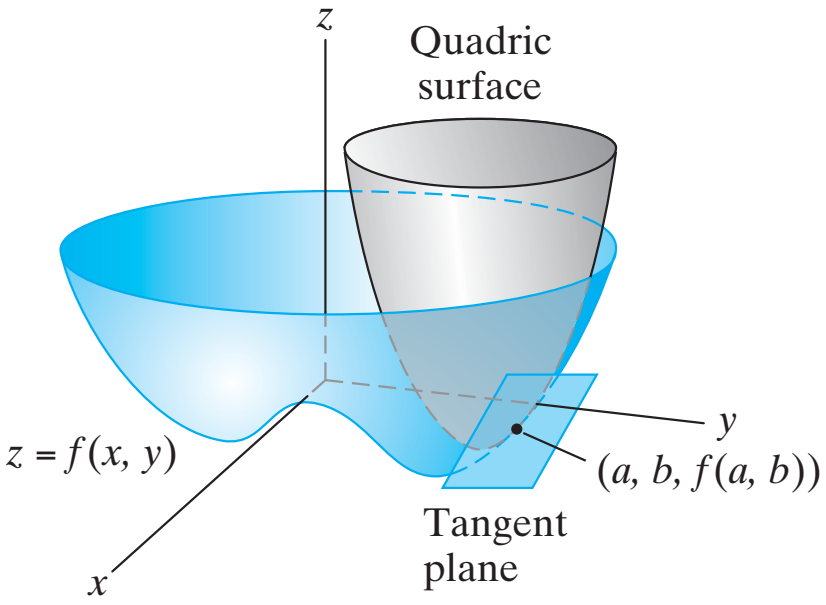


Figure 4.9 The tangent plane and quadric surface.

Exemple 1.134.
Prenons comme fonction à deux variable celle définie par

$$\begin{aligned} f: \ [-1,1]\times[-1,1] &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto \cos(x)\cos(y) \end{aligned}$$

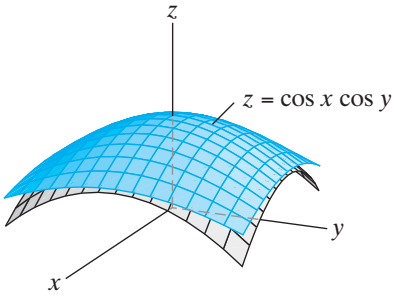


Figure 4.11 The graph of $f(x, y) = \cos x \cos y$ and its Taylor polynomial $p_2(x, y) = 1 - \frac{1}{2}x^2 - \frac{1}{2}y^2$ over the square $\{(x, y) \mid -1 \le x \le 1, -1 \le y \le 1\}$.

Calculons son polynôme de Taylor à l'ordre 2 en $(0,0)$. C'est un point critique donc le gradient y est nul, tandis que la matrice Hessienne au point $(0,0)$ vaut

$$H_{(0,0)}(f) = \begin{pmatrix} -\cos(0)\cos(0) & \sin(0)\sin(0) \\ \sin(0)\sin(0) & -\cos(0)\cos(0) \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$$

Comme $a = b = 0$, on obtient donc, en notant que $h = x - a = x$ et $k = y - b = y$:

$$\begin{aligned} f(x,y) &= f(a,b) + \begin{pmatrix} 0 & 0 \end{pmatrix} \times \begin{pmatrix} h \\ k \end{pmatrix} + \frac{1}{2}\begin{pmatrix} h & k \end{pmatrix} \times \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \times \begin{pmatrix} h \\ k \end{pmatrix} + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right) \\ &= 1 + x \times 0 + y \times 0 - \frac{x^2}{2} \times 1 + xy \times 0 - \frac{y^2}{2} \times 1 + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right) \end{aligned}$$

Le paraboloïde qui approxime le mieux le graphe de la fonction f au voisinage de l'origine $(0,0)$ est donc celui d'équation $z = 1 - \frac{x^2}{2} - \frac{y^2}{2}$.

Exemple 1.135. Nous prenons maintenant la fonction $f(x,y) = x^2+2y^2$ vue dans l'exemple 1.123. Comme c'est un paraboloïde elliptique (défini à partir d'une formule quadratique), nous en déduisons que la meilleure approximation quadratique du graphe de f par la formule de Taylor à l'ordre 2 devrait coïncider avec cette fonction. En effet, soit $(a,b) \in \mathbb{R}^2$, et posons $h = x - a$ et $k = y - b$; nous avons alors :

$$\begin{aligned} f(a+h,b+k) &= f(a,b) + \left\langle \nabla f(a,b), \begin{pmatrix} h \\ k \end{pmatrix} \right\rangle + \frac{1}{2}\begin{pmatrix} h & k \end{pmatrix} \times H_{(a,b)}(f) \times \begin{pmatrix} h \\ k \end{pmatrix} + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right) \\ &= f(a,b) + \begin{pmatrix} 2a & 4b \end{pmatrix} \times \begin{pmatrix} h \\ k \end{pmatrix} + \frac{1}{2}\begin{pmatrix} h & k \end{pmatrix} \times \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \times \begin{pmatrix} h \\ k \end{pmatrix} + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right) \\ &= f(a,b) + 2ah + 4bk + h^2 + 2k^2 + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right) \end{aligned}$$

Remplaçons $f(a,b)$ par a^2+2b^2, ainsi que h et k par leur valeur respective. Nous obtenons alors :

$$\begin{aligned} f(x,y) &= a^2 + 2b^2 + 2a(x-a) + 4b(y-b) + (x-a)^2 + 2(y-b)^2 + \underset{(h,k)\to(0,0)}{o}\left(h^2+k^2\right) \\ &= x^2 + 2y^2 + \underset{(x,y)\to(a,b)}{o}\left(\left\|\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right\|^2\right) \end{aligned}$$

Nous obtenons bien que le développement limité à l'ordre 2 de la fonction polynomiale est elle-même (la fonction petit o est en réalité nulle).

Exemple 1.136. Examinons le développement limité à l'ordre 2 de la fonction $f(x,y) = \ln(x)+y^2$ donnée dans l'exemple 1.72. Soit (a,b) dans le domaine de définition, donc en particulier $a > 0$. La différentielle de f au point (a,b) vaut $df_{(a,b)} = \begin{pmatrix} \frac{1}{a} & 2b \end{pmatrix}$, tandis que la Hessienne vaut $H_{(a,b)}(f) = \begin{pmatrix} -\frac{1}{a^2} & 0 \\ 0 & 2 \end{pmatrix}$. La formule de Taylor nous donne donc, au voisinage du point (a,b) :

$$\begin{aligned} f(x,y) &= f(a,b) + \begin{pmatrix} \frac{1}{a} & 2b \end{pmatrix} \times \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x-a \\ y-b \end{pmatrix}^t \times \begin{pmatrix} -\frac{1}{a^2} & 0 \\ 0 & 2 \end{pmatrix} \times \begin{pmatrix} x-a \\ y-b \end{pmatrix} + \underset{(x,y)\to(a,b)}{o}\left(\left\|\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right\|^2\right) \\ &= \ln(a) + b^2 + \frac{x-a}{a} + 2b(y-b) - \frac{(x-a)^2}{2a^2} + (y-b)^2 + \underset{(x,y)\to(a,b)}{o}\left(\left\|\begin{pmatrix} x-a \\ y-b \end{pmatrix}\right\|^2\right) \end{aligned}$$

La partie d'ordre au plus 1 de cette équation nous donne l'équation du plan tangent au graphe de f au point (a, b) :

$$z = \ln(a) + b^2 + \frac{x-a}{a} + 2b(y-b) = \frac{x}{a} + 2by + \Big(\ln(a) - 1 - b^2\Big)$$

tandis que la partie d'ordre au plus 2 nous donne l'équation du paraboloïde qui approxime le mieux le graphe au voisinage du point (a, b) :

$$z = \ln(a) + b^2 + \frac{x-a}{a} + 2b(y-b) - \frac{(x-a)^2}{2a^2} + (y-b)^2 = -\frac{x^2}{2a^2} + y^2 + \frac{2x}{a} + \left(\ln(a) - \frac{3}{2}\right)$$

2 Intégration des fonctions à deux variables

2.1 Intégrales à paramètres et théorème de Fubini

Rappelons qu'une fonction $u : [a, b] \to \mathbb{R}$ est dite *étagée* ou *en escalier* si elle prend seulement un nombre fini de valeurs, chacune de ces valeurs étant prise sur un nombre fini de sous-intervalles non vides de $[a, b]$, et dans un nombre fini de points isolés. Notamment, il existe $n \in \mathbb{N}^*$ et $n+1$ points du segment $[a, b]$, tels que :

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

Si u est étagée, alors pour tout $1 \leq i \leq n$, il existe $m_i \in \mathbb{R}$ tel que $u(x) = m_i$ pour tout $x \in]x_{i_1}, x_i[$. La valeur de u aux points x_i peut être différente de m_i et m_{i+1}. Toute fonction étagée $u : [a, b] \to \mathbb{R}$ admet une intégrale, définie par :

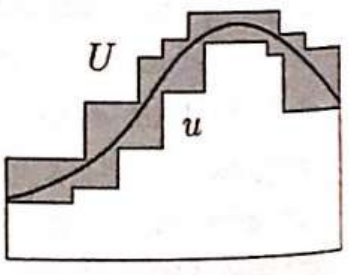
$$\int_a^b u = \int_a^b u(x)\,dx = \sum_{i=1}^{n} m_i(x_i - x_{i-1})$$

On peut vérifier que la valeur de cette intégrale ne dépend pas de la subdivision choisie.

Remarque 2.1. Cela veut dire qu'une fonction en escalier sur $[a, b]$ est une fonction étagée dont les valeurs prises au point a est m_0, au point b est m_n, et au point x_i est soit égale à m_i soit à m_{i+1}, pour tout $1 \leq i \leq n-1$. Autrement dit, une fonction en escalier est une fonction étagée qui est continue à gauche ou à droite en tout point.

Les fonctions étagées nous permettent de définir les intégrales pour des fonctions plus générales. Une fonction $f : [a, b] \to \mathbb{R}$ est *intégrable* (au sens de Riemann) si, pour tout $\epsilon > 0$, il existe deux fonctions étagées $u : [a, b] \to \mathbb{R}$ et $U : [a, b] \to \mathbb{R}$ telles que :

$$u \leq f \leq U \qquad \text{and} \qquad \int_a^b (U - u)(x)\,dx \leq \epsilon$$



Notons que la première condition implique que f est bornée car les fonctions étagées ne prennent qu'un nombre fini de valeurs. Nous allons maintenant donner une caractérisation différente mais équivalente des fonctions intégrables, qui nous permettent de définir la valeur de l'intégrale d'une fonction intégrable. Soit $f : [a, b] \to \mathbb{R}$ une fonction bornée. Alors il existe au moins une fonction étagée $u \leq f$ et une fonction étagée $U \geq f$. Les intégrales des fonctions étagées sur $[a, b]$ sont bien définies donc on peut définir les deux sous-ensembles de $\mathbb{R}$ suivants,

qui sont non-vides :

$$A = \left\{ \int_a^b u(x)\,dx \text{ avec } u : [a,b] \to \mathbb{R} \text{ fonction étagée telle que } u \leq f \right\}$$
$$B = \left\{ \int_a^b U(x)\,dx \text{ avec } U : [a,b] \to \mathbb{R} \text{ fonction étagée telle que } U \geq f \right\}$$

Pour tout $\alpha \in A$, il existe une fonction étagée $u : [a,b] \to \mathbb{R}$ telle que $\alpha = \int_a^b u$, et pour tout $\beta \in B$, il existe une fonction étagée $U : [a,b] \to \mathbb{R}$ telle que $\beta = \int_a^b U$. Comme par définition $u \leq f \leq U$, on a donc $u \leq U$ et donc en intégrant on a que $\alpha \leq \beta$. On en déduit que tous les éléments de A sont inférieurs ou égaux aux éléments de B. En passant à la borne supérieure et inférieure, on en déduit que $\sup(A) \leq \inf(B)$.

Définition 2.2. *Soit $f : [a,b] \to \mathbb{R}$ une fonction bornée. On appelle* intégrale inférieure de f *la borne supérieure* $\sup(A)$ *et* intégrale supérieure de f *la borne inférieure* $\inf(B)$*, et on note :*

$$\int_{[a,b]}^{-} f = \sup(A) \qquad et \qquad \int_{[a,b]}^{+} f = \inf(B)$$

La fonction f est dite Riemann-intégrable *si $\int_{[a,b]}^{-} f = \int_{[a,b]}^{+} f$. Dans ce cas, ce nombre réel est appelé* intégrale de f sur $[a,b]$ *et noté*

$$\int_a^b f \quad ou\ bien \quad \int_a^b f(x)dx$$

On note $\mathcal{R}([a,b])$ l'ensemble des fonctions Riemann-intégrables sur $[a,b]$. Par convention, si $a < b$ on note $\int_b^a f = -\int_a^b f$.

On peut montrer que cette définition de Riemann-intégrabilité avec l'égalité $\sup(A) = \inf(B)$ est équivalente à celle qu'on a donnée plus haut, car la preuve se fait avec les ϵ et la définition des bornes supérieures et inférieures. L'intégrale de $f : [a,b] \to \mathbb{R}$ correspond à l'aire sous la courbe du graphe de f entre les bornes a et b. Il est à noter que les fonctions d'une variable continues par morceaux sur le segment $[a,b]$ sont intégrables au sens de Riemann.

Il existe une technique similaire pour intégrer des fonction à deux variables sur des domaines connexes de $\mathbb{R}^2$. Cela commence par intégrer des fonctions étagées sur des rectangles, puis des fonctions plus générales sur des rectangles, pour enfin intégrer des fonctions sur des domaines (connexes) plus compliqués. L'idée générale restera que l'intégrale d'une fonction à deux variables correspond au *volume* sous le graphe de cette fonction. Dans cette section nous nous contentons de donner l'idée générale de l'intégration des fonctions à deux variables, puis nous verrons dans la section suivante des cas particuliers en nous appuyant sur l'intégration à une variable.

Reprenons la même idée que pour les fonctions à une variable. Soit $a < b$ et $c < d$ deux paires de nombres réels, et soit $D = [a,b] \times [c,d] \subset \mathbb{R}^2$ un rectangle du plan. On dit que la fonction à deux variables $f : D \to \mathbb{R}$ est une *fonction étagée* si ‘

1. il existe $n \in \mathbb{N}^*$ et $n+1$ points du segment $[a,b]$, tels que :

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

2. il existe $m \in \mathbb{N}^*$ et $m+1$ points du segment $[c,d]$, tels que :

$$c = y_0 < y_1 < x_2 < \ldots < y_{m-1} < y_m = d$$

3. et f est constante sur tous les rectangles ouverts du type $]x_{i-1}, x_i[\times]y_{j-1}, y_j[$.

Si f est étagée, alors pour tout $1 \leq i \leq n$ et $1 \leq j \leq m$, il existe $m_{ij} \in \mathbb{R}$ tel que $f(x, y) = m_{ij}$ pour tout $(x, y) \in]x_{i-1}, x_i[\times]y_{j-1}, y_j[$. Toute fonction étagée $f : D \to \mathbb{R}$ admet alors une intégrale sur le rectangle D, dénotée par le signe de double intégrale $\iint_D f(x, y)\, dxdy$ – aussi dénotée $\iint_D f$ – et définie par :

$$\iint_D f(x, y)\, dxdy = \sum_{i=1}^{n} \sum_{j=1}^{m} m_{ij}(x_i - x_{i-1})(y_j - y_{j-1})$$

On note que la double somme est finie donc le membre de droite est bien défini.

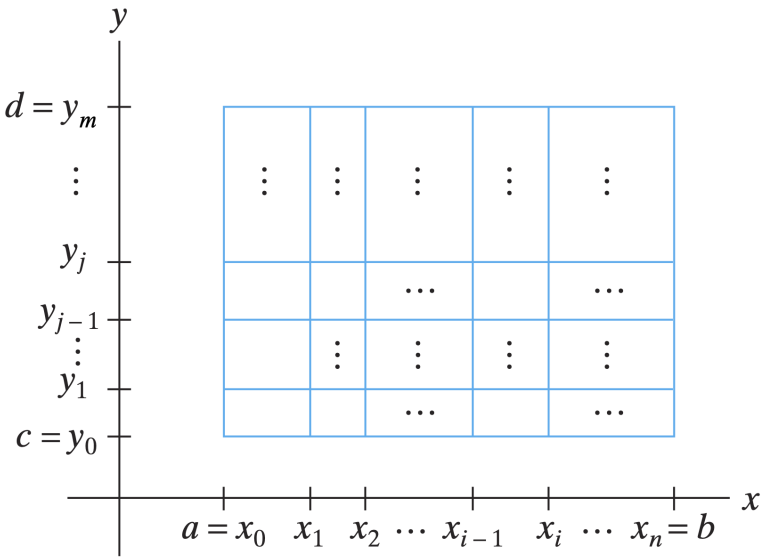


Figure 5.10 A partition of the rectangle $[a, b] \times [c, d]$.

Ensuite, nous pouvons dire qu'une fonction plus générale $f : D \to \mathbb{R}$ est *intégrable* (au sens de Riemann) si pour tout $\epsilon > 0$, il existe deux fonctions étagées $u, U : D \to \mathbb{R}$ telles que $u \leq f \leq U$ et $\iint_D (U - u) \leq \epsilon$. L'intégrale de la fonction f sur le domaine rectangulaire D peut alors se calculer comme dans le cas d'une variable simple. Les intégrales des fonctions étagées sur D sont bien définies donc on peut définir les deux sous-ensembles de $\mathbb{R}$ suivants :

$$A = \left\{ \iint_D u(x, y)\, dxdy \text{ avec } u : D \to \mathbb{R} \text{ fonction étagée telle que } u \leq f \right\}$$
$$B = \left\{ \iint_D U(x, y)\, dxdy \text{ avec } U : D \to \mathbb{R} \text{ fonction étagée telle que } U \geq f \right\}$$

Si f est intégrable sur le domaine rectangulaire D, nous pouvons alors montrer que la borne supérieure de l'ensemble A coincide avec la borne inférieure de l'ensemble B, et nous définissons donc la *double intégrale de f sur D* comme ce nombre unique :

$$\iint_D f(x, y)\, dxdy = \sup(A) = \inf(B)$$

L'intégrale de la fonction f sur le rectangle D consiste alors au *volume* sous le graphe de f (une surface de $\mathbb{R}^3$). A noter que les fonctions de deux variables qui sont continues sur le rectangle D sont intégrables.

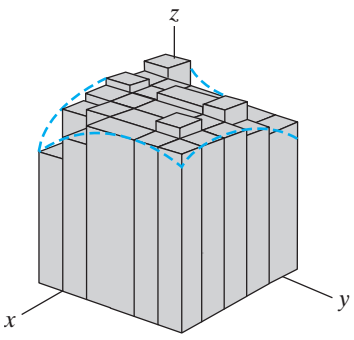


Figure 5.11 The volume under the graph of f is approximated by the Riemann sum.

Remarque 2.3. Bien entendu, comme pour les intégrales des fonctions d'une variable, lorsqu'on dit que c'est le volume 'sous' le graphe de f, cela ne concerne que les fonctions positives. Lorsque la fonction $f : D \to \mathbb{R}$ n'est pas de signe constant sur D, le volume 'sous' le graphe devient le volume sous le graphe sur la partie où f est positive, et le volume au dessus du graphe pour la partie où f est négative.

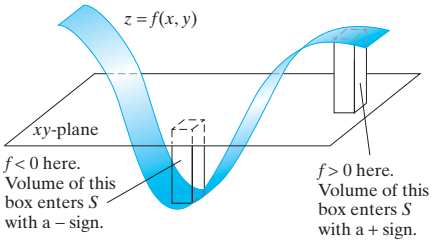


Figure 5.12 The Riemann sum as a signed sum of volumes of boxes.

La théorie générale d'intégration des fonctions à deux variables sur des rectangles est quand même abstraite et nous ne pouvons pas calculer des intégrales dans les cas concrets. Dans la suite de cette section nous allons voir comment calculer des intégrales de fonctions sur des rectangles en utilisant des techniques d'intégration de fonctions à deux variables plus simples que la théorie générale, qui utilisent notamment l'intégrale de Riemann à une variable. Pour cela, nous devons tout d'abord introduire la notion d'intégrale à paramètre.

Soit $I, J \subset \mathbb{R}$ deux intervalle (ouverts, fermés, ou semi-ouverts) . Le produit cartésien $I \times J$ est une partie rectangulaire de $\mathbb{R}^2$. Si I et J sont des intervalles bornés, ce rectangle est borné (potentiellement ouvert sur un ou plusieurs côtés), et si I ou bien J n'est pas borné (c'est à dire s'il contient une borne $\pm\infty$ ou les deux), alors ce rectangle est une bande horizontale ou verticale. Si I et J ne sont pas bornés, cela veut dire que $I \times J$ est un quart de plan ou un demi-plan. Pour l'instant commençons par étudier le cas $I \times [c, d]$, puis nous étudierons le cas plus général $I \times J$.

Soit $f : I \times [c, d] \to \mathbb{R}$ une fonction à deux variables (pas forcément continue). Pour tout $x \in I$ on définit la fonction partielle à une variable $f_x : [c, d] \to \mathbb{R}$ par la formule $f_x(y) = f(x, y)$. Elle n'est pas forcément continue. Si la fonction $f_x : [c, d] \to \mathbb{R}$ est intégrable sur le segment $[c, d]$ (au sens vu ci dessus), alors son intégrale :

$$\varphi(x) = \int_c^d f_x(y)\, dy$$

est une fonction (pas nécessairement continue) de la variable $x \in I$. On dit que c'est une intégrale dépendant du paramètre x. Comme d'habitude dans l'étude de toute fonction, on s'intéresse à savoir si φ est continue, dérivable ou intégrable. Pour cela, il nous faudra des théorèmes similaires à ceux vu sur les séries de fonctions. Nous avons les résultats suivants, que nous ne démontrerons pas (la preuve se trouve dans le livre de Liret-Martinais dans le chapitre sur les intégrales à paramètres) :

Proposition 2.4. *Soit $f : I \times [c, d] \to \mathbb{R}$ une fonction continue de deux variables. En définissant la fonction $\varphi : I \to \mathbb{R}$ par la formule*

$$\varphi(x) = \int_c^d f(x, y)\, dy$$

nous avons les résultats suivants :

1. *la fonction φ est une fonction continue ;*
2. *si la fonction dérivée partielle $(x, y) \mapsto \frac{\partial f}{\partial x}(x, y)$ est aussi continue sur $I \times [c, d]$ alors la fonction φ est de classe $\mathcal{C}^1$ sur $\mathring{I}$ et pour tout $x \in \mathring{I}$, on a :*

$$\varphi'(x) = \int_c^d \frac{\partial f}{\partial x}(x, y)\, dy$$

Exemple 2.5. Soit la fonction à deux variables définie sur $\mathbb{R} \times [0,1]$ par $f(x,y) = \frac{e^{-xy}}{1+y^2}$. On pose $\varphi : \mathbb{R} \to \mathbb{R}$ comme la fonction $\varphi(x) = \int_0^1 f(x,y)dy$. Montrer que f est de classe $\mathcal{C}^2$ et calculer $\varphi(0)$ et $\varphi'(0)$. Montrer que φ est strictement décroissante et convexe. Montrer que

$$\lim_{x\to-\infty} \varphi(x) = +\infty \qquad \text{et} \qquad \lim_{x\to+\infty} \varphi(x) = 0$$

Les solutions sont pages 231-233 du Liret-Martinais.

