Math 255: Analyse pour la Physique:

Notes initialement rédigées par Julien Sabin en 2018–2019
Cours donné par Frédéric Bourgeois en 2021–2022
Puis Guy David en 2022–2024
Puis Stéphane Nonnenmacher en 2024–2027
stephane.nonnenmacher@universite-paris-saclay.fr

L2 Physique
S3 – 2026-2027

Introduction

Ce polycopié a été rédigé par Julien Sabin, puis utilisé avec plaisir et reconnaissance par Frédéric Bourgeois, Guy David puis Stéphane Nonnenmacher.

Guy David a légèrement modifié la version antérieure, et a précédé ses modifications d'un "GD :".

L'idée générale du cours est de vous apprendre à manipuler des fonctions réelles de plusieurs variables (par exemple $f(x, y, z)$), et dans l'idée notamment d'arriver aux formules sur les intégrales multiples que vous utiliserez constamment en physique.

Lisez la table des matières pour vous faire une idée plus précise.

Malheureusement, on n'a pas le temps de vous faire toutes les démonstrations, donc on se contentera souvent de vous donner une idée de comment ça marche, et de vous apprendre à faire les calculs.

En pratique, pour cette année 2025–2026, je mettrai à jour le fichier de temps en temps sur eCampus, qui sera donc légèrement différent du poly distribué à l'avance, mais pas trop. En principe, le dernier poly valable est celui qui est en ligne sur eCampus.

Table des matières

1 Fonctions de plusieurs variables **5**

1.1 Rappels sur les fonctions d'une seule variable 5

1.1.1 Continuité 5

1.1.2 Dérivabilité 10

1.1.3 Développements limités et formule de Taylor 13

1.1.4 Fonctions de 1 variable à valeurs vectorielles 20

1.2 Limite et continuité de fonctions de plusieurs variables 21

1.2.1 Limites 22

1.2.2 Continuité 28

1.3 Dérivées partielles 30

1.3.1 Définition 30

1.3.2 Généralisation aux applications vectorielles 36

1.3.3 Dérivée d'une composition 38

1.3.4 Dérivées d'ordre supérieur 43

1.4 Application à l'étude des extrema locaux 47

1.4.1 Cas des fonctions d'une variable 48

1.4.2 Cas des fonctions de deux (ou n) variables 50

1.5 Analyse vectorielle 56

2 Intégrales doubles, triples **65**

2.1 Rappels sur les intégrales simples 65

2.1.1 Définition et interprétation 65

2.1.2 Calcul des intégrales simples 67

2.2 Intégrales doubles 71

2.2.1 Définition, propriétés 72

2.2.2 Calcul des intégrales doubles 74

2.3 Intégrales triples 82

2.3.1 Définition, propriétés 82

2.3.2 Calcul des intégrales triples 83

3 Intégrales curvilignes **91**
3.1 Courbes paramétrées 91
3.2 Longueur d'une courbe, intégration d'une fonction par rapport à la longueur 96
3.2.1 Longueur d'une courbe 96
3.2.2 Intégration d'une fonction par rapport à la longueur d'une courbe . . 99
3.3 Circulation d'un champ de vecteurs le long d'une courbe orientée 103
3.4 La formule de Green-Riemann 111

4 Intégrales surfaciques **121**
4.1 Surfaces paramétrées 121
4.2 Aire d'une surface, intégration par rapport à l'aire d'une surface 126
4.3 Surfaces orientées 132
4.4 Flux d'un champ de vecteurs à travers une surface orientée 135
4.5 Formules de Stokes et de Green-Ostrogradski 137

Chapitre 1

Fonctions de plusieurs variables

Le but de ce premier chapitre est de développer les outils de dérivation des fonctions de plusieurs variables. On commence par rappeler les notions importantes pour les fonctions d'une variable, avant de les généraliser aux fonctions de plusieurs variables.

1.1 Rappels sur les fonctions d'une seule variable

Dans toute ce paragraphe, on va considérer des fonctions d'une seule variable réelle, c'est-à-dire des fonctions dont le domaine de définition est un sous-ensemble de $\mathbb{R}$, en général un sous-intervalle. Soit donc $I \subset \mathbb{R}$ un sous-ensemble de $\mathbb{R}$ (typiquement, un intervalle comme $]0,1[$). On notera $f : I \to \mathbb{R}$ pour une fonction dont le domaine de définition est I, prenant des valeurs réelles : pour tout $x \in I$, on a $f(x) \in \mathbb{R}$.

1.1.1 Continuité

Définition 1.1.1: Fonction continue en un point

Soit $f : I \to \mathbb{R}$ une fonction, et $a \in I$ un point quelconque de I. On dit que f est *continue en a* si

$$\lim_{\substack{x \to a \\ x \neq a}} f(x) = f(a), \tag{1.1}$$

autrement dit "les valeurs de $f(x)$ se rapprochent de $f(a)$ lorsque x se rapproche de a".

Exemple 1.1.2

— Soit $I = \mathbb{R}$ et $f(x) = x^2$ pour tout $x \in \mathbb{R}$. Pour tout $a \in \mathbb{R}$, on a bien

$$\lim_{x \to a} x^2 = a^2 = f(a).$$

La fonction f est donc continue en a, et ce pour tout $a \in I$.