Exemple 2.6. Soit $h : \mathbb{R} \to \mathbb{R}$ une fonction de classe $\mathcal{C}^\infty$ telle que $h(0) = 0$. Définissons la fonction suivante :

$$\begin{aligned} g : \ \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \begin{cases} \frac{h(x)}{x} & \text{si } x \neq 0 \\ h'(0) & \text{si } x = 0 \end{cases} \end{aligned}$$

Montrons que la fonction g est de classe $\mathcal{C}^\infty$. On peut voir que :

$$\forall\, x \in \mathbb{R} \qquad g(x) = \int_0^1 h'(xy)dy$$

Définissons alors la fonction à deux variables suivante :

$$\begin{aligned} f : \ \mathbb{R} \times [0,1] &\longrightarrow \mathbb{R} \\ (x,y) &\longmapsto h'(xy) \end{aligned}$$

Si on la dérive k fois on a :

$$\forall\, x \in \mathbb{R},\ \forall\, y \in [0,1], \qquad \frac{\partial^k f}{\partial x^k}(x,y) = y^k h'(xy)$$

C'est une fonction continue pour tout $k \in \mathbb{N}$ donc les conditions de la Proposition 2.4 s'appliquent et on voit que la fonction g est de classe $\mathcal{C}^\infty$.

Ce premier résultat nous dit quand la fonction φ est continue et dérivable, dès qu'on intègre f sur le segment $[c,d]$. Dans la suite, nous allons étendre le cas où le segment d'intégration de la variable y est un intervalle plus général, semi-ouvert d'un côté (la borne supérieure par exemple). On peut aussi ouvrir du côté de la borne inférieur avec les mêmes arguments que nous présentons dans la suite de cette section. Dans les deux cas, la continuité et la dérivabilité de la fonction φ est dès lors plus compliquée à montrer. Nous faisons cela car il arrivera d'intégrer sur des rectangles ouverts d'un côté, du type $[a,b] \times [c,d[$ ou de deux côtés, du type $]a,b] \times [c,d[$. Il faut donc savoir définir les intégrales à paramètres dans ce genre de cas. Pour cela revenons sur les intégrales impropres.

Rappelons que dans l'intégration des fonctions réelles à une variable, les intégrales du type $\int_c^d f(x)\,dx$, lorsque f est une fonction définie sur l'intervalle ouvert $]c,d[$ ou semi-ouvert $[c,d[$ où c et d peuvent être un nombre réel ou $\pm\infty$, sont appelées *intégrales impropres*. On ne peut en effet pas utiliser directement la technique usuelle d'intégration avec les fonctions étagées sur ces intervalles pour montrer qu'une fonction est intégrable ou pas intégrable. Ce qu'on fait à la place c'est le processus suivant : on considère le cas où f est *localement intégrable* sur un intervalle J ouvert ou semi-ouvert, c'est à dire intégrable sur tout segment de J (c'est le cas en particulier des fonctions étagées et continues). On note $c = \inf(J)$ et $d = \sup(J)$, ces valeurs pouvant éventuellement valoir l'infini si J n'est pas borné à gauche ou à droite. Avec ces notations, on intègre f sur un segment $[t_-, t_+] \subset J$, et on regarde si la limite de la fonction primitive de f :

$$\lim_{\substack{t_+\to d \\ t_-\to c}} \int_{t_-}^{t_+} f(x)\,dx$$

existe et est finie (c'est à dire un nombre réel). Si tel est le cas, on dit que l'intégrale impropre de f sur J *converge*, et on note la limite $\int_J f(x)\,dx$ ou $\int_c^d f(x)\,dx$. Nous avons un critère pour savoir si une fonction f est intégrable sur un intervalle ouvert ou semi-ouvert J, c'est le critère de convergence majorée : si $f : J \to \mathbb{R}$ est majorée en valeur absolue par une fonction positive $g : J \to \mathbb{R}$ – c'est à dire $|f| \leq g$ – et si l'intégrale impropre $\int_c^d g(x)\,dx$ est convergente, alors l'intégrale impropre $\int_c^d f(x)\,dx$ est convergente (en réalité absolument convergente, comme pour les séries).

Maintenant pour les intégrales à paramètres, nous étudions une fonction f à deux variables continue définie sur $I \times J$ où I et J sont des intervalles de $\mathbb{R}$ (ouvert, fermé ou semi-ouvert). Cette fois-ci, la fonction $\varphi(x) = \int_c^d f(x,y)\,dy$ n'est bien définie qu'à condition que pour tout $x \in I$, l'intégrale impropre $\int_c^d f(x,y)\,dy$ est convergente. La continuité de la fonction φ n'est pas encore garantie. Voici un critère qui permet d'être sûr que φ est définie et continue :

Proposition 2.7. *Soit $f : I \times J \to \mathbb{R}$ une fonction continue de deux variables. Supposons qu'il existe une fonction $g : J \to \mathbb{R}$ continue (et positive) ayant les deux propriétés suivantes :*

1. *pour tout $y \in J$, $|f(x,y)| \leq g(y)$ pour tout $x \in I$, et*
2. *l'intégrale impropre $\int_J g(y)\,dy$ est convergente.*

Alors pour tout $x \in I$, l'intégrale impropre $\int_J f(x,y)\,dy$ est absolument convergente et la fonction $\varphi : I \to \mathbb{R}, x \mapsto \int_J f(x,y)\,dy$ est continue.

Remarque 2.8. Notons que la proposition s'applique automatiquement si la fonction f est bornée sur $I \times J$. Si ce n'est pas le cas, on peut aussi se contenter de prendre les hypothèses de la Proposition 2.7 sur un segment de J pour se ramener à la Proposition 2.4, car la continuité est une propriété locale.

Exemple 2.9. On prend $D = \left(\mathbb{R}_+^*\right)^2 \subset \mathbb{R}^2$ et la fonction définie par $f(x,y) = y^{x-1}e^{-y}$. On définit la fonction Γ *d'Euler* par l'unique fonction définie par :

$$\begin{array}{rcl} \Gamma : \,]0,+\infty[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \displaystyle\int_0^{+\infty} y^{x-1}e^{-y}dy \end{array}$$

D'abord, elle est bien définie car pour tout $x > 0$ nous avons :

$$y^{x-1}e^{-y} \underset{y\to 0}{\sim} \frac{1}{y^{1-x}} \qquad \text{et} \qquad y^{x-1}e^{-y} \underset{y\to+\infty}{=} o\left(\frac{1}{y^2}\right)$$

L'équivalent de gauche est une fonction intégrable en 0 car $1 - x < 0$, et le petit o de droite est clairement intégrable en $+\infty$. L'intégrale de la fonction Γ est donc bien définie pour tout $x > 0$. D'autre part la fonction f est continue sur D. Pour la majoration, on ne peut pas la majorer proprement sur D, mais on va le faire sur un sous-ensemble, en prenant un segment $[a,b] \subset]0,+\infty[$. Car dans ce cas on a, pour tout $x \in [a,b]$ et tout $y \in]0,+\infty[$, que :

$$|f(x,y)| \leq \begin{cases} y^{b-1}e^{-y} & \text{si } y \geq 1 \\ y^{a-1}e^{-y} & \text{si } y < 1 \end{cases} \leq \left(y^{b-1} + y^{a-1}\right)e^{-y}$$

La fonction de droite $g(y) = \left(y^{b-1} + y^{a-1}\right)e^{-y}$ est continue, ne dépend pas de x, majore $|f(x,y)|$ et est intégrable sur $]0,+\infty[$ car $1 - a < 1$ et $1 - b < 1$. La Proposition 2.7 s'applique et on en déduit que Γ est continue sur $[a,b]$. Comme c'est vrai pour tout $0 < a < b$, on a que la fonction Γ est continue sur $]0,+\infty[$.

Remarque 2.10. On observe que pour tout $x > 0$, $\Gamma(x+1) = x\Gamma(x)$, et on en déduit que $\Gamma(n) = n!$.

Exemple 2.11. L'hypothèse que f est dominée par une fonction indépendante de la première variable est importante car si on prend la fonction définie par $f(x,y) = xe^{-xy}$ sur $(\mathbb{R}_+)^2 \subset \mathbb{R}^2$ qui est continue, alors son intégrale vaut $\varphi(0) = 0$ et $\varphi(x) = 1$ si $x > 0$. La fonction φ n'est donc pas continue !

Notez l'analogie avec la convergence normale dans le domaine des séries de fonctions. Soit $(f_n : I \to \mathbb{R})_n$ une suite de fonctions continues. Si jamais il existe une suite d'entiers positifs ayant les deux propriétés suivantes :

1. pour tout $n \in \mathbb{N}$ fixé, $|f_n(x)| \leq a_n$ pour tout $x \in I$ – autrement dit $\|f_n\| \leq a_n$ – et
2. la série $\sum a_n$ est convergente.

On dit alors que la série de fonctions converge normalement. Cela implique que pour tout $x \in I$, la série de nombres réels $\sum f_n(x)$ converge absolument, et donc que la série de fonctions $\sum f_n$ converge simplement vers une fonction $f = \sum_{n=0}^{+\infty} f_n$, et cette fonction est continue. Nous voyons donc que la Proposition 2.7 est une version 'continue' de la convergence normale pour les séries de fonctions.

De cette remarque, nous pouvons généraliser le théorème de dérivation sous le signe somme pour les séries de fonctions, c'est à dire l'interversion de la somme et la dérivation, sous certaines conditions :

$$\left(\sum_n^{+\infty} f_n\right)' = \sum_n^{+\infty} f_n'$$

Cela est possible notamment (mais pas que) lorsque chaque fonction dérivée f_n' est majorée sur sont intervalle de définition par un nombre réel a_n, et que la série positive $\sum a_n$ converge. Si on généralise ces hypothèses en passant du discret ($n \in \mathbb{N}$) au continu ($y \in J$), alors on obtient le résultat suivant :

Proposition 2.12. *Soit $f : I \times J \to \mathbb{R}$ une fonction continue de deux variables (on note $c = \inf(J)$ et $d = \sup(J)$). Supposons que la fonction dérivée partielle $\frac{\partial f}{\partial x}$ existe et est continue sur $\mathring{I}$. On suppose que l'intégrale impropre $\int_c^d f(x,y)\,dy$ converge pour tout $x \in I$ et qu'il existe une fonction $g : J \to \mathbb{R}$ continue (et positive) ayant les deux propriétés suivantes :*

1. *pour tout $y \in J$, $|\frac{\partial f}{\partial x}(x,y)| \leq g(y)$ pour tout $x \in \mathring{I}$, et*
2. *l'intégrale impropre $\int_c^d g(y)\,dy$ est convergente.*

Alors la fonction $\varphi : x \mapsto \int_c^d f(x,y)\,dy$ est de classe C^1 sur $\mathring{I}$ et on a, pour tout $x \in \mathring{I}$:

$$\varphi'(x) = \int_c^d \frac{\partial f}{\partial x}(x,y)\,dy$$

Remarque 2.13. Notons en outre que l'intégrale impropre $\int_c^d \frac{\partial f}{\partial x}(x,y)\,dy$ est absolument convergente.

Exemple 2.14. On prend $I = \mathbb{R}$, $c = 0$ et $d = +\infty$. La fonction que l'on étudie est $f(x,y) = e^{-y^2}\cos(xy)$. Montrer que l'intégrale impropre $\int_0^{+\infty} f(x,y)\,dy$ est (absolument) convergente. On pose ensuite :

$$\varphi(x) = \int_0^{+\infty} e^{-y^2}\cos(xy)\,dy$$

Montrer que φ est de classe C^1.

Exemple 2.15. Nous allons montrer que la fonction Γ de l'Exemple 2.9 est de classe $\mathcal{C}^\infty$. L'idée c'est de dérivée n fois sous le signe intégrale et de majorer à chaque fois. Rappelons la fonction :

$$\begin{array}{rccc} f: & (]0,+\infty[)^2 & \longrightarrow & \mathbb{R} \\ & (x,y) & \longmapsto & y^{x-1}e^{-y} = e^{-y+(x-1)\ln(y)} \end{array}$$

et la fonction Γ définie par intégration sur la variable y :

$$\begin{array}{rccc} \Gamma : &]0,+\infty[& \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \displaystyle\int_0^{+\infty} y^{x-1}e^{-y}dy \end{array}$$

Pour tout $k \in \mathbb{N}$, on a

$$\frac{\partial^k f}{\partial x^k}(x,y) = (\ln(y))^k y^{x-1}e^{-y}$$

C'est une fonction continue en les deux variables sur $D = \left(]0,+\infty[\right)^2$. Soit $0 < a < b$, alors inspirés par l'Exemple 2.9 on a la majoration suivante :

$$\left|\frac{\partial^k f}{\partial x^k}(x,y)\right| \leq (\ln(y))^k \left(y^{b-1} + y^{a-1}\right) e^{-y}$$

La fonction de droite $g_k(y) = (\ln(y))^k \left(y^{b-1} + y^{a-1}\right) e^{-y}$ est continue, ne dépend pas de x, majore $\left|\frac{\partial^k f}{\partial x^k}\right|$ et est intégrable sur $]0,+\infty[$, donc la Proposition 2.7 nous dit que la fonction intégrale $x \mapsto \int_0^{+\infty}(\ln(y))^k y^{x-1}e^{-y}dy$ est continue. Avec la Proposition 2.12, nous pouvons alors voir par récurrence que c'est la dérivée k-ième de la fonction Γ :

$$\forall\, x \in]0,+\infty[\qquad \Gamma'(x) = \int_0^{+\infty} y^{x-1}e^{-y}dy$$

Nous avons étudié les conditions sous lesquelles l'intégrale à paramètre $\varphi : I \to \mathbb{R}$ était continue et dérivable sur I. Son intégration revient à intégrer une fonction d'une variable sur l'intervalle I (ouvert, semi-ouvert ou fermé). Il n'y a pas de problème majeur à cela, à part vérifier le cas échéant que l'intégrale (impropre ou non) $\int_I \varphi(x)dx$ est bien définie. Notons que dans le cas où la fonction $f : I \times [c,d[\to \mathbb{R}$ est prolongeable par continuité à $I \times [c,d]$, la contribution du bord haut du rectangle $I \times \{d\}$ ne compte pas dans l'évaluation de l'intégrale, car le bord est d'aire nulle par rapport à l'aire du rectangle. Plus généralement, intégrer sur un rectangle fermé ou un rectangle ouvert (si f est bien définie et intégrable) ne change pas le résultat de l'intégrale.

Maintenant, précisons un peu la théorie, et prenons le cas particulier où le domaine d'intégration est un rectangle fermé $[a,b]\times[c,d]$, c'est à dire qu'on pose $I = [a,b]$. Soit $f : [a,b]\times[c,d] \to \mathbb{R}$ une fonction continue définie sur un rectangle fermé. Au vu des hypothèses, nous déduisons de la Proposition 2.4 que la fonction $\varphi : [a,b] \to \mathbb{R}$ définie par

$$\varphi(x) = \int_c^d f(x,y)\,dy$$

est une fonction continue de la variable x. Nous pouvons donc l'intégrer sur le segment $[a,b]$ et nous obtenons :

$$A = \int_a^b \varphi(x)\,dx = \int_a^b \left(\int_c^d f(x,y)\,dy\right) dx$$

Mais pour les mêmes raisons, nous pouvons définir de façon symétrique une fonction $\psi : [c,d] \to \mathbb{R}$ par :

$$\psi(x) = \int_a^b f(x,y)\,dx$$

Elle est aussi continue en la variable y et donc intégrable sur le segment $[c,d]$, ce qui donne :

$$B = \int_c^d \psi(y)\,dy = \int_c^d \left(\int_a^b f(x,y)\,dx\right) dy$$

La question naturelle est de se demander si $A = B$, c'est à dire si on peut intervertir l'ordre d'intégration, ce que nous allons explorer maintenant. Nous voyons qu'il y a deux façons de calculer l'intégrale de la fonction f (continue) sur le rectangle $[a, b] \times [c, d]$:

— intégrer par rapport à y, en laissant x fixé, puis intégrer le résultat par rapport à x, ou

— intégrer par rapport à x, en laissant y fixé, puis intégrer le résultat par rapport à y.

Lorsqu'on fixe $x = x_0 \in [a, b]$, le point (x_0, y) parcourt le segment vertical d'abscisse x_0 lorsque y varie entre c et d (cf la figure 2 ci dessous). Lorsqu'on intègre la fonction f par rapport à la variable y en premier, c'est donc comme si on intégrait la fonction au dessus du segment $\{x_0\} \times [c, d]$ et le résultat – qui correspond à $\varphi(x_0)$ – est l'aire de la tranche verticale au dessus du segment $\{x_0\} \times [c, d]$ (figure verte de la deuxième image). Ensuite, lorsqu'on intègre la fonction φ sur la variable x, on se retrouve à sommer les aires de toutes les tranches vertes. On obtient comme résultat que l'intégrale :

$$A = \int_a^b \varphi(x)\, dx = \int_a^b \left(\int_c^d f(x, y)\, dy \right) dx$$

est le volume sous le graphe de la fonction f.

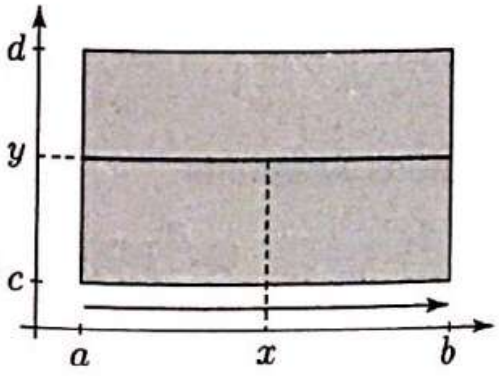


figure 1

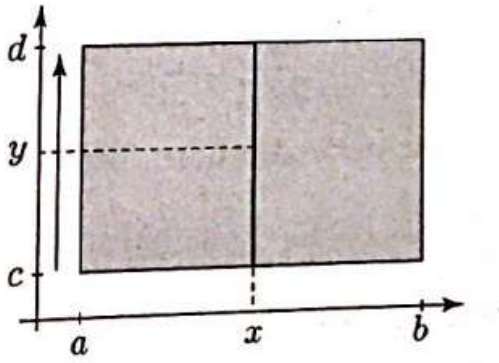
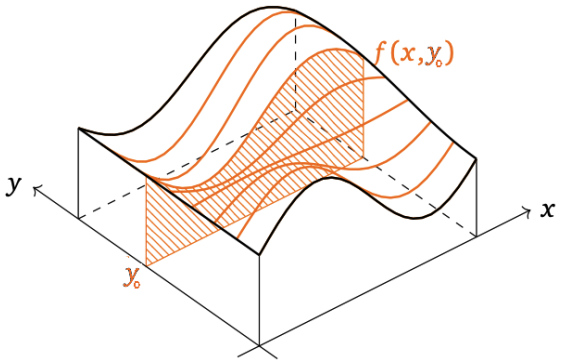
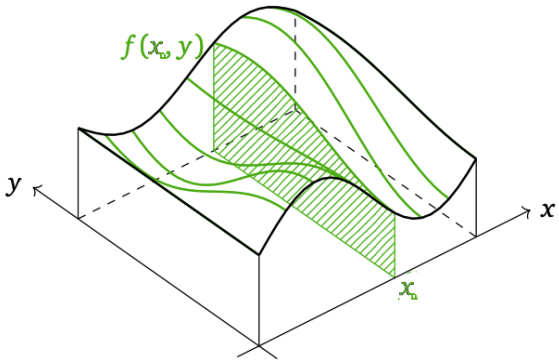


figure 2





Au contraire, dans la seconde manière, lorsqu'on fixe $y = y_0 \in [c, d]$, le point (x, y_0) parcourt le segment horizontal d'ordonnée y_0 lorsque x varie entre a et b (cf la figure 1 ci dessus). Donc lorsqu'on intègre la fonction f par rapport à la variable x en premier, c'est donc comme si on intégrait la fonction au dessus du segment $[a, b] \times \{y_0\}$ et le résultat – qui correspond à $\psi(y_0)$ – est l'aire de la tranche verticale au dessus du segment $[a, b] \times \{y_0\}$ (figure orange de la deuxième

image). Ensuite, lorsqu'on intègre la fonction ψ sur la variable x, on se retrouve à sommer les aires de toutes les tranches oranges. On obtient comme résultat que l'intégrale :

$$B = \int_c^d \psi(y)\,dy = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy$$

est elle aussi le volume sous le graphe de la fonction f. On a donc prouvé graphiquement (intuitivement) que $A = B$.

Theorème de Fubini. *Soit $f : [a,b] \times [c,d] \to \mathbb{R}$ une fonction continue définie sur un rectangle fermé. Alors on a*

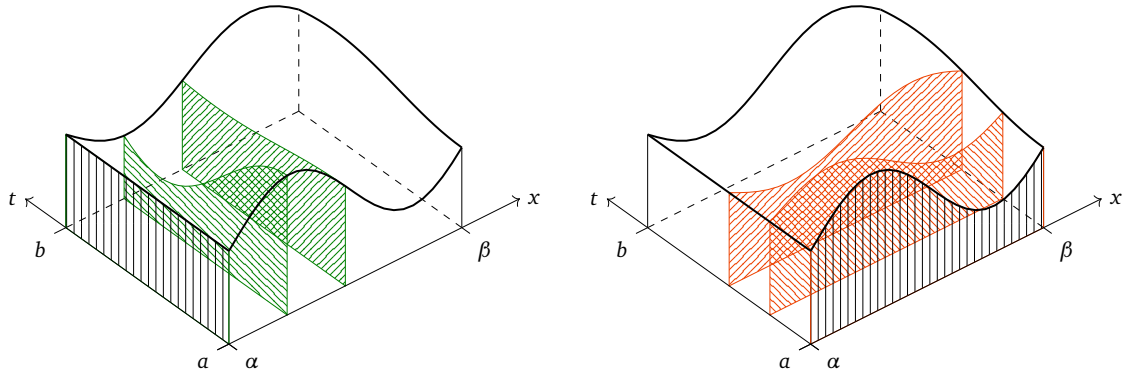
$$\int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy$$

Démonstration. La preuve rigoureuse est intéressante et se trouve pages 250-251 du manuel Liret-Martinais. □

Définition 2.16. *Lorsque $f : [a,b] \times [c,d] \to \mathbb{R}$ est une fonction continue, le nombre*

$$\int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy$$

se note $\iint_{[a,b]\times[c,d]} f(x,y)\,dxdy$ et s'appelle l'intégrale double de f sur $[a,b] \times [c,d]$.

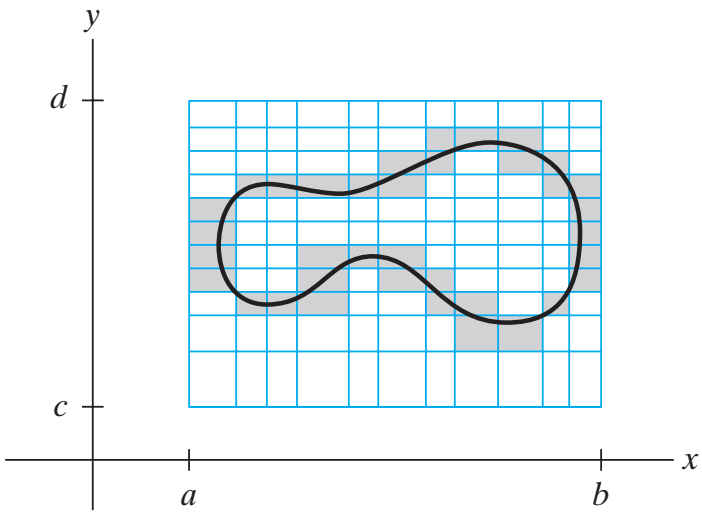


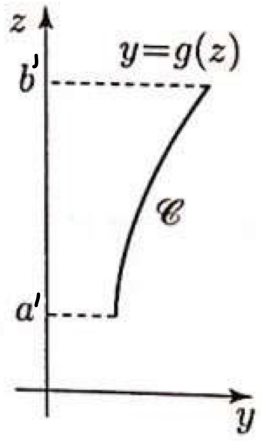
Ici, pour nos fonctions de deux variables, on calcule d'abord l'aire d'une tranche parallèle à l'axe des t (en vert sur la figure), puis on fait la somme (c'est-à-dire on effectue une seconde intégration) des aires de toutes les tranches (qui ont en fait une épaisseur infinitésimale). On pourrait faire la même opération en commençant par les tranches parallèles à l'axe des x (en rouge sur la figure). Le théorème de Fubini affirme que ces deux méthodes conduisent à la même valeur. Ce nombre correspond au volume sous la portion de surface.

2.2 Intégration sur des domaines élémentaires de $\mathbb{R}^2$

Nous avons présenté la théorie générale de l'intégration des fonctions à deux variables sur des rectangles. Il y a un problème qu'il reste à résoudre : comment faire pour des domaines non rectangulaires ? Nous pouvons couper le domaine de définition en plusieurs rectangles, et il restera des zones près des bords qui ne seront pas rectangulaires, mais qui seront du type décrit par le Théorème des fonctions implicites, comme chaque morceau de courbe dans les rectangles de la figure suivante à gauche, avec un zoom sur un morceau dans la figure de droite. Nous savons calculer les intégrales doubles sur les rectangles complets, et nous devons maintenant

définir les intégrales doubles sur les domaines qui sont rectangulaires sur un bord, et bordés par une courbe sur un autre comme sur la figure de droite.





La double intégrale sur les rectangles se fait en deux étapes : d'abord intégrer $f(x,y)$ sur un segment (chemin fermé) vertical ou horizontal, puis intégrer selon la deuxième variable. L'idée d'intégrer d'abord sur un chemin par rapport à une variable puis par rapport à l'autre variable est une technique générale que nous allons maintenant étudier. Pour cela nous pouvons étendre la théorie de l'intégrale à paramètre à des domaines du plan qui ne sont pas des rectangles.

Définition 2.17. *On dit qu'un sous-ensemble $D \subset \mathbb{R}^2$ est un* domaine élémentaire *de $\mathbb{R}^2$ s'il peut être décrit par l'un des deux types suivants :*

— Type 1 : $D = \{(x,y) \,|\, a \leq x \leq b\,, \gamma(x) \leq y \leq \delta(x)\}$ *où* $\gamma, \delta \in \mathcal{C}^0([a,b])$

— Type 2 : $D = \{(x,y) \,|\, \alpha(y) \leq x \leq \beta(y),\ c \leq y \leq d\}$ *où* $\alpha, \beta \in \mathcal{C}^0([c,d])$

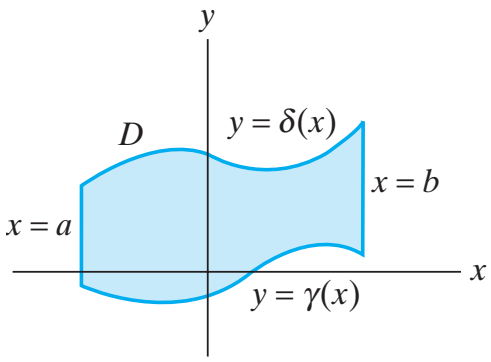


Figure 5.22 A type 1 elementary region.

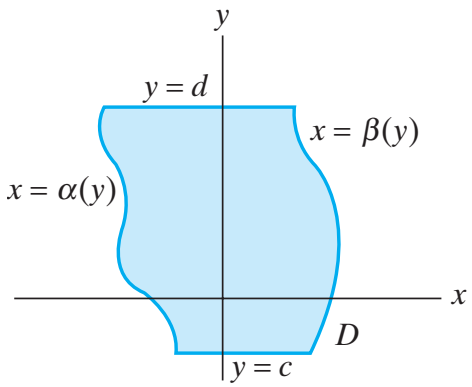
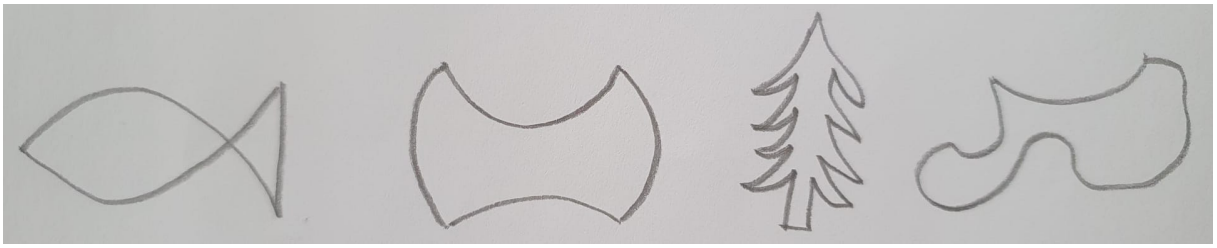


Figure 5.23 A type 2 elementary region.

Remarque 2.18. De la définition, on voit que les domaines élémentaires de type 1 et de type 2 sont *compacts*, c'est à dire *fermés bornés* (car nous sommes en dimension finie), et *connexes (par arc)*, c'est à dire que deux points (a,b) et (a',b') de D sont toujours joignables par une courbe paramétrée. On se limite à ce type de domaine par la suite.

Pour caractériser les domaines de type 1, ce sont les domaines compacts connexes D de $\mathbb{R}^2$ qui satisfont la propriété suivante : pour tout $a \leq x \leq b$, le segment *vertical* $\{x\} \times [\gamma(x), \delta(x)]$ est

entièrement contenu dans D. Cela revient à dire que D peut être recouvert de segments verticaux. Pour caractériser les domaines de type 2, ce sont les domaines compacts connexes D de $\mathbb{R}^2$ qui satisfont la propriété suivante : pour tout $c \leq y \leq d$, le segment *horizontal* $[\alpha(x), \beta(x)] \times \{y\}$ est entièrement contenu dans D. Cela revient à dire que D peut être recouvert de segments horizontaux. Dans les figures ci dessous, le sapin et le dernier dessin ne sont ni de type 1 ni de type 2, le poisson et la deuxième figure sont de type 1 mais pas de type 2.



Définition 2.19. *Soit D un domaine élémentaire de $\mathbb{R}^2$ et $f : D \to \mathbb{R}$ une fonction continue.*

1. *Si D est de type 1, on définit l'intégrale de f sur D comme l'intégrale à paramètre suivante :*

$$\iint_D f dA = \int_a^b \left(\int_{\gamma(x)}^{\delta(x)} f(x,y) dy \right) dx$$

2. *Si D est de type 2, on définit l'intégrale de f sur D comme l'intégrale à paramètre suivante :*

$$\iint_D f dA = \int_c^d \left(\int_{\beta(y)}^{\alpha(y)} f(x,y) dx \right) dy$$

Remarque 2.20. Le symbole dA dénote l'élément d'aire infinitésimal, c'est $dxdy$ en coordonnées cartésiennes. C'est bien une aire, puisque dx et dy sont des longueurs donc $dxdy$ est une aire. On peut donc écrire l'intégrale $\iint_D f dA$ comme $\iint_D f(x,y)\, dxdy$ (c'est juste une notation).

Proposition 2.21. *Soit D un domaine élémentaire de $\mathbb{R}^2$ et $f : D \to \mathbb{R}$ une fonction continue. Si D est à la fois de type 1 et de type 2, alors les deux définitions de l'intégrale de f données ci-dessus coincident.*

Exemple 2.22. Voir les exemples pages 252-253 du manuel Liret-Martinais ou la sous-section 5.2 du livre de Susan Jane Colley. Il est important de noter que souvent, une façon d'intégrer est plus simple que l'autre donc il faut être malin pour choisir la bonne.

Corollaire 2.23. *Soit D un domaine élémentaire de $\mathbb{R}^2$ de type 1 et de type 2. Alors l'intégrale de f sur D mesure le volume (positif ou négatif) "sous" le graphe de f. En particulier, l'aire de D est l'intégrale de la fonction constante $f : D \to \mathbb{R}, (x,y) \to 1$, c'est à dire :*

$$\textit{Aire du domaine } D \; = \iint_D dA$$

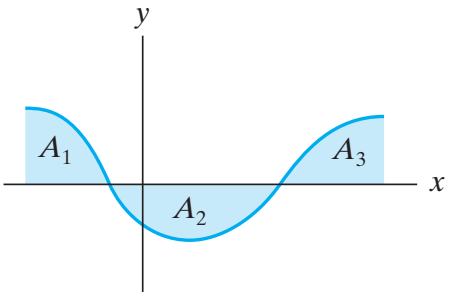


Figure 5.13 If A_1, A_2, A_3 represent the values of the shaded areas, then $\int_a^b f(x)\,dx = A_1 - A_2 + A_3$.

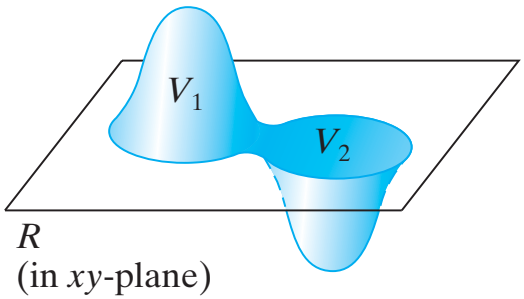
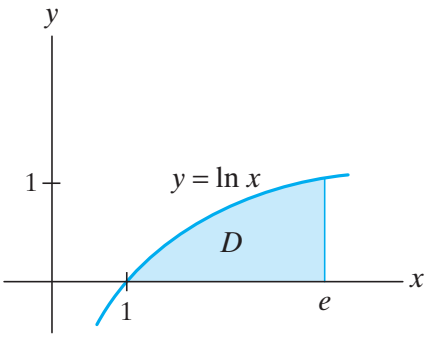


Figure 5.14 If V_1, V_2 represent the volumes of the shaded regions, then $\iint_R f(x, y)\,dA = V_1 - V_2$.

Exemple 2.24. Soit D le domaine délimité par l'axe horizontal et le graphe de la fonction logarithme, entre $x = 1$ et $x = e$ (voir dessin). On prend comme fonction $f : D \to \mathbb{R}, (x, y) \mapsto 1$ la fonction constante à 1. Son intégrale devra donc être l'aire du domaine D considéré. C'est un domaine élémentaire de type 1 et de type 2, avec les fonctions suivantes :



$$\alpha(y) = e^y, \quad \beta(y) = e, \quad \gamma(x) = 0, \quad \delta(x) = \ln(x)$$

On commence par intégrer par rapport à la variable y puis x :

$$\int_1^e \left(\int_0^{\ln(x)} 1 dy \right) dx = \int_1^e \ln(x)\,dx = [x\ln(x) - x]_1^e = 1$$

On fait ensuite l'intégration en changeant l'ordre d'intégration :

$$\int_0^1 \left(\int_{e^y}^e 1 dx \right) dy = \int_0^1 (e - e^y)\,dy = [ey - e^y]_0^1 = 1$$

On retrouve le même résultat dans les deux cas, et l'aire du domaine D vaut 1.

Exemple 2.25. Soit D le cercle unité et soit $f : D \to \mathbb{R}, (x, y) = 1$. L'intégrale de f devra donc être l'aire du cercle de centre $(0, 0)$ et de rayon 1. On observe que le disque est un domaine élémentaire de type 1 et de type 2. Nous pouvons calculer l'intégrale de f de deux façon différentes, données par la Définition 2.19.

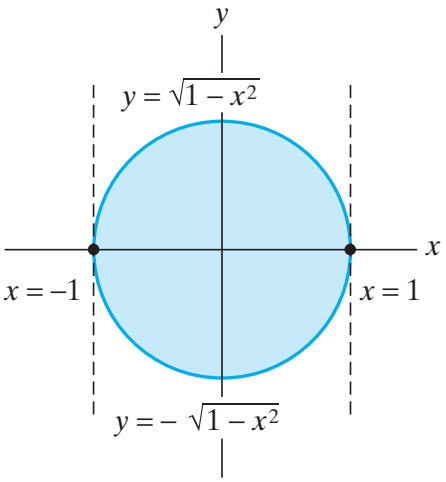


Figure 5.25 The unit disk D as a type 1 region.

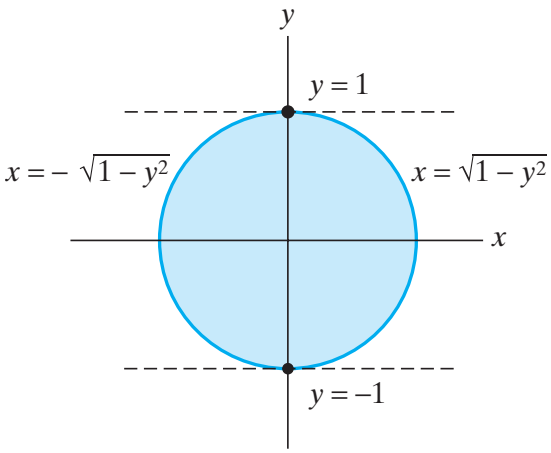


Figure 5.26 The unit disk D as a type 2 region.

Comme montré sur le dessin, nous avons les fonctions suivantes :

$$\alpha(y) = -\sqrt{1-y^2}, \qquad \beta(y) = \sqrt{1-y^2}, \qquad \gamma(x) = -\sqrt{1-x^2}, \qquad \delta(x) = \sqrt{1-x^2}$$

Ceci qui nous donne, pour la première intégrale de la Définition 2.19 :

$$\int_{-1}^{1}\left(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 1dy\right)dx = \int_{-1}^{1}[y]_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}}dx = \int_{-1}^{1} 2\sqrt{1-x^2}dx$$

et on a de manière similaire, en intégrant d'abord par rapport à la variable x :

$$\int_{-1}^{1}\left(\int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} 1dx\right)dy = \int_{-1}^{1}[x]_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}}dx = \int_{-1}^{1} 2\sqrt{1-y^2}dy$$

Ces deux intégrales sont identiques car la variable d'intégration n'a pas d'importance. Pour calculer l'intégrale on doit passer en coordonnées polaires. On le fera plus tard lorsqu'on saura changer de variables.

Exemple 2.26. Pour un calcul de volume, voir aussi pages 256-258 du Liret-Martinais.

Remarque 2.27. Changer l'ordre d'intégration sur un domaine élémentaire D qui est à la fois de type 1 et de type 2 peut rendre possible un calcul d'intégrable auparavant impossible. Prenons par exemple le domaine D défini de façon équivalente par les deux définitions suivantes :

$$D = \{(x,y)\,|\,0 \le x \le 4\,, 0 \le y \le \sqrt{x}\}$$
$$D = \left\{(x,y)\,|\,y^2 \le x \le 4,\ 0 \le y \le 2\right\}$$

Si on note $f : D \to \mathbb{R}, (x,y) \mapsto y\cos(x^2)$, il est impossible de calculer l'intégrale $\int_0^2\left(\int_{y^2}^4 f(x,y)dx\right)dy$ mais on peut facilement calculer $\int_0^4\left(\int_0^{\sqrt{x}} f(x,y)dy\right)dx$ (voir Example 2 de la section 5.3 de Susan Jane Colley).

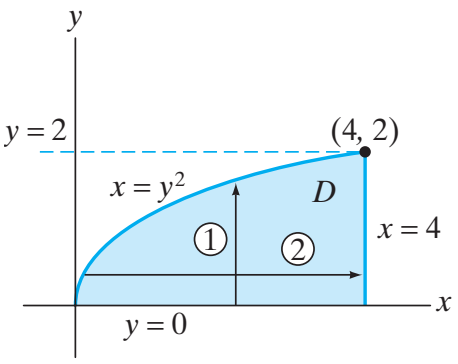


Figure 5.50 Note that $x = y^2$ corresponds to $y = \sqrt{x}$ over the region shown.

On s'intéresse maintenant aux applications inversibles $T : \mathbb{R}^2 \to \mathbb{R}^2$ (pas forcément linéaire). Pour distinguer les deux espaces, on note (u, v) les coordonnées du premier espace et (x, y) les coordonnées du second. On peut noter alors l'application T avec ses fonctions composantes (voir le paragraphe avant la Proposition 1.4) :

$$T(u, v) = (x(u, v), y(u, v))$$

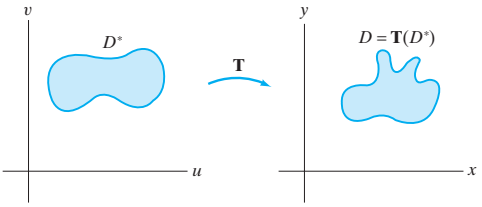


Figure 5.73 The transformation $\mathbf{T}(u, v) = (x(u, v), y(u, v))$ takes the subset D^* in the uv-plane to the subset $D = \{(x, y) \mid (x, y) = \mathbf{T}(u, v) \text{ for some } (u, v) \in D^*\}$ of the xy-plane.

On cherche à étudier comment l'aire d'un domaine change sous la transformation de l'espace induite par un isomorphisme $T : \mathbb{R}^2 \to \mathbb{R}^2$. On va d'abord étudier le cas d'une application linéaire. Une application est linéaire lorsqu'elle peut être représentée par une matrice 2×2. Dans ce cas, l'application est inversible si et seulement si le déterminant de la matrix est non nul. Une rotation, une homothétie ou une symétrie par rapport à un axe sont des applications linéaires inversibles. Une translation n'est pas une application linéaire mais elle est inversible.

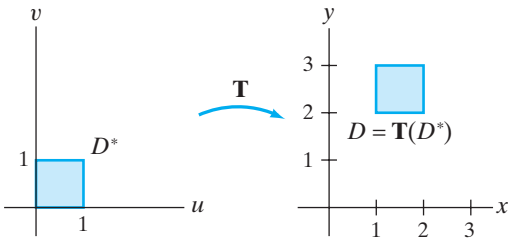


Figure 5.74 The image of $D^* = [0, 1] \times [0, 1]$ is $D = [1, 2] \times [2, 3]$ under the translation $\mathbf{T}(u, v) = (u + 1, v + 2)$ of Example 1.

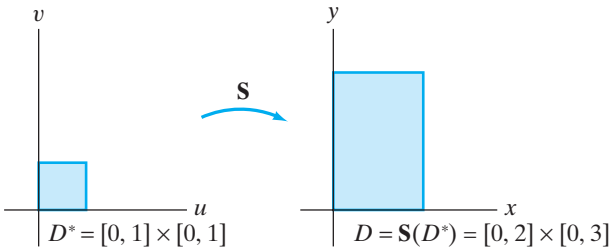


Figure 5.75 The transformation $\mathbf{S}$ of Example 2 is a scaling by a factor of 2 in the horizontal direction and 3 in the vertical direction.

Soit $U : \mathbb{R}^2 \to \mathbb{R}^2$ une application linéaire, représentée par une matrice $U = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. La transformation linéaire envoie

— le point $(1, 0)$ (ou vecteur $\vec{i} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$) sur le point (a, c) (ou vecteur $\vec{a} = \left(\begin{smallmatrix} a \\ c \end{smallmatrix}\right)$),

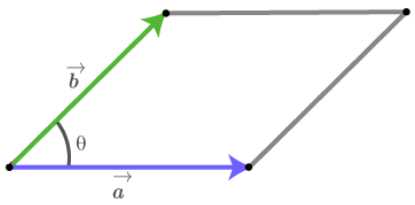
— le point $(0, 1)$ (ou vecteur $\vec{j} = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$) sur le point (b, d) (ou vecteur $\vec{b} = \left(\begin{smallmatrix} b \\ d \end{smallmatrix}\right)$).

Ainsi, le carré $[0, 1] \times [0, 1]$ dont les bords sont donnés par $\vec{i}$ et $\vec{j}$ est envoyé sur un parallélogramme dont les bords sont donnés par $\vec{a}$ et $\vec{b}$. L'aire d'un parallélogramme est donnée par le produit d'un côté par la hauteur, et est donnée par le produit vectoriel entre $\vec{a}$ et $\vec{b}$:

$$\text{Aire} = \|\vec{a}\| \left\|\vec{b}\right\| \left|\sin\left(\theta(\vec{a}, \vec{b})\right)\right|$$

où $\theta(\vec{a}, \vec{b})$ est l'angle entre les deux vecteurs.

$|\vec{a} \times \vec{b}| = |\vec{a}| \cdot |\vec{b}| \sin\theta$



Cette aire peut aussi se calculer avec le déterminant de la matrice U :

$$\text{Aire} = \left|\det \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right| = |ad - bc|$$

Un rectangle d'une aire A donnée devient sous l'action de la transformation U un parallélogramme d'aire $|\det(U)| \times A$. Maintenant, comme expliqué dans le Corollaire 2.23, l'aire d'un domaine est habituellement calculée par l'intégrale sur ce domaine de la fonction constante égale à 1. L'intégrale est une somme infinie de rectangles infinitésimaux, et chaque petit rectangle a

sont aire modifiée par $|\det(U)|$. Ainsi, si on part d'un domaine D^* dans l'espace $\mathbb{R}^2$ de coordonnées (u, v) et qu'on note $D = U(D^*)$ le domaine image dans $\mathbb{R}^2$ de coordonnées (x, y), alors l'aire de D est donnée par :

$$\text{Aire du domaine } D \ = \iint_D dxdy = \iint_{D^*} |\det(U)|dudv = |\det(U)| \iint_{D^*} dudv \tag{2.1}$$

Symboliquement, on a que l'aire du rectangle infinitésimal à l'arrivée – notée $dxdy$ – est égale à l'aire du rectangle infinitésimal au départ – notée $dudv$ – multipliée par la valeur absolue du déterminant de U, c'est à dire :

$$dxdy \simeq |\det(U)|dudv \tag{2.2}$$

Dans notre exemple ici, la transformation U est une application linéaire donc son déterminant est constant et peut sortir de l'intégrale dans l'Equation (2.1) donc on a :

$$\text{Aire du domaine } D \ = |\det(U)| \times \text{Aire du domaine } D^*$$

Cependant dans le cas d'une transformation plus générale $T : \mathbb{R}^2 \to \mathbb{R}^2$ qui n'est pas forcément linéaire, la transformation (2.2) dépendra du point (u, v). On cherche une formule similaire à cette équation, qui implique T d'une façon ou d'une autre. En effet, un rectangle infinitésimal d'aire $dudv$ est envoyé par la transformation T sur un parallélogramme infinitésimal. Pour calculer, son aire, il faut voir que la transformation entre un rectangle infinitésimal et son image est une transformation locale. Cela veut dire que ce n'est pas la fonction T qui va apparaitre mais bien sa différentielle. Et nous pouvons le justifier par notre connaissance du cas des fonctions à une variable où la dérivée est le facteur multiplicateur des intervalle infinitésimaux.

En effet, rappelons comment cela fonctionne en dimension 1. Soit $A < B$ deux réels et soit $g : [A, B] \to \mathbb{R}$ une fonction à une variable, continue donc intégrable sur le segment $[a, b]$. Maintenant supposons que nous avons une fonction bijective $S : [A, B] \to [a, b]$ de classe $\mathcal{C}^1$, à valeurs dans un segment $[a, b]$. Supposons d'abord que $A < B$ et $S(A) = a < b = S(B)$, dans ce cas on a que S est strictement croissante, et donc que $S' > 0$. On peut alors faire un changement de variable dans l'intégrale, en notant x la coordonnée sur $[a, b]$ et u la coordonnée sur $[A, B]$:

$$\int_a^b g(x)dx = \int_a^b g(S(u))d(S(u)) = \int_A^B g(S(u))S'(u)du$$

Notons que si $S(B) = a < b = S(A)$ c'est à dire si la bijection S est décroissante et que $S' < 0$, alors nous avons :

$$\int_a^b g(x)dx = \int_a^b g(S(u))d(S(u)) = \int_B^A g(S(u))S'(u)du = \int_A^B g(S(u))|S'(u)|du$$

Si on synthétise les deux cas, on trouve donc la formule, valide dans tous les cas :

$$\int_a^b g(x)dx = \int_A^B g(S(u))|S'(u)|du \tag{2.3}$$

Nous voyons que la valeur absolue de la dérivée de S multiplie l'élément de longueur infinitésimal dans l'intégrale de droite, de façon à ce qu'on a $dx \simeq |S'(u)|du$. Pour une intégrale double, le raisonnement est le même, c'est la différentielle de la bijection $T : D^* \to D$ qui joue ce rôle. Or la différentielle est une application linéaire, donc comme vu précédemment, le coefficient qui multiplie les aires est $|\det(dT_{(u,v)})|$.

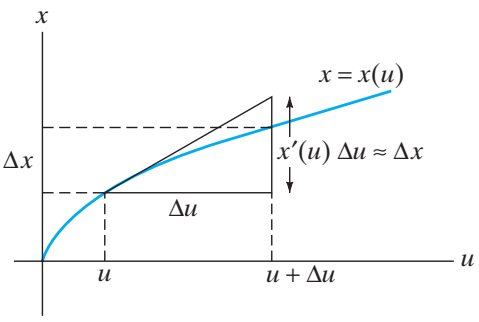


Figure 5.84 As $\Delta u = du \to 0$, $\Delta x \to dx = x'(u)\,\Delta u$. Thus, the factor $x'(u)$ measures how length in the u-direction relates to length in the x-direction.

Théorème 2.28. Changement de variables pour intégrales doubles. *Soit D et D^* deux domaines élémentairs du plan. Soit $T : \mathbb{R}^2 \to \mathbb{R}^2$ une application de classe $\mathcal{C}^1$ qui envoie D^* sur D de façon bijective. Alors pour toute fonction $f : D \to \mathbb{R}$ continue sur D, nous avons le changement de variables suivant sous l'intégrale double :*

$$\iint_D f(x,y)dxdy = \iint_{D^*} f(T(u,v)) \left|\det\left(J_{(u,v)}(T)\right)\right| dudv$$

Démonstration. A proof can be found in pages 362-364 in the book by Susan Jane Colley. □

Corollaire 2.29. *Soit D et D^* deux domaines élémentairs du plan. Soit $T : \mathbb{R}^2 \to \mathbb{R}^2$ une application de classe $\mathcal{C}^1$ qui envoie D^* sur D de façon bijective. Alors l'aire du domaine D peut être calculée comme*

$$\textit{Aire du domaine } D \ = \iint_D dxdy = \iint_{D^*} \left|\det(J_{(u,v)}(T))\right| dudv$$

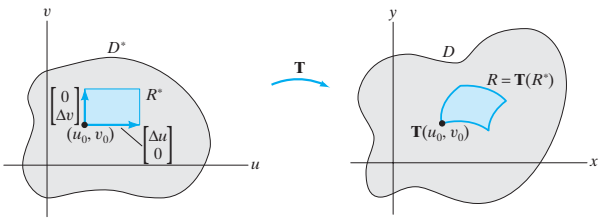


Figure 5.92 **T** takes a rectangle R^* inside D^* to a region R inside D.

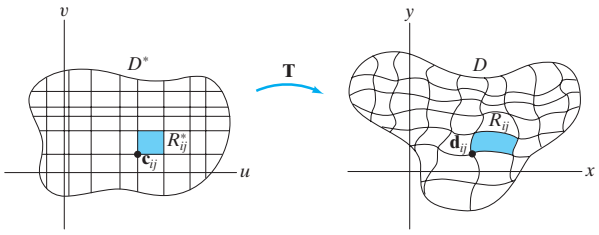


Figure 5.94 A partition of D^* gives rise to a partition of D.

Exemple 2.30. Un exemple important est le passage aux coordonnées polaires. Le rectangle $D^* = [0, R] \times [0, 2\pi[$ dans le plan $\mathbb{R}^2$ de coordonnées (r, θ) est envoyé sur le cercle D, de centre $(0, 0)$ et de rayon R, dans le plan $\mathbb{R}^2$ avec coordonnées (x, y) avec la transformation bijective et de classe $\mathcal{C}^\infty$ suivante :

$$\begin{array}{rccc} T: & D^* & \longrightarrow & D \\ & (r,\theta) & \longmapsto & (r\cos(\theta), r\sin(\theta)) \end{array}$$

L'intégrale d'une fonction f définie sur le cercle de rayon R revient à intégrer sur le rectangle :

$$\iint_D f(x,y)dxdy = \iint_{D^*} f(r\cos(\theta), r\sin(\theta)) \left|\det\left(J_{(r,\theta)}(T)\right)\right| drd\theta$$

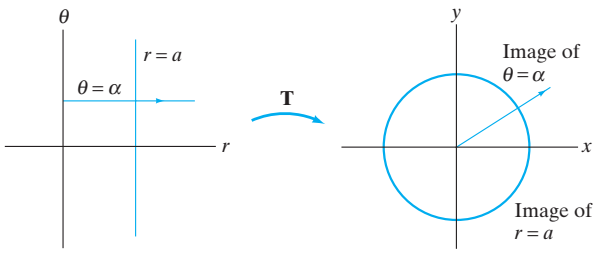


Figure 5.81 The images of lines in the $r\theta$-plane under the transformation $\mathbf{T}(r, \theta) = (r\cos\theta, r\sin\theta)$.

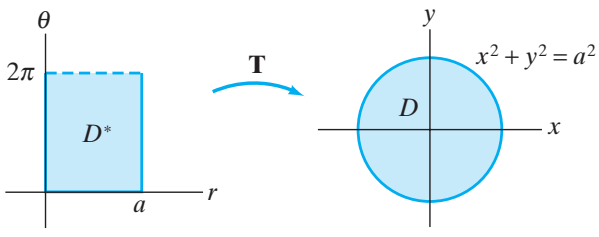


Figure 5.89 **T** maps the (nonclosed) rectangle D^* to the disk D of radius a.

Pour cela il faut calculer la matrice Jacobienne de la fonction T, elle est donnée par :

$$J_{(r,\theta)}(T) = \begin{pmatrix} \frac{\partial T^1}{\partial r} & \frac{\partial T^1}{\partial \theta} \\ \frac{\partial T^2}{\partial r} & \frac{\partial T^2}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{pmatrix}$$

Son déterminant est donc $\det\left(J_{(r,\theta)}(T)\right) = r(\cos(\theta)^2 + \sin(\theta)^2) = r$. On a donc :

$$\iint_D f(x,y)dxdy = \iint_{D^*} f(r\cos(\theta), r\sin(\theta))rdrd\theta$$

En particulier, l'élément d'aire infinitésimal en coordonnées polaires est $rdrd\theta$. En application, calculons l'aire d'un disque de rayon R en nous appuyant sur l'Exemple 2.25 :

$$\int_{-R}^{R}\left(\int_{-\sqrt{R^2-y^2}}^{\sqrt{R^2-y^2}} dx\right)dy = \int_0^{2\pi}\left(\int_0^R rdr\right)d\theta = \int_0^{2\pi}\frac{R^2}{2}d\theta = \pi R^2$$

Remarque 2.31. Plus généralement, pour tout domaine $D \subset \mathbb{R}^2$ qui est borné dans un cercle de rayon R, peut s'écrire comme l'image d'un domaine $D^* \subset \mathbb{R}^2$ inclus dans le rectangle $[0,R]\times[0,2\pi[$. Grâce à cela, on peut souvent calculer une double intégrale facilement, si la fonction à intégrer est sympathique en coordonnées polaires.

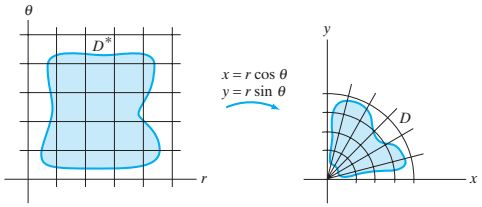


Figure 5.96 The polar-rectangular transformation takes rectangles in the $r\theta$-plane to wedges of disks in the xy-plane.

Exemple 2.32. Nous voulons calculer l'intégrale $\int_0^{+\infty} e^{-t^2}dt$. Pour cela observons que :

$$\left(\int_0^R e^{-t^2}dt\right)^2 = \int_0^R e^{-x^2}dx \times \int_0^R e^{-y^2}dy = \int_0^R\int_0^R e^{-x^2}e^{-y^2}dxdy = \iint_{[0,R]^2} e^{-(x^2+y^2)}dxdy$$

Le carré $[0,R]^2$ est le quart supérieur droit du carré $[-R,R]^2$, et la fonction $(x,y)\mapsto e^{-(x^2+y^2)}$ est intégrable sur chaque quart et :

$$\begin{aligned}\iint_{[0,R]^2} e^{-(x^2+y^2)}dxdy &= \iint_{[0,R]\times[-R,0]} e^{-(x^2+y^2)}dxdy \\ &= \iint_{[-R,0]\times[0,R]} e^{-(x^2+y^2)}dxdy = \iint_{[-R,0]\times[-R,0]} e^{-(x^2+y^2)}dxdy\end{aligned}$$

Ainsi l'intégrale sur le carré complet $[-R,R]^2$ est 4 fois l'intégrale sur le carré $[0,R]^2$. De plus, le carré $[-R,R]^2$ contient le disque de centre $(0,0)$ et de rayon R qu'on note C_R, et est contenu dans le disque de centre $(0,0)$ et de rayon $\sqrt{2}R$, qu'on note $C_{\sqrt{2}R}$. On a alors l'encadrement :

$$\iint_{C_R} e^{-(x^2+y^2)}dxdy \;\le\; 4\iint_{[0,R]^2} e^{-(x^2+y^2)}dxdy \;\le\; \iint_{C_{\sqrt{2}R}} e^{-(x^2+y^2)}dxdy \qquad (2.4)$$

En faisant le changement de variables pour passer aux polaires, on a x^2+y^2 devient r^2 et l'élément d'aire infinitésimal $dxdy$ devient $rdrd\theta$, ce qui donne :

$$\iint_{C_R} e^{-(x^2+y^2)}dxdy = \iint_{[0,R]\times[0,2\pi[} e^{-r^2}rdrd\theta = \int_0^R e^{-r^2}rdr\times\int_0^{2\pi}d\theta = \left[-\frac{e^{-r^2}}{2}\right]_0^R\times[\theta]_0^{2\pi} = \pi\left(1-e^{-R^2}\right)$$

L'encadrement (2.4) peut alors se récrire :

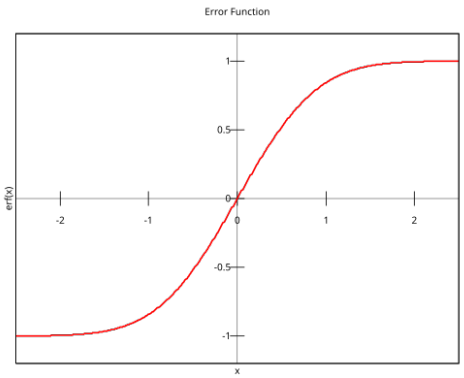
$$\pi\left(1-e^{-R^2}\right) \;\le\; 4\iint_{[0,R]^2} e^{-(x^2+y^2)}dxdy \;\le\; \pi\left(1-e^{-2R^2}\right)$$

En faisant tendre R vers l'infini, le membre de gauche et celui de droite tendent vers π. On en déduit par le théorème des encadrements que :

$$\int_0^{+\infty} e^{-t^2}dt = \frac{\sqrt{\pi}}{2}$$

Remarque 2.33. L'intégrale de l'Exemple 2.32 définit une fonction primitive, qui n'a pas d'expression en termes des fonctions usuelles, et qui s'appelle la *fonction d'erreur* :

$$\forall\, x \geq 0 \qquad \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2} dt$$



Exemple 2.34. On peut trouver plein d'exemples de changements de variables dans la Section 5.3 du livre de Susan Jane Colley.

2.3 Intégrales curvilignes et intégrales vectorielles, théorème de Green

Cette section est dédiée à utiliser tout ce qu'on sait jusqu'ici pour aboutir à un théorème important en géométrie différentielle : le Théorème de Stokes, qui relie une intégrale de surface à une intégrale de contour, sur la frontière de cette surface. Ce théorème, et la dualité "surface $\longleftrightarrow$ frontière" a de profondes implications en physique théorique et en mathématiques. Nous allons donc étudier le cas le plus simple : une surface dont la frontière est une courbe, le tout dans l'espace $\mathbb{R}^3$. Commençons par rappeler certaines informations sur les courbes paramétrées, avec plus de précision dans le cours MAT102.

Définition 2.35. *On dit que qu'une courbe paramétrée $\gamma : I \to \mathbb{R}^2$ de classe (au moins) $\mathcal{C}^1$ est* régulière *si*

$$\forall\, t \in I \qquad \left\| \overrightarrow{\gamma'}(t) \right\| \neq 0$$

On dit qu'une courbe paramétrée $\gamma : I \to \mathbb{R}^2$ régulière est paramétrée normalement *si*

$$\forall\, s \in I \qquad \left\| \overrightarrow{\gamma'}(s) \right\| = 1$$

Dans ce cas on appelle le paramètre s une abscisse curviligne *de l'arc paramétré γ.*

Exemple 2.36. Le temps propre d'une particule en relativité restreinte est une abscisse curviligne de sa ligne d'univers.

Définition 2.37. *Soit I, J deux intervalles ouverts de $\mathbb{R}$. Une bijection $u : I \to J$ est dite $\mathcal{C}^1$-difféomorphisme si la fonction u est de classe $\mathcal{C}^1$ et sa fonction réciproque $u^{-1} : J \to I$ est aussi de classe $\mathcal{C}^1$.*

Une fonction est une bijection si et seulement si elle est strictement monotone (croissante ou décroissante) et dans ce cas sa fonction réciproque a le même sens de variation. Attention être un $\mathcal{C}^k$-difféomorphisme est une condition plus forte car on demande que la fonction réciproque est aussi de classe $\mathcal{C}^k$. Il existe des bijection qui ne sont pas des difféomorphismes. C'est le cas de la fonction $f : x \mapsto x^3$, qui est une bijection continue, mais ce n'est pas un difféomorphisme car la fonction réciproque $f^{-1} : x \mapsto \sqrt[3]{x}$ n'est pas dérivable en 0. Cela vient de la formule :

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

On voit que lorsque x tend vers 0, le membre de droite tend vers l'infini. On comprend donc qu'un difféomorphisme est non seulement strictement monotone, mais que sa dérivée ne s'annule jamais.

Définition 2.38. *Soit $\gamma_1 : I \to \mathbb{R}^2$ et $\gamma_2 : J \to \mathbb{R}^2$ deux courbes paramétrées de classe $\mathcal{C}^1$. On dit que γ_1 est* $\mathcal{C}^1$-équivalente *à γ_2 si il existe un $\mathcal{C}^1$-difféomorphisme $u : I \to J$ tel que*

$$\forall\, t \in I \qquad \gamma_1(t) = \gamma_2(u(t))$$

Dans ce cas on dit que γ_2 est un paramétrage admissible *ou un* reparamétrage *de γ_1 (et inversement). Si le $\mathcal{C}^1$-difféomorphisme $u : I \to J$ est strictement croissant (resp. décroissante), on dit que γ_2 est* positivement *(resp.* négativement*)* équivalente *à γ_1.*

Proposition 2.39. *Toute courbe paramétrée $\gamma_1 : I \to \mathbb{R}^2$ régulière est positivement $\mathcal{C}^1$-équivalente à une courbe $\gamma_2 : J \to \mathbb{R}^2$ de classe $\mathcal{C}^1$ paramétrée normalement. De plus, le $\mathcal{C}^1$-difféomorphisme $u : I \to J$ satisfait :*

$$\forall\, t \in I, \qquad u'(t) = \left\|\overrightarrow{\gamma_1'}(t)\right\|$$

Démonstration. Une abscisse curviligne est une primitive de la fonction $t \mapsto \left\|\overrightarrow{\gamma'}(t)\right\|$. □

Définition 2.40. *On dit qu'une courbe paramétrée $\gamma : [a, b] \to \mathbb{R}^2$ est de classe $\mathcal{C}^1$ (resp. régulière)* par morceaux *si*

1. *il existe $n \in \mathbb{N}^*$ et $n + 1$ points du segment $[a, b]$, tels que :*

$$a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$$

2. *la restriction de γ à chaque intervalle ouvert $]x_{k-1}, x_k[$ est de classe $\mathcal{C}^1$ (resp. régulière) et prolongeable par continuité aux extrémités gauche et droite.*

Soit I un intervalle quelconque. On dit qu'une courbe paramétrée $\gamma : I \to \mathbb{R}^2$ est de classe $\mathcal{C}^1$ (resp. régulière) par morceaux *sur I si elle est de classe $\mathcal{C}^1$ (resp. régulière)* par morceaux *sur tout segment inclus dans I.*

Remarque 2.41. En particulier, aux points x_k, la courbe paramétrée γ_1 est continue mais pas forcément dérivable. Elle peut être cependant dérivable à gauche et à droite de ces points de jonctions, mais ses dérivées à gauche et à droite ne coincident pas forcément, et ne coincident pas forcément non plus avec le prolongement par continuité de la dérivée définie sur les segments ouverts $]x_{k-1}, x_k[$.

Exemple 2.42. Soit $\gamma : \mathbb{R} \to \mathbb{R}^2, t \mapsto (t, |t|)$ la courbe paramétrée associée à la fonction valeur absolue, c'est à dire dont le support est le graphe de la valeur absolue. Si on prend $I_1 =]-\infty, 0]$ et $I_2 =]0, +\infty[$ alors γ est $\mathcal{C}^1$ par morceaux vis à vis de cette partition.

Observons qu'une courbe régulière par morceaux est nécessairement $\mathcal{C}^1$ par morceaux, sur la même subdivision. Notons que l'intérêt de cette définition est que la courbe paramétrée est définie sur un intervalle. Par exemple la parmétrisation habituelle du folium de Descartes (voir cours MAT102) n'est pas définie sur un intervalle, bien qu'elle soit de classe $\mathcal{C}^\infty$ sur son domaine de définition. Dans la suite on suppose que toutes les courbes paramétrées qu'on étudie sont définies sur un intervalle.

Définition 2.43. *Soit $[a, b]$ et $[c, d]$ deux segments et $u : [a, b] \to [c, d]$ un homéomorphisme (c'est à dire une bijection continue dont la réciproque est continue). Soit $\gamma_1 : [a, b] \to \mathbb{R}^2$ et $\gamma_2 : [c, d] \to \mathbb{R}^2$ deux courbes paramétrées de classe $\mathcal{C}^1$ par morceaux. On dit qu'elles sont* $\mathcal{C}^1$-équivalentes vis à vis de $u : [a, b] \to [c, d]$ *si, pour toute subdivision $a = x_0 < x_1 < x_2 < \ldots < x_{n-1} < x_n = b$ de $[a, b]$ qui satisfait la Définition 2.40 pour γ_1, la restriction de $u : [a, b] \to [c, d]$ à chaque intervalle ouvert $]x_k, x_{k+1}[$ – dénotée u_k – satisfait les conditions suivantes :*

— *c'est un $\mathcal{C}^1$-difféomorphisme dont la dérivée est prolongeable par continuité en x_k et x_{k+1},*

— *la restriction de* γ_1 *à* $]x_{k-1}, x_k[$ *est* $\mathcal{C}^1$*-équivalente à la restriction de* γ_2 *à* $u(]x_{k-1}, x_k[)$ *via la restriction* u_k*.*

Soit I *et* J *deux intervalles quelconques et soit* $u : I \to J$ *un homéomorphisme. Soit* $\gamma_1 : I \to \mathbb{R}^2$ *et* $\gamma_2 : J \to \mathbb{R}^2$ *deux courbes paramétrées de classe* $\mathcal{C}^1$ *par morceaux. On dit qu'elles sont* $\mathcal{C}^1$-équivalentes vis à vis de $u : I \to J$ *si, pour tout segment* $[a, b] \subset I$*, elles sont* $\mathcal{C}^1$*-équivalentes vis à vis de* $u : [a, b] \to u([a, b])$*.*

Remarque 2.44. Comme $u : [a, b] \to [c, d]$ est un homéomorphisme, il est strictement croissant ou décroissant.

Définition 2.45. *Soit* $f : D \to \mathbb{R}$ *une fonction continue et soit* $\gamma : I \to D$ *une courbe paramétrée de classe* $\mathcal{C}^1$ *par morceaux. On définit l'*intégrale curviligne scalaire de f le long du chemin γ *l'intégrale sur* I *(si elle existe) de la fonction continue par morceaux* $t \mapsto f(\gamma(t))\left\|\overrightarrow{\gamma'}(t)\right\|$*, et on note :*

$$\int_\gamma f ds = \int_I f(\gamma(t))\left\|\overrightarrow{\gamma'}(t)\right\| dt$$

Remarque 2.46. L'intégrale existe toujours si $I = [a, b]$ car toute fonction continue par morceaux est intégrable sur un segment.

Si γ est paramétré normalement, i.e. si t est une abscisse curviligne – qu'on note habituellement s – alors on a $\left\|\overrightarrow{\gamma'}(s)\right\| = 1$ pour tout s, et on obtient l'intégrale $\int_\gamma f ds = \int_a^b f(\gamma(s))ds$ ce qui fait sens. Et si on a un arc paramétré régulier alors on sait qu'il existe un arc paramétré normalement qui lui est $\mathcal{C}^1$-équivalent par un $\mathcal{C}^1$-difféomorphisme dont la dérivée est la vitesse est $\left\|\overrightarrow{\gamma'}(t)\right\|$ par la Proposition 2.39. Et comme $ds = d(s(t)) = s'(t)dt = \left\|\overrightarrow{\gamma'}(t)\right\|dt$, on retrouve la formule de la Définition 2.45.

Définition 2.47. *Soit* $\vec{F} : D \to \mathbb{R}^2$ *un champ de vecteurs continu, et soit* $\gamma : I \to D$ *une courbe paramétrée de classe* $\mathcal{C}^1$ *par morceaux. On définit l'*intégrale curviligne vectorielle de $\vec{F}$ le long du chemin γ *– ou* circulation de $\vec{F}$ le long du chemin γ *– l'intégrale sur* I *(si elle existe) de la fonction continue par morceaux* $t \mapsto \left\langle \vec{F}(\gamma(t)), \overrightarrow{\gamma'}(t) \right\rangle$*, et on note :*

$$\int_\gamma \vec{F} \cdot d\vec{s} = \int_I \left\langle \vec{F}(\gamma(t)), \overrightarrow{\gamma'}(t) \right\rangle dt \tag{2.5}$$

Remarque 2.48. L'intégrale existe toujours si $I = [a, b]$ car toute fonction continue par morceaux est intégrable sur un segment.

L'intégrale de gauche de la Définition 2.47 est une notation. Si s est une abscisse curviligne, on peut penser à la quantité $d\vec{s}$ comme une quantité vectorielle infinitésimale, c'est à dire comme $ds\,\vec{T}(s)$, où ds est une quantité infinitésimale scalaire (la même que dans la Définition 2.45) et $\vec{T}(s)$ est le vecteur tangent (unitaire) à la courbe au temps s (en supposant un paramétrage normal). Le produit $\vec{F}\cdot d\vec{s}$ est donc le produit scalaire $\left\langle \vec{F}, \vec{T}(s)\right\rangle ds$, qui extrait la composante du champ de vecteurs $\vec{F}$ qui est tangent au vecteur tangent à la courbe $\vec{T}$.
L'interprétation physique des intégrale curvilignes vectorielles est de voir $\vec{F}$ comme une force et γ comme la trajectoire d'une particule, soumise à cette force. L'intégrale $\int_\gamma \vec{F}\cdot d\vec{s}$ peut être compris comme le *travail* effectué par la force $\vec{F}$ sur la particule de trajectoire γ. On mesure ce travail avec la composante du champ de vecteurs $\vec{F}$ qui est tangent au vecteur tangent à la courbe $\vec{T}$.

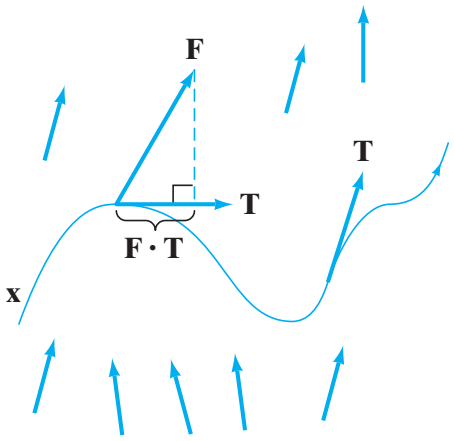


Figure 6.5 The vector line integral $\int_{\mathbf{x}} \mathbf{F}\cdot d\mathbf{s}$ equals $\int_{\mathbf{x}} (\mathbf{F}\cdot\mathbf{T})\,ds$, the scalar line integral of the tangential component of $\mathbf{F}$ along the path.

Proposition 2.49. *Soit $\vec{F} : D \to \mathbb{R}^2$ un champ de vecteurs continu, et soit $\gamma : I \to D$ une courbe paramétrée de classe $\mathcal{C}^1$ par morceaux. En notant $F^1 : D \to \mathbb{R}$ et $F^2 : D \to \mathbb{R}$ les deux composantes du champ de vecteurs $\vec{F}$, on a l'égalité suivante :*

$$\int_\gamma \vec{F}\cdot d\vec{s} = \int_I \Big(F^1(\gamma(t))x'(t) + F^2(\gamma(t))y'(t)\Big)\,dt$$

Remarque 2.50. On note souvent la dernière intégrale $\int_\gamma F^1(x,y)dx + F^2(x,y)dy$ pour souligner que l'intégrale $\int_\gamma \vec{F}\cdot d\vec{s}$ peut en fait se calculer par rapport à une intégrale sur x et une sur y, tant qu'on reste sur la courbe γ. C'est une notation qui est utile pour se rappeler de la formule de la Proposition 2.49, notamment, en se rappelant que $dx = x'(t)dt$ et $dy = y'(t)dt$.

Démonstration. C'est une conséquence directe de l'Equation (2.5). En effet en rappelant que $\gamma(t) = (x(t), y(t))$, on déduit que $\overrightarrow{\gamma'}(t) = \binom{x'(t)}{y'(t)}$, et donc on a :

$$\left\langle \vec{F}(\gamma(t)), \overrightarrow{\gamma'}(t)\right\rangle dt = \Big(F^1(x(t),y(t))x'(t) + F^2(x(t),y(t))y''(t)\Big)\,dt = F^1(x,y)dx + F^2(x,y)dy$$

D'où le résultat. □

Exemple 2.51. Beaucoup d'exemples et d'explications plus précises et profondes dans la Section 6.1 du livre de Susan James Colley.

Il existe une autre fonction dépendant de $\vec{F}$ et de $\overrightarrow{\gamma'}$ que nous pouvons créer et intégrer le long de la courbe paramétrée γ : c'est la fonction déterminant $t \mapsto \det(\vec{F}(\gamma(t)), \gamma'(t))$ où $\det\left(\binom{a}{c}, \binom{b}{d}\right) = ad - bc$ est le déterminant de deux vecteurs, c'est à dire celui de la matrice $\left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right)$. Or ce déterminant d'une matrice 2×2 peut être interprété comme un *produit vectoriel* entre deux vecteurs bidimensionnels (les vecteurs colonnes). Rappelons que le *produit vectoriel* entre deux vecteurs bimensionnels $\binom{a}{c}$ et $\binom{b}{d}$ est le scalaire (nombre réel) défini par :

$$\binom{a}{c} \wedge \binom{b}{d} = ad - bc = \det\begin{pmatrix} a & b \\ c & d\end{pmatrix} \tag{2.6}$$

Attention à ne pas confondre avec le produit vectoriel tridimensionnel défini dans la Définition 1.110, qui donne un vecteur. Cela nous permet de définir une nouvelle intégrale vectorielle.

Définition 2.52. *Soit $\vec{F} : D \to \mathbb{R}^2$ un champ de vecteurs continu, et soit $\gamma : I \to D$ une courbe paramétrée de classe $\mathcal{C}^1$ par morceaux. On définit l'*intégrale curviligne vectorielle de $\vec{F}$ à travers le chemin γ – *ou* flux de $\vec{F}$ à travers le chemin γ – *l'intégrale sur I (si elle existe) de la fonction continue par morceaux $t \mapsto -\vec{F}(\gamma(t)) \wedge \overrightarrow{\gamma'}(t)$, et on note :*

$$\int_\gamma \vec{F} \cdot \vec{n}\, ds = -\int_I \vec{F}(\gamma(t)) \wedge \overrightarrow{\gamma'}(t) dt$$

Le signe devant l'intégrale peut s'expliquer. Le déterminant entre un vecteur $\vec{u}$ et un vecteur $\vec{v}$ revient à prendre "moins" le produit scalaire entre $\vec{u}$ et le vecteur normal à $\vec{v}$ tourné de $\frac{\pi}{2}$ (on peut le vérifier avec les composantes de $\vec{u}$ et $\vec{v}$) :

$$\vec{u} \wedge \vec{v} = -\langle \vec{u}, \vec{v} \rangle$$

Si on note $\vec{N}_\gamma(t)$ le vecteur normal au vecteur vitesse $\overrightarrow{\gamma'}(t)$, tourné de $\frac{\pi}{2}$, et de norme $\left\|\overrightarrow{\gamma'}(t)\right\|$, on a :

$$\int_I \vec{F}(\gamma(t)) \wedge \overrightarrow{\gamma'}(t) dt = -\int_I \left\langle \vec{F}(\gamma(t)), \vec{N}_\gamma(t) \right\rangle dt$$

L'intégrale de la Définition 2.52 mesure donc le *flux* du champ de vecteurs $\vec{F}$ à travers la courbe paramétrée γ. Cette intégrale est donc importante. Maintenant tournons nous vers la reparamétrisation.

Proposition 2.53. *Soit $D \subset \mathbb{R}^2$ un domaine du plan. Soit $\gamma_1 : I \to D$ une courbe paramétrée de classe $\mathcal{C}^1$ par morceaux, et soit $\gamma_2 : J \to D$ une courbe $\mathcal{C}^1$-équivalente à γ_1. Pour toute fonction continue $f : D \to \mathbb{R}$, nous avons l'égalité :*

$$\int_{\gamma_1} f ds = \int_{\gamma_2} f ds$$

Démonstration. On va faire la preuve dans le cas où γ est de classe $\mathcal{C}^1$ sur $[a, b]$, car le cas général se fait en reproduisant le raisonnement sur plusieurs intervalles finis. On note $u : [a, b] \to [c, d]$ le $\mathcal{C}^1$-difféomorphisme tel que $\gamma_1(t) = \gamma_2(u(t))$ pour tout $t \in [a, b]$. En utilisant l'identité $\overrightarrow{\gamma_1'}(u(t)) = u'(t) \times \overrightarrow{\gamma_2'}(u(t)$ dérivée de l'égalité $\gamma_1(t) = \gamma_2(u(t))$, l'Equation (2.3) donne :

$$\begin{aligned}\int_{\gamma_2} f ds = \int_c^d f(\gamma_2(T)) \left\|\overrightarrow{\gamma_2'}(T)\right\| dT &= \int_a^b f(\gamma_2(u(t))) \left\|\overrightarrow{\gamma_2'}(u(t))\right\| |u'(t)| dt \\ &= \int_a^b f(\gamma_1(t)) \left\|\overrightarrow{\gamma_1'}(u(t))\right\| dt = \int_{\gamma_1} f ds\end{aligned}$$

Ceci est bien le résultat demandé. □

Proposition 2.54. *Soit $D \subset \mathbb{R}^2$ un domaine du plan. Soit $\gamma_1 : I \to D$ une courbe paramétrée de classe $\mathcal{C}^1$ par morceaux, et soit $\gamma_2 : J \to D$ une courbe $\mathcal{C}^1$-équivalente à γ_1. Pour tout champ de vecteurs continu $\vec{F} : D \to \mathbb{R}^2$, nous avons les égalités suivantes :*

- *$\int_{\gamma_1} \vec{F} \cdot d\vec{s} = \int_{\gamma_2} \vec{F} \cdot d\vec{s}$ si γ_1 et γ_2 sont positivement orientés ;*
- *$\int_{\gamma_1} \vec{F} \cdot d\vec{s} = -\int_{\gamma_2} \vec{F} \cdot d\vec{s}$ si γ_1 et γ_2 sont négativement orientés.*

Nous avons les mêmes résultats avec l'intégrale $\int_\gamma \vec{F} \cdot \vec{n}\, ds$ de la Définition 2.52.

Démonstration. On reprend les mêmes conventions et observations que dans la preuve de la Proposition 2.53. Le $\mathcal{C}^1$-difféomorphisme $u : [a,b] \to [c,d]$ est soit (strictement) croissant, soit décroissant. Prenons d'abord le cas croissant, donc que $u(a) = c$ et $u(b) = d$ et $|u'(t)| = u'(t)$. On obtient donc, en effectuant un changement de variable et par linéarité du produit scalaire :

$$\begin{aligned}\int_{\gamma_2} \vec{F} \cdot d\vec{s} = \int_c^d \left\langle \vec{F}(\gamma_2(T)), \overrightarrow{\gamma_2'}(T) \right\rangle dT &= \int_a^b \left\langle \vec{F}(\gamma_2(u(t))), \overrightarrow{\gamma_2'}(u(t)) \right\rangle |u'(t)| dt \\ &= \int_a^b \left\langle \vec{F}(\gamma_1(t)), \overrightarrow{\gamma_1'}(t) \right\rangle dt = \int_{\gamma_1} \vec{F} \cdot d\vec{s}\end{aligned}$$

Prenons ensuite le cas décroissant, donc que $u(a) = d$ et $u(b) = c$ et $|u'(t)| = -u'(t)$, et on a :

$$\begin{aligned}\int_{\gamma_2} \vec{F} \cdot d\vec{s} = \int_c^d \left\langle \vec{F}(\gamma_2(T)), \overrightarrow{\gamma_2'}(T) \right\rangle dT &= \int_a^b \left\langle \vec{F}(\gamma_2(u(t))), \overrightarrow{\gamma_2'}(u(t)) \right\rangle |u'(t)| dt \\ &= -\int_a^b \left\langle \vec{F}(\gamma_2(u(t))), \overrightarrow{\gamma_2'}(u(t)) \right\rangle u'(t) dt \\ &= -\int_a^b \left\langle \vec{F}(\gamma_1(t)), \overrightarrow{\gamma_1'}(t) \right\rangle dt = -\int_{\gamma_1} \vec{F} \cdot d\vec{s}\end{aligned}$$

C'est bien le résultat demandé. □

Maintenant observons la chose suivante : supposons que $\gamma : I \to \mathbb{R}^2$ est une courbe paramétrée qui n'est pas injective, c'est à dire qu'il existe $t_1 \neq t_2$ tel que $\gamma(t_1) = \gamma(t_2)$. Cela se manifeste géométriquement pas le fait que le support de γ a un point d'auto-intersection (attention la réciproque est fausse : un point d'auto-intersection du support peut exister même si la courbe paramétrée est injective, voir 2.60). Nous souhaitons travailler sur des courbes qui n'ont pas d'auto-intersection.

Définition 2.55. *On dit qu'une courbe paramétrée $\gamma : [a,b] \to \mathbb{R}^2$ est* fermée *si $\gamma(a) = \gamma(b)$. On dit qu'une courbe paramétrée $\gamma : [a,b] \to \mathbb{R}^2$ est* simple *si elle est injective sur le segment $[a,b]$. On dit qu'une courbe paramétrée $\gamma : [a,b] \to \mathbb{R}^2$ est simple et fermée si elle est fermée et injective sur l'intervalle ouvert $]a,b[$.*

Exemple 2.56. Prenons la courbe $\gamma_0 : [0, 4\pi] \to \mathbb{R} \to \mathbb{R}^2, t \mapsto (\cos(t), \sin(t))$. Son support est le cercle de centre l'origine et de rayon 1 mais elle le parcourt deux fois donc elle n'est pas injective. Ainsi γ_0 est fermée mais pas simple. Par contre les deux courbes suivantes :

$$\gamma_1 : [0, 2\pi] \to \mathbb{R} \to \mathbb{R}^2, t \mapsto (\cos(t), \sin(t)) \quad \text{et} \quad \gamma_2 : [0, 2\pi] \to \mathbb{R} \to \mathbb{R}^2, t \mapsto (\cos(t), -\sin(t))$$

sont simples et fermées. L'une parcourt le cercle dans le sens trigonométrique, et l'autre dans le sens inverse trigonométrique.

Proposition 2.57. *La propriété d'être une courbe simple ou fermée est préservée par $\mathcal{C}^1$-équivalence.*

Définition 2.58. *On appelle* courbe plane *tout sous-ensemble du plan $\mathbb{R}^2$ qui est le support d'une courbe paramétrée. Une courbe plane est dite :*

- fermée *si elle est le support d'une courbe paramétrée fermée ;*
- sans auto-intersection *si elle ne se coupe pas*[2] *;*
- de Jordan *si elle est fermée et sans auto-intersection ;*

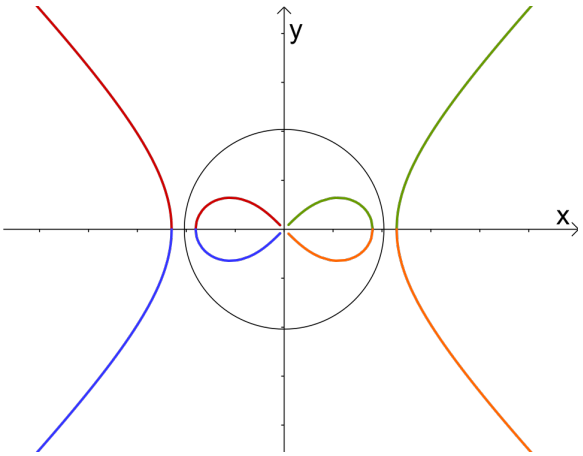
2. Nous ne pouvons pas donner de définition mathématique plus précise à notre niveau pour l'instant.

— lisse *si elle est le support d'une courbe paramétrée de classe* $\mathcal{C}^1$ *;*

— lisse par morceaux *si elle est le support d'une courbe paramétrée de classe* $\mathcal{C}^1$ *par morceaux.*

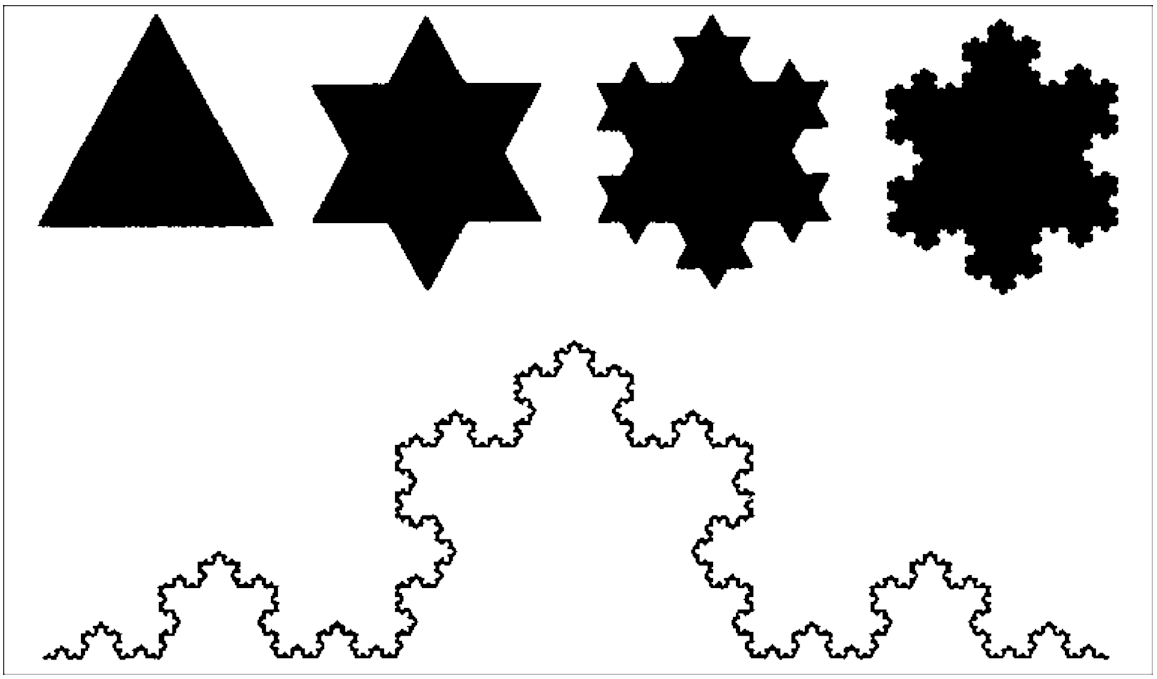
Remarque 2.59. Attention à ne pas confondre *courbe plane* (un sous-ensemble de $\mathbb{R}^2$), et *courbe paramétrée* (une application continue $\gamma : I \to \mathbb{R}^2$). Le support d'une courbe paramétrée est une courbe plane, et toute courbe plane est le support d'une courbe paramétrée. Pour une courbe plane donnée, il peut exister plusieurs courbes paramétrées dont elle est le support (dessin). Ce sont des *paramétrisations* différentes de cette courbe plane.

Remarque 2.60. Le fait d'être sans auto-intersection ne peut pas se traduire uniquement pas l'injectivité de $\gamma : I \to \mathbb{R}^2$. Le Lemniscate de Bernoulli est un exemple d'une courbe paramétrée injective définie sur un intervalle ouvert, et dont le support présente cependant un point d'auto-intersection. Cele vient du fait que son domaine de définition est un intervalle ouvert. Elle est paramétrée par (avec $0 < a < 1$) :



$$\begin{aligned}\gamma : \left]-\tfrac{\pi}{2}, \tfrac{3\pi}{2}\right[&\longrightarrow \mathbb{R}^2 \\ t &\longmapsto \left(a\frac{\cos(t)}{1+\sin(t)^2}, a\frac{\sin(t)\cos(t)}{1+\sin(t)^2}\right)\end{aligned}$$

Exemple 2.61. Notre cerveau a tendance à imaginer des courbes à l'allure très simple alors qu'en réalité il existe des courbes planes sans auto-intersections qui ne sont pas dessinables mais continues. C'est notamment le cas de courbes qui ne sont pas lisses par morceaux comme le *Flocon de Koch* (voir Wikipedia). La longueur de cette courbe est infinie mais l'aire du domaine borné qu'elle délimite est finie. Plus généralement, une courbe de Jordan peut être vue comme un élastique circulaire déformé à volonté, de manière que deux points distincts ne se touchent jamais. L'élastique est supposé être infiniment élastique. Autrement dit, il peut acquérir une longueur infinie. Cela revient à dire que le dessinateur de l'introduction est supposé être capable de tracer une ligne infiniment rapidement avec une précision infinie (voir la page Wikipedia du Théorème de Jordan).



Proposition 2.62. *Une courbe de Jordan est le support d'une courbe paramétrée simple fermée.*

Démonstration. Notons C la courbe plane de Jordan. Comme elle est fermée, il existe une courbe paramétrée fermée $\gamma : [a, b] \to \mathbb{R}^2$ telle que $C = \mathrm{Im}(\gamma)$. Comme elle n'a pas de point d'auto-intersection, l'application γ est injective (sauf aux extrémités a et b) donc simple. □

Soit $\gamma : [a, b] \to \mathbb{R}^2$ une courbe paramétrée simple. La fonction vectorielle γ induit un sens de parcours du support de γ (l'image de γ). Inversement, étant donné une courbe plane sans auto-intersection dénotée C, on peut choisir de la parcourir dans un sens ou dans l'autre. Une *orientation* de la courbe plane C est un choix de sens de parcours de C (il n'en existe que deux car il n'y a pas d'auto-intersections). Comme C est le support d'une courbe paramétrée injective $\gamma : I \to \mathbb{R}^2$, l'orientation de C peut être la même que celle induite par le paramétrage γ – on dit que la paramétrisation est *cohérente* avec l'orientation de C – ou celle inverse à celle induite par la paramétrisation γ, et on dit que que celle-ci est *incohérente* avec l'orientation de C.

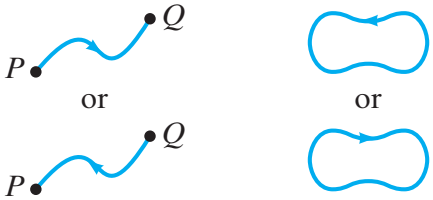


Figure 6.11 The possible orientations for a nonclosed and a closed, simple curve.

Exemple 2.63. Soit C le cercle de centre $(0, 0)$ et de rayon 1. On prend comme orientation le sens de parcours trigonométrique. Reprenons les courbes paramétrées γ_1 et γ_2 de l'Exemple 2.56. La paramétrisation γ_1 est cohérente avec l'orientation de C, et la paramétrisation γ_2 est incohérente.

Remarque 2.64. Par définition, remarquons qu'une courbe orientée est nécessairement sans auto-intersections, car de telles auto-intersections posent problème pour choisir l'orientation. Par la suite, si on parle de courbe plane orientée, on pourra omettre de dire qu'elles sont sans auto-intersection (c'est sous-entendu, comme dans la proposition suivante et après).

Définition 2.65. *Soit C une courbe plane orientée lisse par morceaux et soit $f : \mathbb{R}^2 \to \mathbb{R}$ une fonction continue. On définit l'intégrale de f sur C par*

$$\int_C f ds = \int_\gamma f ds$$

où $\gamma : I \to \mathbb{R}^2$ est n'importe quelle paramétrisation (injective) de classe $\mathcal{C}^1$ par morceaux de la courbe plane C, i.e. telle que $C = \mathrm{Im}(\gamma)$.

Soit $\vec{F} : \mathbb{R}^2 \to \mathbb{R}^2$ un champ de vecteurs continu. On définit l'intégrale de $\vec{F}$ le long de C par :

$$\int_C \vec{F} \cdot d\vec{s} = \int_\gamma \vec{F} \cdot d\vec{s}$$

et l'intégrale de $\vec{F}$ à travers C par :

$$\int_C \vec{F} \cdot \vec{n} ds = \int_\gamma \vec{F} \cdot \vec{n} ds$$

où $\gamma : I \to \mathbb{R}^2$ est n'importe quelle paramétrisation (injective) de classe $\mathcal{C}^1$ par morceaux de la courbe plane C, i.e. telle que $C = \mathrm{Im}(\gamma)$, qui est cohérente *avec l'orientation de C.*

Si la courbe plane orientée lisse par morceaux C est fermée, i.e. si C est une courbe de Jordan, on note $\oint_C$ le signe intégral pour marquer le fait qu'on intègre sur une courbe fermée.

Remarque 2.66. Dans le premier cas, on peut prendre n'importe quelle paramétrisation de C car l'invariance de l'intégrale curviligne scalaire par $\mathcal{C}^1$-équivalence est garantie par la Proposition 2.53. Dans le second cas, l'intégrale curviligne vectorielle change de signe si on a deux choix de courbes paramétrées négativement $\mathcal{C}^1$-équivalentes, comme montré dans la Proposition 2.54.

Exemple 2.67. Pour la courbe C, prenons le cercle de rayon 1 orienté dans le centre trigonométrique et les courbes paramétrées γ_1 et γ_2 de l'Exemple 2.56. Rappelons la discussion sur leur orientation de l'Exemple 2.63. Prenons la fonction vectorielle définie par $\vec{F}(x,y) = x\vec{i} + y\vec{j}$. C'est un champ de vecteur radial (comme les rayons du soleil). Le champ de vecteur est donc orthogonal au cercle C. Nous avons deux intégrales à vérifier. D'abord, l'intégrale le long de C, dont le résultat ne change pas selon qu'on prenne γ_1 ou γ_2 :

$$\int_C \vec{F} \cdot d\vec{s} = \int_{\gamma_i} \vec{F} \cdot d\vec{s} = \int_0^{2\pi} \left\langle \vec{F}(\gamma_i(t)), \overrightarrow{\gamma_2'}(t) \right\rangle dt = 0$$

car on a $\left\langle \vec{F}(\gamma_i(t)), \overrightarrow{\gamma_2'}(t) \right\rangle = 0$ pour tout $t \in [0, 2\pi]$ par orthogonalité.

D'autre part, pour l'intégrale à travers C, nous remarquons d'abord que pour γ_1 qui parcourt le cercle dans le sens trigonométrique, le vecteur normal à la courbe $\vec{N_1}(t)$ est tourné de $\frac{\pi}{2}$ par rapport au vecteur vitesse $\overrightarrow{\gamma_1'}(t)$ et pointe vers l'intérieur du cercle. Tandis que pour γ_2, qui parcourt le cercle dans le sens horaire, le vecteur normal à la courbe $\vec{N_2}(t)$ est tourné de $\frac{\pi}{2}$ par rapport au vecteur vitesse $\overrightarrow{\gamma_2'}(t)$ et pointe donc vers l'extérieur du cercle. Plus précisément on a donc $\vec{N_2}(t) = -\vec{N_1}(2\pi - t)$, ainsi que $\gamma_2(t) = \gamma_1(2\pi - t)$. Calculons l'intégrale de $\vec{F}$ à travers C pour la paramétrisation γ_1 qui est cohérente avec l'orientation de C :

$$\int_C \vec{F} \cdot \vec{n}\, ds = \int_{\gamma_1} \vec{F} \cdot \vec{n}\, ds = -\int_0^{2\pi} \vec{F}(\gamma_1(t)) \wedge \overrightarrow{\gamma_1'}(t) dt = \int_0^{2\pi} \left\langle \vec{F}(\gamma_1(t)) \cdot \vec{N_1}(t) \right\rangle dt$$

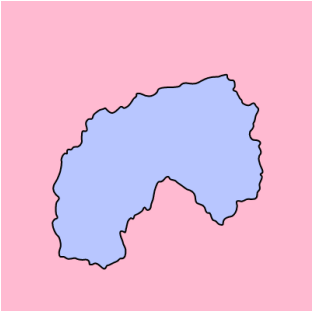
La norme du vecteur $\vec{N_1}(t)$ est celle du vecteur vitesse de γ_1 donc 1. La norme du champ de vecteurs $\vec{F}$ est $\sqrt{x^2+y^2}$ donc sur le cercle C, sa norme est 1. Le vecteur $\vec{F}(\gamma_1(t))$ est radial pointant vers l'extérieur, donc de sens opposé au vecteur $\vec{N_1}(t)$, on a donc $\int_C \vec{F} \cdot \vec{n} ds = -1$.

Calculons l'intégrale de $\vec{F}$ à travers γ_2 qui est incohérente avec l'orientation de C :

$$\begin{aligned}
\int_{\gamma_2} \vec{F} \cdot \vec{n}\, ds &= -\int_0^{2\pi} \vec{F}(\gamma_2(t)) \wedge \overrightarrow{\gamma_2'}(t) dt = \int_0^{2\pi} \left\langle \vec{F}(\gamma_2(t)) \cdot \vec{N_2}(t) \right\rangle dt \\
&= \int_{2\pi}^{0} \left\langle \vec{F}(\gamma_2(2\pi - u)) \cdot \vec{N_2}(2\pi - u) \right\rangle d(2\pi - u) \\
&= -\int_0^{2\pi} \left\langle \vec{F}(\gamma_1(u)) \cdot (-\vec{N_1}(u)) \right\rangle (-du) \\
&= -\int_C \vec{F} \cdot \vec{n}\, ds
\end{aligned}$$

On voit bien que le choix d'une paramétrisation incohérente avec l'orientation de C amène un signe $-$ devant l'intégrale. On retrouve le résultat de la Proposition 2.54.

Remarque 2.68. Le flocon de Koch est une courbe de Jordan, mais il n'existe pas de courbe paramétrée $\mathcal{C}^1$ par morceaux dont elle est le support car c'est une courbe fractale. On ne peut donc pas intégrer une fonction sur le flocon de Koch.



Theorème de Jordan. *Le complémentaire d'une courbe de Jordan C dans $\mathbb{R}^2$ est formé d'exactement deux composantes connexes distinctes, l'une bornée – appelée* domaine de Jordan *– et l'autre non. Toutes deux ont pour frontière la courbe de Jordan C.*

Si C est une courbe de Jordan, elle délimite un domaine du plan borné, dénoté D, dont elle est la frontière : c'est le domaine de Jordan associé à C. Faisons un choix d'orientation de C. On dit que la courbe C est *positivement orientée* si le domaine D dont C est la frontière se localise à la gauche lorsqu'on parcourt C selon l'orientation choisie. On dit qu'elle est *négativement orientée* sinon. Inversement, soit D un domaine connexe du plan, dont on suppose que la frontière est constituée d'une ou de plusieurs courbes de Jordan disjointes. Dans ce dernier cas le domaine a des "trous". L'orientation des courbes frontières se fait selon la même règle que précédemment : elles sont positivement orientées si D se trouve à la gauche en parcourant les courbes frontières.

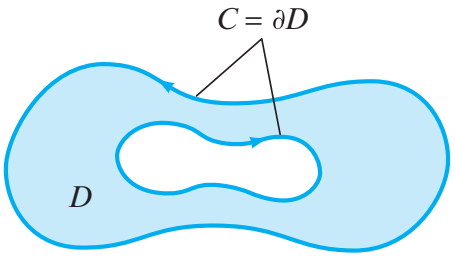


Figure 6.18 The shaded region D has a boundary consisting of two simple, closed curves, each of class C^1, whose union we call C.

Remarque 2.69. Si la frontière ∂D est faite de plusieurs courbes de Jordan disjointes $C_1, \ldots, C_n$ alors l'intégrale $\oint_{\partial D}$ est la somme des intégrales $\oint_{C_1} + \ldots + \oint_{C_n}$.

Theorème de Green. *Soit D un domaine borné du plan dont la frontière forme une ou plusieurs courbes de Jordan lisses par morceaux positivement orientées. Soit $\vec{F} : \mathbb{R}^2 \to \mathbb{R}^2$ un champ de vecteurs de classe $\mathcal{C}^1$, de composantes $F^1 : \mathbb{R}^2 \to \mathbb{R}$ et $F^2 : \mathbb{R}^2 \to \mathbb{R}$, respectivement. Alors nous avons (sous réserve que les intégrales existent) :*

$$\oint_{\partial D} \vec{F} \cdot d\vec{s} = \iint_D \left(\frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \right) dA \tag{2.7}$$

Démonstration. Voir pages 433-436 du livre de Susan Jane Colley, avec une note historique à la fin. □

On peut utiliser le Théorème de Green pour calculer des aires de domaines. Supposons que $\frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} = 1$, dans ce cas le membre de droite de l'Equation (2.7) mesure l'aire du domaine D. On peut la calculer à partir du membre de gauche si on trouve un champ de vecteurs continu $\vec{F}$ tel que $\frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} = 1$. Soit $u \in [0,1]$, et posons $F^1(x,y) = -uy$, $F^2(x,y) = (1-u)x$ satisfait la condition. Autrement dit, pour tout $u \in [0,1]$ on a (on rappelle la Proposition 2.49 et la Remarque 2.50 pour les notations) :

$$\text{Aire de } D \; = \iint_D dxdy = \oint_{\partial D} (-uy)dx + (1-u)xdy = \oint_{\partial D} xdy - u \oint_{\partial D} xdy + ydx$$

Exemple 2.70. Par exemple prenons la courbe paramétrée qui engendre une ellipse $\gamma : [0, 2\pi] \to \mathbb{R}^2, t \mapsto (a\cos(t), b\sin(t))$, où $a, b > 0$. On note D le domaine qu'elle contient et alors $\partial D = \mathrm{Im}(\gamma)$ est l'ellipse qui l'entoure et elle est positivement orientée. On prend le paramètre $u = \frac{1}{2}$, ce qui donne $F^1(x,y) = -\frac{y}{2}$ et $F^2(x,y) = \frac{x}{2}$ et donc :

$$\begin{aligned} \text{aire de l'ellipse } &= \frac{1}{2} \oint_{\partial D} xdy - ydx \\ &= \frac{1}{2} \oint_{\partial D} a\cos(t)d(b\sin(t)) - b\sin(t)d(a\cos(t)) \\ &= \frac{ab}{2} \int_0^{2\pi} \Big(\cos(t)^2 + \sin(t)^2\Big) dt = ab\pi \end{aligned}$$

C'est une manière élégante de calculer l'aire d'une ellipse de paramètres. On note que un cercle de rayon a est une ellipse pour laquelle $b = a$, et la formule $ab\pi$ devient l'aire du cercle $a^2\pi$.

Notons que rigoureusement on a (voir la Définition 2.65 et 2.47) :

$$\text{Aire de } D \ = \oint_{\partial D} \vec{F} \cdot d\vec{s} = \int_0^{2\pi} \left\langle \vec{F}(\gamma(t)), \overrightarrow{\gamma'}(t) \right\rangle dt = \int_0^{2\pi} \left\langle \tfrac{1}{2}\begin{pmatrix} -y(t) \\ x(t) \end{pmatrix}, \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \right\rangle dt$$

Avec cela on obtient le même résultat.

2.4 Surfaces paramétrées et théorème de Stokes

Dans la suite nous allons étendre toutes nos connaissances des courbes paramétrées aux surfaces paramétrées. Rappelons que nous pouvons voir le graphe d'une fonction réelle continue à une variable $f : I \to \mathbb{R}$ comme le support (l'image) d'une courbe paramétrée $\gamma_f : I \to \mathbb{R}^2, t \mapsto (t, f(t))$. Les graphes des fonctions réelles représentent donc des cas particuliers du support de courbes paramétrées. Et il existe des courbes paramétrées dont le support n'est pas le graphe d'une fonction réelle, comme par exemple $\gamma : \mathbb{R} \to \mathbb{R}, t \mapsto (0, t)$, qui est l'axe vertical des ordonnées. Nous allons suivre la même approche avec les fonctions continues à deux variables : leur graphe (dans $\mathbb{R}^3$) sont le support d'un certain type de fonctions à valeurs dans $\mathbb{R}^3$ qu'on appelle *surfaces paramétrées*. Bien sûr il existe des surfaces paramétrées dont le support n'est pas le graphe d'une fonction à deux variables. Le tableau suivant permet de suivre l'analogie :

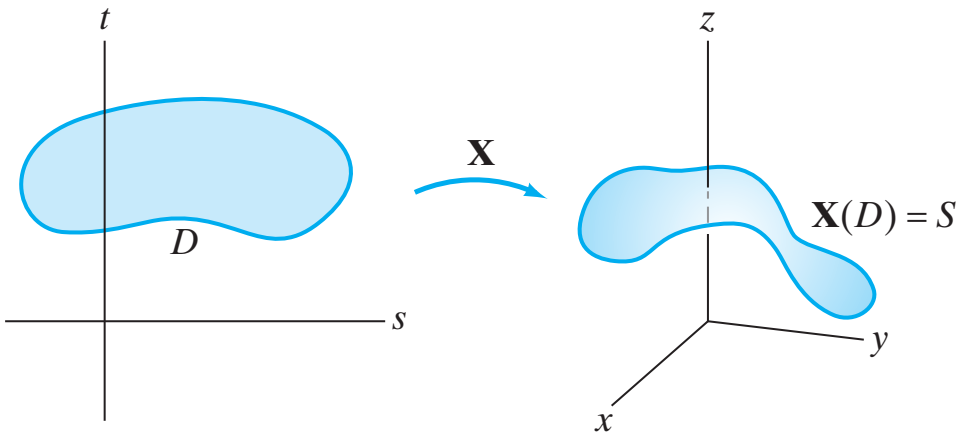
Application	$\gamma : I \to \mathbb{R}^2$	$\boldsymbol{X} : D \to \mathbb{R}^3$
Domaine de définition	intervalle ou segment de $\mathbb{R}$	domaine ouvert ou compact de $\mathbb{R}^2$
Support	courbe dans $\mathbb{R}^2$	surface dans $\mathbb{R}^3$
Cas particulier	graphe de fonctions réelles si $g : I \to \mathbb{R}$, $\gamma_g(t) = (t, g(t))$	graphe de fonctions à deux variables si $f : D \to \mathbb{R}$, $\boldsymbol{X}_f(s,t) = (s, t, f(s,t))$
Dérivabilité	droite tangente dirigée par $\overrightarrow{\gamma'}(t_0)$ au point $\gamma(t_0)$	plan tangent engendré par $\frac{\partial \boldsymbol{X}}{\partial s}(s_0,t_0)$ et $\frac{\partial \boldsymbol{X}}{\partial t}(s_0,t_0)$ au point $\boldsymbol{X}(s_0,t_0)$
Intégrabilité scalaire	$\int_\gamma f ds = \int_I f(\gamma(t)) \left\| \overrightarrow{\gamma'}(t) \right\| dt$ avec $f : \mathbb{R}^2 \to \mathbb{R}$	$\iint_D f(\boldsymbol{X}(s,t)) \left\| \frac{\partial \boldsymbol{X}}{\partial s}(s,t) \wedge \frac{\partial \boldsymbol{X}}{\partial t}(s,t) \right\| dsdt$ avec $f : \mathbb{R}^3 \to \mathbb{R}$
Intégrabilité vectorielle	$\int_\gamma \vec{F} \cdot \vec{n} ds$ avec $\vec{F} : \mathbb{R}^2 \to \mathbb{R}^2$	$\iint_D \left\langle \vec{F}(\boldsymbol{X}(s,t)), \frac{\partial \boldsymbol{X}}{\partial s}(s,t) \wedge \frac{\partial \boldsymbol{X}}{\partial t}(s,t) \right\rangle dsdt$ avec $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$

Remarque 2.71. On ne peut généraliser qu'un seul type d'intégral vectoriel aux surfaces paramétrées : celles de flux. En effet en un point d'une surface paramétrée il existe une infinité de vecteurs tangents (donc on ne peut pas faire une intégrale de circulation) mais un seul vecteur normal (donc on peut faire une intégrale de flux).

Définition 2.72. *Soit D un domaine connexe (par arcs) de $\mathbb{R}^2$. On appelle* surface paramétrée *toute application continue $\mathbf{X} : D \to \mathbb{R}^3$ qui est injective sur l'intérieur du domaine $\mathring{D}$. En particulier, la fonction $\mathbf{X} : D \to \mathbb{R}^3$ a trois* fonctions composantes *$x : D \to \mathbb{R}$, $y : D \to \mathbb{R}$ et $z : D \to \mathbb{R}$ qui satisfont :*

$$\mathbf{X}(s,t) = (x(s,t), y(s,t), z(s,t))$$

On appelle l'image de $\mathbf{X}$ le support *de la surface paramétrée par $\mathbf{X}$, et on la dénote habituellement par S ou $S_{\mathbf{X}}$. Ce sont des sous-ensembles de $\mathbb{R}^3$.*



Remarque 2.73. On suppose que la fonction $\boldsymbol{X}$ est injective sur l'intérieur du domaine $\mathring{D}$, car on veut éviter des situations de surfaces avec auto-intersections, qui sont plutôt difficiles à visualiser et à étudier.

Exemple 2.74. Un cas particulier arrive quand on a une fonction continue à deux variables $f : D \to \mathbb{R}$. Dans ce cas on pose $\boldsymbol{X}(s,t) = (s,t,f(x,t))$. Le support de $\boldsymbol{X}$ est le graphe de la fonction f dans $\mathbb{R}^3$.

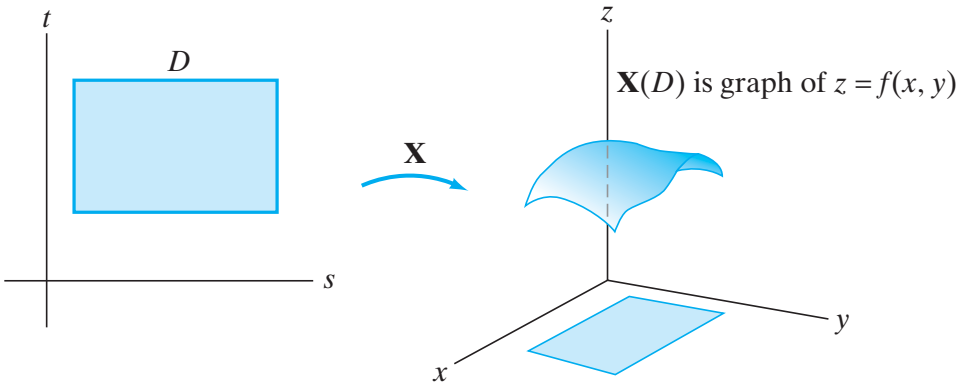


Figure 7.5 The graph of $z = f(x, y)$ as a parametrized surface.

Exemple 2.75. La sphère de Riemann est un exemple non trivial de surface paramétrée qui n'est pas le graphe d'une fonction à deux variables. Le domaine de définition est le plan $D = \mathbb{R}^2$ et on définit la surface paramétrée $\boldsymbol{X} : \mathbb{R}^2 \to \mathbb{R}^3$ par :

$$\forall\, s,t \in \mathbb{R}^2 \qquad \boldsymbol{X}(s,t) = \left(\frac{2s}{1+s^2+t^2}, \frac{2t}{1+s^2+t^2}, 1 - \frac{2}{1+s^2+t^2}\right)$$

Le support (l'image) de cette surface paramétrée est une sphère de centre l'origine, et dont on a enlevé le pôle nord, c'est à dire le point $(0,0,1)$. Cette fonction montre que le plan $\mathbb{R}^2$ est isomorphe à la sphère moins un point. Cela a des conséquences importante en géométrie complexe et algébrique. En particulier on voit qu'on peut compactifier le plan $\mathbb{R}^2$ en lui ajoutant un point, qu'on note habituellement $\{\infty\}$. Ce point symbolise le fait d'être à l'infini de l'origine. Comme $\boldsymbol{X}$ est injective, elle est inversible, et sa fonction inverse est la *projection stéréographique* de la sphère sur le plan (voir Wikipedia pour comprendre sa définition).

Les fonctions composantes de la surface paramétrée $\boldsymbol{X} : D \to \mathbb{R}^3$ sont des fonctions à deux variables $x, y, z : D \to \mathbb{R}$. Si $\boldsymbol{X} : D \to \mathbb{R}^3$ est différentiable en un point (a,b), les fonctions composantes peuvent être dérivées par rapport à la première et la seconde variable (et même

par rapport à n'importe quel vecteur $\vec{v}$ comme vu dans la Définition 1.23). Si on dérive les fonctions coordonnées dans les directions principales $\vec{i}$ (direction de l'axe des abscisses s) et $\vec{j}$ (direction de l'axe des ordonnées t), cela définit deux vecteurs de $\mathbb{R}^3$ qui sont tangents au support de $\boldsymbol{X}$ au point $\boldsymbol{X}(a,b)$:

$$\vec{T}_s(a,b) = \frac{\partial \boldsymbol{X}}{\partial s}(a,b) = \begin{pmatrix} \frac{\partial x}{\partial s}(a,b) \\ \frac{\partial y}{\partial s}(a,b) \\ \frac{\partial z}{\partial s}(a,b) \end{pmatrix} \qquad \text{et} \qquad \vec{T}_t(a,b) = \frac{\partial \boldsymbol{X}}{\partial t}(a,b) = \begin{pmatrix} \frac{\partial x}{\partial t}(a,b) \\ \frac{\partial y}{\partial t}(a,b) \\ \frac{\partial z}{\partial t}(a,b) \end{pmatrix}$$

Nous allons expliquer pourquoi ils sont tangents au support de $\boldsymbol{X}$, en définissant le plan tangent à $S_{\boldsymbol{X}}$ au point (a,b).

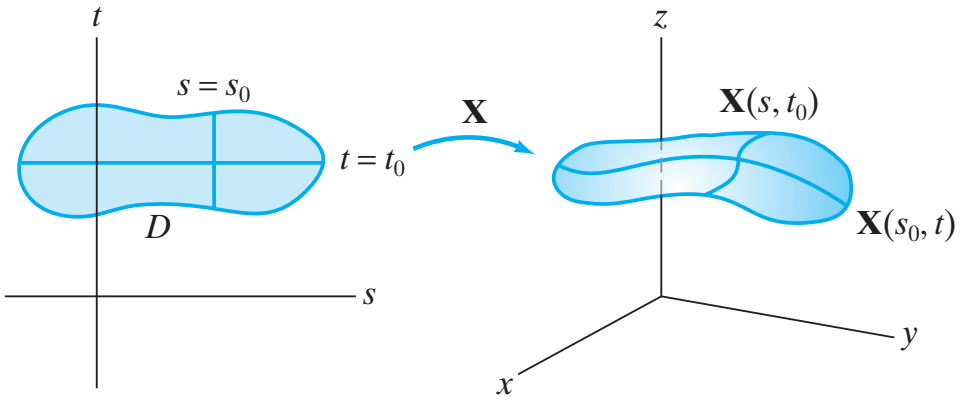


Figure 7.6 The coordinate curves of a parametrized surface.

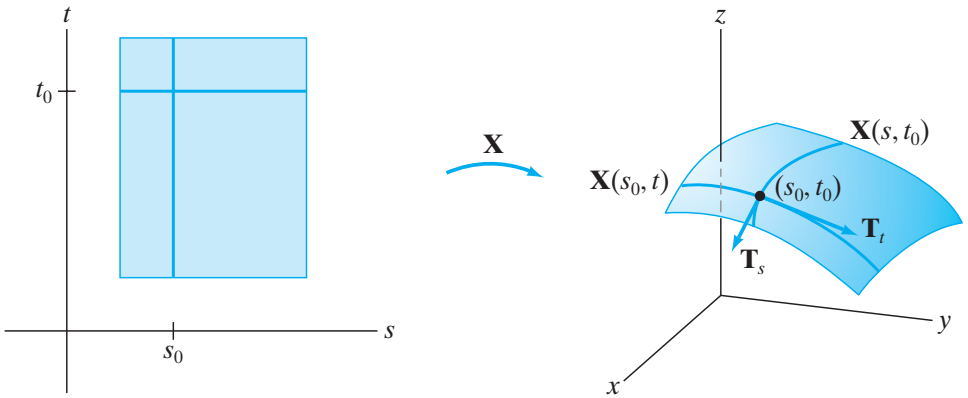


Figure 7.10 The tangent vectors $\mathbf{T}_s$ and $\mathbf{T}_t$ to the coordinate curves.

Exemple 2.76. Prenons la sphère de Riemann de l'Exemple 2.75. Les vecteurs $\vec{T}_s$ et $\vec{T}_t$ sont donnés par :

$$\vec{T}_s(s,t) = \frac{2}{(1+s^2+t^2)^2}\begin{pmatrix} 1-s^2+t^2 \\ -2st \\ 2s \end{pmatrix} \qquad \text{et} \qquad \vec{T}_t(s,t) = \frac{2}{(1+s^2+t^2)^2}\begin{pmatrix} -2st \\ 1+s^2-t^2 \\ 2t \end{pmatrix}$$

Un champ de vecteurs normal à la sphère est un champ de vecteurs radial, c'est à dire du type $\vec{n}(x,y,z) = x\vec{i} + y\vec{j} + z\vec{k}$. Au point $\boldsymbol{X}(s,t)$, on obtient :

$$\vec{n}(\boldsymbol{X}(s,t)) = \frac{1}{1+s^2+t^2}\begin{pmatrix} 2s \\ 2t \\ -1+s^2+t^2 \end{pmatrix}$$

Calculons le produit scalaire $\left\langle \vec{T}_s(s,t), \vec{n}(\boldsymbol{X}(s,t)) \right\rangle$ à un point (a,b). On obtient :

$$\begin{aligned}\left\langle \vec{T}_s(a,b), \vec{n}(\boldsymbol{X}(a,b)) \right\rangle &= \frac{2}{(1+a^2+b^2)^3} \left\langle \begin{pmatrix} 1-a^2+b^2 \\ -2ab \\ 2a \end{pmatrix}, \begin{pmatrix} 2a \\ 2b \\ -1+a^2+b^2 \end{pmatrix} \right\rangle \\ &= \frac{2}{(1+a^2+b^2)^3} \underbrace{\Big((1-a^2+b^2)\times 2a - 2ab \times 2b + 2a \times (-1+a^2+b^2)\Big)}_{=0}\end{aligned}$$

On voit donc que le vecteur $\vec{T}_s(a,b)$ est normal au vecteur radial $\vec{n}(\boldsymbol{X}(a,b))$, et donc il est tangent à la sphère. De même pour $\vec{T}_t(a,b)$.

Cet exemple inspire d'ailleurs une définition plus générale. En effet, du fait que $\vec{T}_s(a,b)$ et $\vec{T}_t(a,b)$ sont tous les deux tangents au support de $\boldsymbol{X}$, leur produit vectoriel est un vecteur normal aux deux (voir Définition 1.110), dès lors qu'il n'est pas nul. Rappelons que le produit vectoriel des deux vecteurs tridimensionnels $\vec{T}_s(a,b)$ et $\vec{T}_t(a,b)$ est donné par :

$$\vec{T}_s(a,b) \wedge \vec{T}_t(a,b) = \begin{pmatrix} \frac{\partial y}{\partial s}(a,b)\frac{\partial z}{\partial t}(a,b) - \frac{\partial z}{\partial s}(a,b)\frac{\partial y}{\partial t}(a,b) \\ -\frac{\partial x}{\partial s}(a,b)\frac{\partial z}{\partial t}(a,b) + \frac{\partial z}{\partial s}(a,b)\frac{\partial x}{\partial t}(a,b) \\ \frac{\partial x}{\partial s}(a,b)\frac{\partial y}{\partial t}(a,b) - \frac{\partial y}{\partial s}(a,b)\frac{\partial x}{\partial t}(a,b) \end{pmatrix} \tag{2.8}$$

Exemple 2.77. Prenons la suite de l'Exemple 2.76 et calculons le produit vectoriel des deux vecteurs tangents. Si on fait les calculs, on obtient la formule suivante :

$$\vec{T}_s(a,b) \wedge \vec{T}_t(a,b) = -\frac{4}{(1+a^2+b^2)^3} \begin{pmatrix} 2a \\ 2b \\ -1+a^2+b^2 \end{pmatrix} = -\frac{4}{(1+a^2+b^2)^2} \vec{n}(\boldsymbol{X}(a,b))$$

Le produit vectoriel des deux vecteurs tangent à la sphère est donc bien normal comme attendu (et pointe vers l'intérieur).

Définition 2.78. *La surface paramétrée* $\mathbf{X} : D \to \mathbb{R}^3$ *est dite* lisse au point $(a,b) \in \mathring{D}$ *si :*

- *la fonction* $\mathbf{X}$ *est de classe* $\mathcal{C}^1$ *dans un voisinage du point* (a,b) *et*
- *le vecteur de* $\mathbb{R}^3$ *suivant est non nul :*

$$\vec{N}(a,b) = \vec{T}_s(a,b) \wedge \vec{T}_t(a,b)$$

Si $\mathbf{X}$ *est lisse en tout point de* $\mathring{D}$ *alors on dit que la surface paramétrée* $\mathbf{X}$ *est* lisse*, et on appelle* vecteur normal standard *associé à la paramétrisation* $\mathbf{X}$ *l'application vectorielle continue suivante :*

$$\begin{aligned} \vec{N} : \ \mathring{D} &\longrightarrow \mathbb{R}^3 \\ (s,t) &\longmapsto \vec{N}(s,t) = \vec{T}_s(s,t) \wedge \vec{T}_t(s,t) \end{aligned}$$

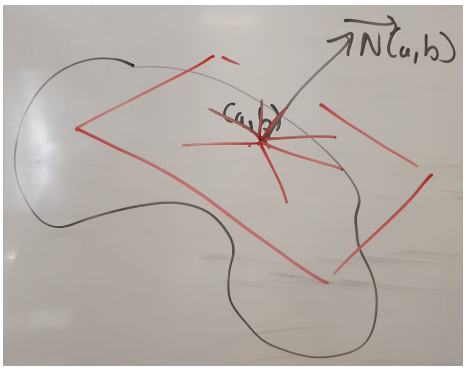
Remarque 2.79. Etre lisse veut dire quelque chose de plus qu'être $\mathcal{C}^1$: on demande une condition sur la colinéarité des deux vecteurs tangents canoniques $\vec{T}_s$ et $\vec{T}_t$. Attention, il n'y a aucune raison pour laquelle les vecteurs $\vec{T}_s(a,b), \vec{T}_t(a,b)$ et $\vec{N}(a,b)$ soient unitaires.

D'autre part, comme on a un vecteur normal à la surface, on peut définir le plan tangent au support de $\boldsymbol{X}$. Et même en déduire l'équation de son plan tangent, comme on l'a vu dans la discussion qui conduit à la Définition 1.106.

Définition 2.80. *Soit* $\mathbf{X} : D \to \mathbb{R}^3$ *une surface paramétrée lisse au point* $(a,b) \in \mathring{D}$. *Le plan de l'espace passant au point* $\mathbf{X}(a,b) = (x(a,b), y(a,b), z(a,b))$ *et orthogonal au vecteur* $\vec{N}(a,b)$ *est appelé* plan tangent *au support de* $\mathbf{X}$ *au point* $\mathbf{X}(a,b)$, *et son équation est donnée par :*

$$(x - x(a,b))N^1(a,b) + (y - y(a,b))N^2(a,b) + (z - z(a,b))N^3(a,b) = 0$$

où N^1, N^2, N^3 *sont les composantes du vecteur normal* $\vec{N}$.



Remarque 2.81. Il est équivalent de dire que le plan tangent est normal au vecteur $\vec{N}(a,b)$, ou qu'il est engendré par $\vec{T}_s(a,b)$ et $\vec{T}_t(a,b)$.

Exemple 2.82. Calculons l'équation du plan tangent à la sphère de Riemann de l'Exemple 2.75 en un point $\boldsymbol{X}(a,b)$. Le vecteur normal est donné dans l'Exemple 2.77. On obtient l'équation suivante :

$$2ax + 2by + (-1 + a^2 + b^2)z - (1 + a^2 + b^2) = 0$$

En application, on peut prendre le pôle sud, c'est à dire $a = 0$ et $b = 0$, c'est à dire $\boldsymbol{X}(0,0) = (0,0,-1)$. Le plan tangent a pour équation $z = -1$, c'est bien le plan horizontal d'altitude $z = -1$.

Nous souhaitons maintenant calculer l'aire d'une surface paramétrée $\boldsymbol{X} : D \to \mathbb{R}^3$. Pour cela, il faut une double intégrale sur le domaine de définition D, mais qui dépend évidemment des fonctions composantes de $\boldsymbol{X}$. La fonction $\boldsymbol{X}$ déforme le domaine de définition pour l'emmener sur le support $\mathrm{Im}(\boldsymbol{X})$. Il y a donc un facteur de déformation/distorsion. Nous avons déjà vu ce type de cas, c'était dans la discussion qui mène jusqu'à la Proposition 2.28. Malheureusement, la fonction $\boldsymbol{X} : D \to \mathbb{R}^3$ envoie un domaine de dimension 2 dans un espace de dimension 3 (même si son image est une surface, donc de dimension 2). Sa matrice Jacobienne est donc une matrice 3×2. On ne peut pas calculer le déterminant d'une matrice 3×2 donc on ne peut pas reproduire l'Equation 2.29. Par contre, on peut réinterpréter le déterminant de l'Equation 2.29 comme suit :

$$\det\left(J_{(u,v)}(T)\right) = \det\begin{pmatrix} \frac{\partial T^1}{\partial u}(u,v) & \frac{\partial T^1}{\partial v}(u,v) \\ \frac{\partial T^2}{\partial u}(u,v) & \frac{\partial T^2}{\partial v}(u,v) \end{pmatrix} = \det\begin{pmatrix} \frac{\partial T}{\partial u}(u,v) & \frac{\partial T}{\partial v}(u,v) \end{pmatrix}$$

Or avec l'Equation (2.6) et la discussion autour, nous pouvons exprimer ce déterminant comme un produit vectoriel des deux vecteurs bidimensionnels, et nous avons donc l'égalité suivante :

$$\det\left(J_{(u,v)}(T)\right) = \frac{\partial T}{\partial u}(u,v) \wedge \frac{\partial T}{\partial v}(u,v)$$

Ceci nous permet de récrire l'Equation 2.29 comme :

$$\text{Aire du domaine } D \;= \iint_D dxdy = \iint_{D^*} \left|\frac{\partial T}{\partial u}(u,v) \wedge \frac{\partial T}{\partial v}(u,v)\right| dudv \tag{2.9}$$

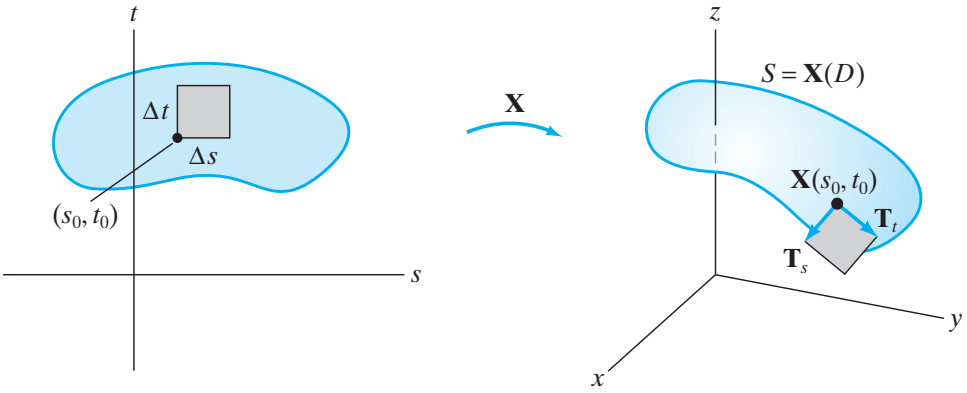
Nous allons donc appliquer cette formule au cas d'une surface paramétrée dans $\mathbb{R}^3$, en reproduisant le Corollaire 2.29 pour le calcul de l'aire d'une surface. En 2 dimensions nous avions une application $T : D^* \to D$ dont l'image est un domaine bidimensionnel de $\mathbb{R}^2$, et maintenant on a une application $\boldsymbol{X} : D \to S_{\boldsymbol{X}}$ dont l'image est une surface bidimensionnelle de $\mathbb{R}^3$. On ne peut pas écrire le déterminant de la matrice Jacobienne de l'application $\boldsymbol{X}$ car sa matrice Jacobienne n'est pas carrée, mais on peut écrire un produit vectoriel des deux vecteurs $\vec{T}_s = \frac{\partial \boldsymbol{X}}{\partial s}(s,t)$ et $\vec{T}_t = \frac{\partial \boldsymbol{X}}{\partial t}(s,t)$ comme on l'a vu dans l'Equation (2.9). Cela remplace le déterminant de la Jacobienne. Plus précisément on a le dictionnaire suivant :

Application	$T : D^* \to \mathbb{R}^2$	$\boldsymbol{X} : D \to \mathbb{R}^3$
Domaine de définition	D^*	D
Image	$D \subset \mathbb{R}^2$	$S_{\boldsymbol{X}} \subset \mathbb{R}^3$
Facteur multiplicateur	$\left\lvert \frac{\partial T}{\partial u}(u,v) \wedge \frac{\partial T}{\partial v}(u,v) \right\rvert$	$\left\lVert \vec{T}_s(s,t) \wedge \vec{T}_t(s,t) \right\rVert$

Grâce à ce tableau, nous pouvons définir l'aire du support d'une surface paramétrée en nous appuyant sur l'Equation (2.9) ce qui généralise le Corollaire 2.29 :

Définition 2.83. *Soit* $\mathbf{X} : D \to \mathbb{R}^3$ *une surface paramétrée de classe* $\mathcal{C}^1$*. L'aire du support de la surface paramétrée* $\mathbf{X}$ *est donnée par :*

$$\text{Aire de } S_{\mathbf{X}} = \iint_D \left\lVert \vec{T}_s(s,t) \wedge \vec{T}_t(s,t) \right\rVert ds dt \tag{2.10}$$



Exemple 2.84. Reprenons l'Exemple 2.74 où on suppose la fonction $f : D \to \mathbb{R}$ de classe $\mathcal{C}^1$. Nous avons que $\vec{T}_s = \left(1, 0, \frac{\partial f}{\partial s}(s,t)\right)$ et $\vec{T}_t = \left(0, 1, \frac{\partial f}{\partial t}(s,t)\right)$. On obtient donc (on a déjà vu le calcul dans la Proposition 1.111) :

$$\vec{T}_s(s,t) \wedge \vec{T}_t(s,t) = \begin{pmatrix} -\nabla f(s,t) \\ 1 \end{pmatrix}$$

Ceci donne donc :

$$\text{Aire de } S_{\boldsymbol{X}_f} = \iint_D \sqrt{\left(\frac{\partial f}{\partial s}(s,t)\right)^2 + \left(\frac{\partial f}{\partial t}(s,t)\right)^2 + 1}\, ds dt$$

Exemple 2.85. Reprenons la sphère de Riemann de l'Exemple 2.75, où on a donné le vecteur normal dans l'Exemple 2.77, et calculons son aire. On devrait trouver 4π car son rayon vaut 1. La norme du vecteur normal est :

$$\left\lVert \vec{N}(s,t) \right\rVert = \frac{4}{(1+s^2+t^2)^3} \sqrt{4s^2 + 4t^2 + (-1+s^2+t^2)^2} = \frac{4}{(1+s^2+t^2)^2}$$

On obtient donc, en passant en polaire, en écrivant $s = r\cos(\theta)$ et $t = r\sin(\theta)$:

$$\begin{aligned}
\text{Aire de la sphère de rayon 1} &= \iint_{\mathbb{R}^2} \left\|\vec{N}(s,t)\right\| dsdt \\
&= \iint_{\mathbb{R}^2} \frac{4}{(1+s^2+t^2)^2} dsdt \\
&= 4\int_0^{2\pi}\int_0^{+\infty} \frac{1}{(1+r^2\cos^2(\theta)+r^2\sin^2(\theta))^2} r dr d\theta \\
&= 4\left(\int_0^{2\pi} d\theta\right)\cdot\left(\int_0^{+\infty} \frac{rdr}{(1+r^2)^2}\right) \\
&= 4\times 2\pi\times\left[-\frac{1}{2(1+r^2)}\right]_0^{+\infty} = 4\pi
\end{aligned}$$

On retrouve bien le résultat que l'aire de la sphère de rayon 1 est 4π.

L'Equation (2.10) est une généralisation de l'Equation (2.9), donc on devrait retrouver le Corollaire 2.29 à partir de la Définition 2.83 si on a une application $\boldsymbol{X} : D^*\mathbb{R}^3$ dont l'image est un domaine D du plan $\mathbb{R}^2$ horizontal d'équation $z = 0$. Dans ce cas, $\boldsymbol{X}$ est d'altitude constante et dans ce cas, on a :

$$\vec{T}_s(a,b) = \begin{pmatrix} \frac{\partial x}{\partial s}(a,b) \\ \frac{\partial y}{\partial s}(a,b) \\ 0 \end{pmatrix} \qquad \text{et} \qquad \vec{T}_t(a,b) = \begin{pmatrix} \frac{\partial x}{\partial t}(a,b) \\ \frac{\partial y}{\partial t}(a,b) \\ 0 \end{pmatrix}$$

l'Equation (2.10) appliquée à ces deux vecteurs tangents donne, en utilisant l'Equation (2.8) :

$$\left\|\vec{T}_s(s,t)\wedge\vec{T}_t(s,t)\right\| = \left|\frac{\partial x}{\partial s}(a,b)\times\frac{\partial y}{\partial t}(a,b) - \frac{\partial y}{\partial s}(a,b)\times\frac{\partial y}{\partial t}(a,b)\right| = \left|\begin{pmatrix}\frac{\partial x}{\partial s}(a,b)\\ \frac{\partial y}{\partial s}(a,b)\end{pmatrix}\wedge\begin{pmatrix}\frac{\partial x}{\partial s}(a,b)\\ \frac{\partial y}{\partial s}(a,b)\end{pmatrix}\right|$$

On voit bien que le facteur multiplicateur de l'intégrale de l'Equation (2.10) dans ce cas particulier redonne le facteur multiplicateur de l'Equation (2.9). C'est important en mathématiques, lorsqu'on généralise une idée, de vérifier que, si on applique cette idée au cas particulier qui a donné l'inspiration, on retrouve bien le résultat qui a donné l'inspiration.

Nous pouvons dès lors généraliser la Définition 2.45 et la Définition 2.52 en remplaçant la courbe paramétrée $\gamma : I \to \mathbb{R}^2$ par une surface paramétrée $\boldsymbol{X} : D \to \mathbb{R}^3$, en suivant la conversion suivante (où on suit la généralisation de la Définition 2.83 pour les intégrales scalaires) :

Intégrale scalaire	Définition 2.45 $\int_I f(\gamma(t))\left\|\overrightarrow{\gamma'}(t)\right\| dt$ avec $f : \mathbb{R}^2 \to \mathbb{R}$	Définition 2.86 $\iint_D f(\boldsymbol{X}(s,t))\left\|\frac{\partial \boldsymbol{X}}{\partial s}(s,t)\wedge\frac{\partial \boldsymbol{X}}{\partial t}(s,t)\right\| dsdt$ avec $f : \mathbb{R}^3 \to \mathbb{R}$
Intégrale vectorielle	Définition 2.52 $\int_\gamma \vec{F}\cdot\vec{n}ds$ avec $\vec{F} : \mathbb{R}^2 \to \mathbb{R}^2$	Définition 2.87 $\iint_D \left\langle \vec{F}(\boldsymbol{X}(s,t)), \frac{\partial \boldsymbol{X}}{\partial s}(s,t)\wedge\frac{\partial \boldsymbol{X}}{\partial t}(s,t)\right\rangle dsdt$ avec $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$

Alors pourquoi généralise-t-on la Définition 2.52 et pas la Définition 2.47 ? Car on ne peut pas généraliser la Définition 2.47. En effet, dans cette définition le produit scalaire en le champ de vecteurs se prend avec le vecteur tangent $\overrightarrow{\gamma'}$. Cependant avec une surface paramétrée, il n'y a pas un unique vecteur tangent en un point de la surface, mais une infinité (c'est le plan tangent). La seule intégrale vectorielle que l'on peut faire en 3 dimensions, c'est donc celle qui prend le produit scalaire entre un champ de vecteur $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ et le vecteur normal canonique $\vec{N} = \vec{T}_s \wedge \vec{T}_t$. Plus précisément nous avons :

Définition 2.86. *Soit $f : \mathbb{R}^3 \to \mathbb{R}$ une fonction continue et soit $\mathbf{X} : D \to \mathbb{R}^3$ une surface paramétrée lisse. On définit l'*intégrale surfacique scalaire de f sur la surface $\boldsymbol{X}$ *l'intégrale surD (si elle existe) de la fonction continue $(s,t) \mapsto f(\mathbf{X}(s,t))\left\|\vec{N}(s,t)\right\|$, et on note :*

$$\iint_{\mathbf{X}} f dS = \iint_D f(\mathbf{X}(s,t))\left\|\vec{N}(s,t)\right\| dsdt$$

Définition 2.87. *Soit $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs continu, et soit $\mathbf{X} : D \to \mathbb{R}^3$ une surface paramétrée lisse. On définit l'*intégrale surfacique vectorielle de $\vec{F}$ à travers la surface S *– ou* flux de $\vec{F}$ à travers la surface S *– l'intégrale sur D (si elle existe) de la fonction continue $(s,t) \mapsto \left\langle \vec{F}(\mathbf{X}(s,t)), \vec{N}(s,t) \right\rangle$, et on note :*

$$\iint_{\mathbf{X}} \vec{F} \cdot d\vec{S} = \iint_D \left\langle \vec{F}(\mathbf{X}(s,t)), \vec{N}(s,t) \right\rangle dsdt$$

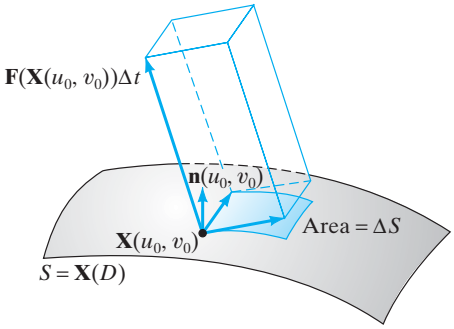


Figure 7.19 The amount of fluid transported across a small piece of S during a brief time interval Δt may be approximated by the volume of a parallelepiped.

Remarque 2.88. Si on voit $\vec{F}$ comme le champ de vitesse d'un fluide, alors le flux de $\vec{F}$ à travers la surface S correspond à son débit (voir p476 du livre de Susan Colley).

Remarque 2.89. Comme pour les courbes paramétrées, on peut définir la notion de reparamétrisation d'une surface paramétrée. Etre lisse est une propriété invariante sous reparamétrisation. Comme pour les Propositions 2.53 et 2.54, on peut montrer que les intégrales surfaciques scalaires sont invariantes sous reparamétrisation, tandis que les intégrales surfaciques vectorielles peuvent ou non changer de signe. Les détails se trouvent pages 477-479 du livre de Susan Jane Colley.

Définition 2.90. *Une* surface de l'espace *est un sous-ensemble S de $\mathbb{R}^3$ qui est le support (c'est à dire l'image) d'une surface paramétrée $\mathbf{X} : D \to \mathbb{R}$, i.e. $S = S_{\mathbf{X}}$. Elle est dite* lisse *si elle est le support d'une surface paramétrée lisse.*

Une surface de l'espace lisse S est dite *orientable* si, en tout point de $S \backslash \partial S$ (S privé de sa frontière), on peut faire un unique choix de vecteur unitaire normal à S (i.e. au plan tangent à S en ce point) qui varie continument sur la surface $S \backslash \partial S$. Plus précisément, si on peut définir une application $\vec{n} : S \to \mathbb{R}^3$ telle que pour une (toute) paramétrisation lisse $\boldsymbol{X} : D \to \mathbb{R}^3$ de S – i.e. telle que $S = S_{\boldsymbol{X}}$, la composition $\vec{n} \circ \boldsymbol{X} : D \to \mathbb{R}^3$ est une fonction continue dont l'image $\boldsymbol{X}(\vec{s}, t)$ en un point (s,t) est un vecteur unitaire normal au plan tangent à S au point $\boldsymbol{X}(s,t)$. On dit que S est *non-orientable* sinon. Une *orientation* d'une surface de l'espace lisse orientable S est le choix d'un tel champ de vecteur $\vec{n} : S \to \mathbb{R}^3$. Une surface de l'espace lisse orientable munie d'une orientation est dite *orientée*.

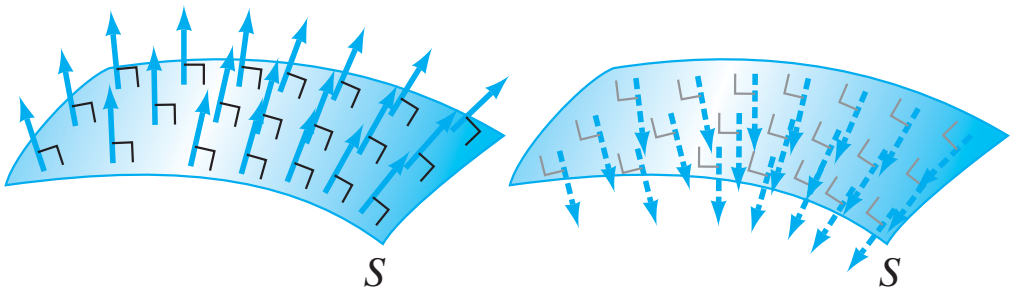


Figure 7.21 The connected orientable surface S shown with its two possible orientations.

Remarque 2.91. D'après la Remarque 2.89, être lisse ne dépend pas du choix de paramétrisation $\boldsymbol{X} : D \to \mathbb{R}^3$ lisse de S. De même le fait d'être orientable est indépendant du choix de paramétrisation.

Exemple 2.92. Il existe des surfaces non-orientables : c'est le cas du ruban de Möbius défini par :

$$\forall\, s \in [0, 2\pi],\ \forall\, t \in \left[-\tfrac{1}{2}, \tfrac{1}{2}\right] \qquad \boldsymbol{X}(s,t) = \begin{cases} x(s,t) = (1 + t\cos\left(\frac{s}{2}\right))\cos(s) \\ y(s,t) = (1 + t\cos\left(\frac{s}{2}\right))\sin(s) \\ z(s,t) = t\sin\left(\frac{s}{2}\right) \end{cases}$$

On a la propriété que $\boldsymbol{X}(0,t) = \boldsymbol{X}(2\pi, -t)$, et donc que $\vec{N}(0,t) = -\vec{N}(2\pi, -t)$. Ainsi, même si le vecteur normal standard varie de manière continue sur la surface, il n'est pas défini de manière unique ! On dit que cette surface n'a qu'un seul côté (elle est non-orientable).

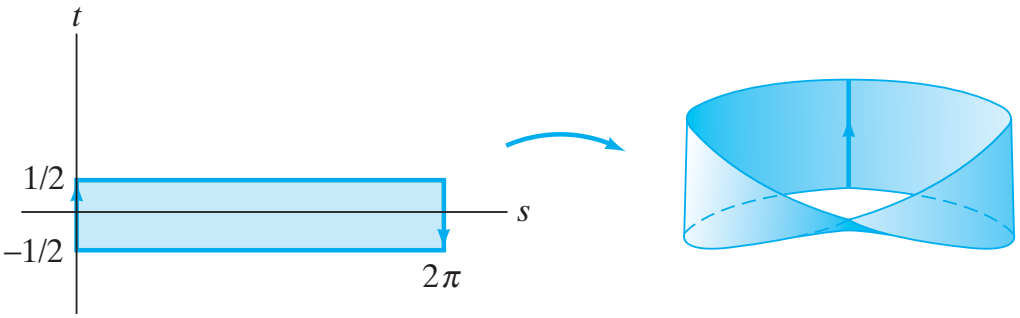
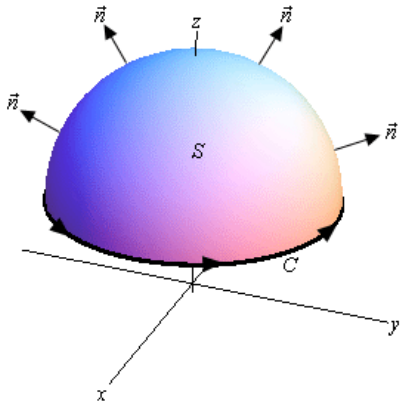


Figure 7.24 Gluing the ends of a strip of paper so that the arrows align provides a model of the Möbius strip.

Bien sûr, une surface paramétrée lisse $\boldsymbol{X} : D \to \mathbb{R}^3$ induit une orientation naturelle sur son support, celle induite par le vecteur normal standard associé à la paramétrisation $\boldsymbol{X}$ tel que défini dans la Définition 2.78. Soit maintenant S une surface de l'espace lisse orientée, et $\boldsymbol{X} : D \to \mathbb{R}^3$ une paramétrisation de S, i.e. telle que $S = S_{\boldsymbol{X}}$. Si le champ de vecteurs $\vec{n} : S \to \mathbb{R}^3$ qui définit l'orientation de S est *positivement colinéaire* avec le vecteur normal standard associé à la paramétrisation $\boldsymbol{X}$ tel que défini dans la Définition 2.78, on dit que l'orientation de $S_{\boldsymbol{X}}$ (ou de $\boldsymbol{X}$) est cohérente avec celle de S. Si le champ de vecteurs $\vec{n} : S \to \mathbb{R}^3$ qui définit l'orientation de S est *négativement colinéaire* avec le vecteur normal standard associé à la paramétrisation $\boldsymbol{X}$ tel que défini dans la Définition 2.78, on dit que l'orientation de $S_{\boldsymbol{X}}$ (ou de $\boldsymbol{X}$) est incohérente avec celle de S. Comme les surfaces sont suffisamment régulières, il ne peut y avoir que deux choix.

Une surface de l'espace qui n'a pas de frontière est dite *fermée*. Une surface de l'espace qui n'est pas fermée a donc une frontière, qui est une courbe de l'espace. La question repose maintenant de choisir une orientation de la surface et une orientation de la courbe frontière qui soient *cohérentes* entre elles. Qu'est-ce que cela veut dire ? Soit S une surface de l'espace orientable et soit C sa frontière, une courbe de l'espace qu'on suppose orientable (c'est à dire qu'on peut choisir un sens de parcours). Choisissons une orientation de S, c'est à dire la donnée d'un unique vecteur normal en tout point qui varie de manière continue, c'est à dire une application continue $\vec{n} : S \to \mathbb{R}^3$ qui est normal à S en tout point de S (sauf peut être à la frontière), comme défini dans la discussion sous la Définition 2.90. Pour choisir l'orientation de la courbe C, on va suivre la *règle de la main droite* ou du *trièdre direct* :

- imaginer que $\vec{n}$ suit le pouce de la main droite, et imaginer qu'on s'approche de la frontière de S avec cette orientation ;
- mettre l'index parallèle à la courbe C, et il y a deux façons possibles : soit avec le doigt majeur "inward" qui pointe vers l'intérieur de la surface S, soit avec le majeur "outward", qui pointe vers l'extérieur de la surface S ;
- l'orientation *cohérente* de la courbe C vis à vis de l'orientation de la surface S est de faire le choix du majeur "inward".



Dans ce cas on dit que la frontière C est orientée de manière cohérente avec l'orientation de S. Avec cette règle de la main droite, nous pouvons même orienter des surfaces lisses par morceaux, et orienter leurs frontière de manière cohérente.

Définition 2.93. *Une* surface paramétrée lisse par morceaux *est une famille finie de surfaces paramétrées lisses* $\mathbf{X}_i : D_i \to \mathbb{R}^3$. *Son support est l'union des supports de chaque* $\mathbf{X}_i$. *Une surface de l'espace lisse par morceaux est le support d'une surface paramétrée lisse par morceaux.*

Le problème lorsqu'on a une surface de l'espace lisse par morceaux, c'est que se passe-t-il à la frontière entre deux parties lisses ? comment choisir une orientation ? Il est en réalité possible d'orienter la surface dans son ensemble. Voici comment : supposons que S_1 et S_2 soient deux surfaces de l'espace lisses qui se rejoignent le long d'une courbe de l'espace commune C. Soient $\vec{n}_1$ et $\vec{n}_2$ les vecteurs normaux unitaires respectifs qui déterminent l'orientation de S_1 et S_2 (c'est à dire définis de manière unique et qui varie continument, comme définis dans la discussion sous la Définition 2.90). Alors, $\vec{n}_1$ et $\vec{n}_2$ déterminent chacun l'orientation de C par la règle de la main droite. Pour le visualiser, pour $j = 1, 2$, pointez le pouce de votre main droite le long de $\vec{n}_j$; la direction de vos doigts indiquera l'orientation de C. Si C reçoit des orientations opposées de $\vec{n}_1$ et $\vec{n}_2$, alors S_1 et S_2 sont orientées de manière cohérente ; si C reçoit la même orientation, alors S_1 et S_2 sont orientées de manière incohérente.

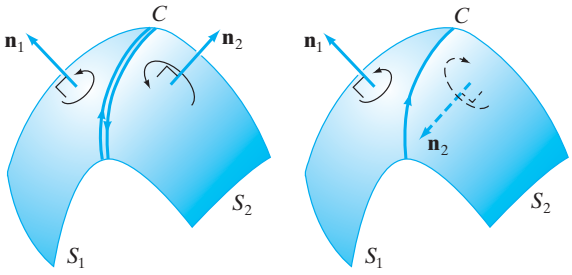


Figure 7.28 The piecewise smooth surface $S = S_1 \cup S_2$ oriented consistently on the left and inconsistently on the right.

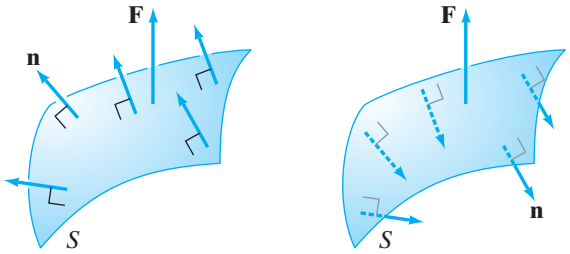


Figure 7.26 How flux depends on orientation. On the left, the surface S is oriented by unit normal vectors so that $\mathbf{F} \cdot \mathbf{n}$ is positive at every point. Hence, $\iint_S \mathbf{F} \cdot d\mathbf{S} = \iint_S \mathbf{F} \cdot \mathbf{n}\, dS$ is positive. On the right, S is given the opposite orientation so that the flux $\iint_S \mathbf{F} \cdot d\mathbf{S} < 0$.

Définition 2.94. *Soit S une surface de l'espace lisse par morceaux, considérée comme l'union de surfaces S_i lisses orientées. On dit que S est* orientée *si l'orientation de la frontière de chaque surface S_i est cohérente avec l'orientation de S_i, et si les orientations des surfaces S_i et S_j sont cohérentes entre elles, pour tout i, j.*

Grâce à cette définition, nous pouvons désormais intégrer des fonctions et des champs de vecteurs sur des surfaces de l'espace lisses par morceaux. Comme pour les courbes planes orientées, un choix d'orientation n'a pas de conséquences sur une intégrale surfacique scalaire, mais a des conséquences sur une intégrale surfacique vectorielle, car prendre $\vec{N}$ ou $-\vec{N}$ change le signe de l'intégrale vectorielle. Cela permet de définir une version de la Définition 2.65 adaptée aux surfaces de l'espace et de généraliser le Théorème de Green à tout type de surface de l'espace.

Définition 2.95. *Soit S une surface de l'espace lisse par morceaux orientée et soit $f : \mathbb{R}^3 \to \mathbb{R}$ une fonction continue. On définit l'intégrale de f sur S par*

$$\iint_S f dS = \iint_{\mathbf{X}} f dS$$

où $\mathbf{X} : D \to \mathbb{R}^3$ est n'importe quelle paramétrisation lisse par morceaux de la surface S, i.e. telle que $S = S_{\mathbf{X}}$.

Soit $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs continu. On définit l'intégrale de $\vec{F}$ sur S par :

$$\iint_S \vec{F} \cdot d\vec{S} = \iint_{\mathbf{X}} \vec{F} \cdot d\vec{S}$$

où $\mathbf{X} : D \to \mathbb{R}^2$ est n'importe quelle paramétrisation lisse par morceaux de la surface de l'espace S, qui est cohérente avec l'orientation de S.

Si la surface de l'espace orientée est fermée, on note $\oiint_S$ le signe intégral pour marquer le fait qu'on intègre sur une surface fermée.

Theorème de Stokes. *Soit S une surface bornée de l'espace lisse par morceaux orientée. Soit $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs de classe $\mathcal{C}^1$, de composantes $F^1 : \mathbb{R}^3 \to \mathbb{R}$, $F^2 : \mathbb{R}^3 \to \mathbb{R}$ et $F^3 : \mathbb{R}^3 \to \mathbb{R}$, respectivement. Alors nous avons (sous réserve que les intégrales existent) :*

$$\oint_{\partial S} \vec{F} \cdot d\vec{s} = \iint_S \begin{pmatrix} \frac{\partial F^3}{\partial y} - \frac{\partial F^2}{\partial z} \\ -\frac{\partial F^1}{\partial z} + \frac{\partial F^3}{\partial x} \\ \frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \end{pmatrix} \cdot d\vec{S}$$

Attention il ne faut pas apprendre cette formule qui est trop compliquée ! ! C'est juste pour montrer l'analogie avec le Théorème de Green. Nous allons introduire de nouveaux opérateurs qui nous permettent de récrire ces deux théorèmes de façon compacte.

Définition 2.96. *Soit $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs de classe $\mathcal{C}^1$. Alors*

— *la* divergence *de $\vec{F}$ est la trace de la matrice Jacobienne de $\vec{F}$:*

$$\begin{aligned} \operatorname{div}(\vec{F}) : \ \mathbb{R}^3 &\longrightarrow \mathbb{R} \\ (x,y,z) &\longmapsto \frac{\partial F^1}{\partial x}(x,y,z) + \frac{\partial F^2}{\partial y}(x,y,z) + \frac{\partial F^3}{\partial z}(x,y,z) \end{aligned}$$

— *le* rotationnel *de $\vec{F}$ (*curl *en anglais) est le champ de vecteur continu défini par :*

$$\begin{aligned} \overrightarrow{\operatorname{rot}}(\vec{F}) : \ \mathbb{R}^3 &\longrightarrow \mathbb{R}^3 \\ (x,y,z) &\longmapsto \begin{pmatrix} \frac{\partial F^3}{\partial y} - \frac{\partial F^2}{\partial z} \\ -\frac{\partial F^1}{\partial z} + \frac{\partial F^3}{\partial x} \\ \frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \end{pmatrix}(x,y,z) \end{aligned}$$

Remarque 2.97. On peut voir le rotationnel de $\vec{F}$ de manière non rigoureuse mais métaphorique comme un "produit vectoriel" :

$$\overrightarrow{\operatorname{rot}}(\vec{F}) = \begin{pmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{pmatrix} \wedge \vec{F}$$

Theorème de Stokes. *Soit S une surface bornée de l'espace lisse par morceaux orientée. Soit $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs de classe $\mathcal{C}^1$, de composantes $F^1 : \mathbb{R}^3 \to \mathbb{R}$, $F^2 : \mathbb{R}^3 \to \mathbb{R}$ et $F^3 : \mathbb{R}^3 \to \mathbb{R}$, respectivement. Alors nous avons (sous réserve que les intégrales existent) :*

$$\oint_{\partial S} \vec{F} \cdot d\vec{s} = \iint_S \overrightarrow{\operatorname{rot}}(\vec{F}) \cdot d\vec{S}$$

Démonstration. Voir pages 500-503 du livre de Susan Jane Colley. □

Corollaire 2.98. Théorème de Green. *Lorsque S est une surface de l'espace horizontale contenue dans le plan d'équation $z = 0$ et orientée par le vecteur $\vec{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$, et que $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$ est un champ de vecteur, alors le Théorème de Stokes est le Théorème de Green.*

Démonstration. En effet dans ce cas S est un domaine de $\mathbb{R}^2$, sa frontière est une courbe plane, et $d\vec{S} = \vec{k}dA$. Nous avons alors sur le plan d'équation $z = 0$:

$$\overrightarrow{\operatorname{rot}}(\vec{F}) \cdot d\vec{S} = \left\langle \begin{pmatrix} \frac{\partial F^3}{\partial y} - \frac{\partial F^2}{\partial z} \\ -\frac{\partial F^1}{\partial z} + \frac{\partial F^3}{\partial x} \\ \frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \end{pmatrix}(x,y,0), \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle dA = \left(\frac{\partial F^2}{\partial x} - \frac{\partial F^1}{\partial y} \right)(x,y,0)dxdy$$

C'est bien l'intégrande de l'intégrale surfacique du Théorème de Green. □

Nous avons quelques identités remarquables avec les opérateurs différentiels divergence et rotationnels. Premièrement, on remarque que le Laplacien d'une fonction $f : \mathbb{R}^3 \to \mathbb{R}$ se définit à partir de la divergence de son gradient :

$$\Delta f = \operatorname{div}(\nabla f) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

C'est la trace de la Hessienne de la fonction f. Maintenant, observons que pour toute fonction $f : \mathbb{R}^3 \to \mathbb{R}$ de classe au moins $\mathcal{C}^2$ et tout champ de vecteurs $\vec{F} : \mathbb{R}^3 \to \mathbb{R}^3$de classe au moins $\mathcal{C}^2$, on a :

$$\overrightarrow{\operatorname{rot}}(\nabla f) = \vec{0} \qquad \text{et} \qquad \operatorname{div}(\overrightarrow{\operatorname{rot}}(\vec{F})) = 0$$

Ce sont deux observations fondamentales et un premier résultat de ce qu'on appelle l'*algèbre homologique*. On peut symboliser les deux équations ci dessus par le diagramme suivant, où on a noté $\mathfrak{X}(\mathbb{R}$ l'espace vectoriel des champs de vecteurs sur $\mathbb{R}^3$:

$$\mathcal{C}^\infty(\mathbb{R}^3) \xrightarrow{\nabla} \mathfrak{X}(\mathbb{R}^3) \xrightarrow{\overrightarrow{\text{rot}}} \mathfrak{X}(\mathbb{R}^3) \xrightarrow{\text{div}} \mathcal{C}^\infty(\mathbb{R}^3)$$

Ce type de diagramme est appelé *complexe de chaines*, et sa particularité est que si on compose deux flèches, on est identiquement nul. On appelle ce complexe de chaines le *complexe de de Rham* (pour $\mathbb{R}^3$). Il existe des complexes de chaines avec des espaces vectoriels plus généraux dans beaucoup de domaines des mathématiques et de la physique, et ils jouent un rôle important dans la compréhension des problèmes rencontrés, par exemple le complexe de Dolbeault, de Hochschild, de Floer, etc.

Dans ce contexte, le théorème de Stokes admet une formulation en dimension quelconque sur des surfaces quelconques, qu'on peut écrire comme :

$$\int_{\partial M} \omega = \int_M d\omega$$

Ici, M est une surface de dimension n, ω est une fonction définie sur M qui a des propriétés particulières et qu'on appelle *forme différentielle*. L'opérateur d qui agit sur ω est un opérateur de complexe de chaine, comme ∇, $\overrightarrow{\text{rot}}$ ou div. Le Théorème de Stokes en dimension n prend ainsi tout se sens dans le contexte des complexes de chaine, et représente un premier exemple important d'une notion mathématique qu'on appelle *homologie* et *cohomologie*, qui est centrale dans des dizaines de domaines en mathématiques. C'est la porte d'entrée vers des résultats très profonds en géométrie différentielle, avec même des applications en physique mathématique. En effet, on peut interpréter les équations de Schwinger-Dyson comme la formule de Stokes sur l'espace (de dimension infinie) des configurations des champs quantiques.