— Soit $f : \mathbb{R} \to \mathbb{R}$ la fonction définie par $f(x) = 0$ pour tout $x < 0$ et par $f(x) = 1$ pour

tout $x \geq 0$. Alors,

$$\lim_{\substack{x\to 0\\ x<0}} f(x) = 0, \quad \lim_{\substack{x\to 0\\ x>0}} f(x) = 1,$$

donc f n'admet pas de limite en 0 : en particulier, f n'est pas continue en 0. On notera cette fonction $\mathbb{1}_{\mathbb{R}_+}$, c'est la *fonction indicatrice de* $\mathbb{R}_+$.

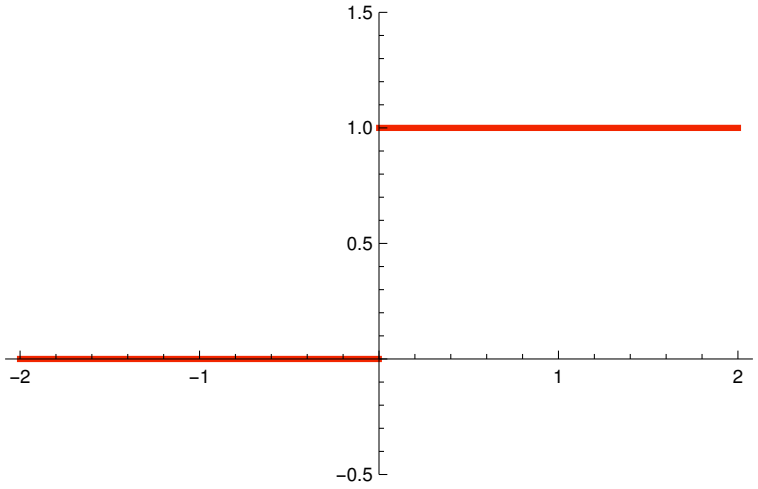

FIGURE 1.1 – Graphe de la fonction f discontinue en 0

Remarque 1.1.1 Le dernier exemple montre un phénomène important : x peut tendre vers a de plusieurs façons différentes (ici, x tend vers 0 par la gauche ou par la droite). Pour que f soit continue en a, il faut en particulier considérer toutes les façons possibles pour x de tendre vers a et que dans tous les cas, on trouve la même limite. Ce n'était pas le cas dans l'exemple.

Remarque 1.1.2 La définition rigoureuse de limite peut se faire via des suites : $f(x)$ tend vers une limite ℓ lorsque x tend vers a si et seulement si, pour toute suite $(x_n)_{n\geq 1}$ convergeant vers a avec $x_n \neq a$, la suite $(f(x_n))$ converge vers ℓ. En particulier, la suite (x_n) peut tendre vers a de n'importe quelle manière (par la gauche, par la droite, osciller entre les deux, à n'importe quelle vitesse, etc..).

Par exemple, la fonction

$$f : \begin{cases} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \begin{cases} \sin(1/x) & \text{si } x \neq 0 \\ 0 & \text{si } x = 0 \end{cases} \end{cases}$$

dessinée ci-dessous n'est pas continue en 0 : on a $f(1/(2\pi n)) = 0$ et $f(1/(\pi/2 + 2\pi n)) = 1$ pour tout $n \in \mathbb{N}$, et les suites $(1/(2\pi n))$ et $(1/(\pi/2 + 2\pi n))$ convergent bien vers 0. Ici, on a deux suites qui convergent vers zéro par valeurs positives mais la valeur de la fonction f le long de ces suites est différente. Ceci est dû au fait que f oscille fortement au voisinage de 0.

Notation En raison de l'hypothèse $x_n \neq a$, on parle de *limite épointée*, ou *limite par valeurs différentes*, et on notera

$$\lim_{\substack{x \to a \\ x \neq a}} f(x) = \ell.$$

Remarque 1.1.3 Cette définition par les suites n'est pas forcément la plus simple à utiliser. Quand il s'agit de démontrer quelque chose au sujet des limites ou de la continuité d'une fonction, c'est la définition avec ε et δ qui fait foi. On dit que la fonction $f : I \to \mathbb{R}$ (où $I \subset \mathbb{R}$ est un intervalle) a la limite ℓ au point $a \in I$ quand

$$\begin{gathered} \text{pour tout nombre } \varepsilon > 0, \text{ il existe } \delta > 0 \text{ tel que} \\ |f(x) - \ell| < \varepsilon \text{ pour tout } x \in I \setminus \{a\} \text{ tel que } |x - a| < \delta. \end{gathered} \tag{1.2}$$

Autrement dit : pour tout degré de précision $\varepsilon > 0$, si on veut que $f(x)$ soit proche de ℓ avec la précision ε, il suffit de demander que x soit assez proche de a (donc, à distance $\leq \delta$).

Et donc, pour $f : I \to \mathbb{R}$ et $a \in I$, on dit que f est continue au point a quand la limite $\ell = f(a)$, autrement dit si :

$$\begin{gathered} \text{pour tout } \varepsilon > 0, \text{ il existe } \delta > 0 \text{ tel que} \\ |f(x) - f(a)| < \varepsilon \text{ pour tout } x \in I \text{ tel que } |x - a| < \delta. \end{gathered} \tag{1.3}$$

Remarque 1.1.4 Pour la limite on a exclu a de la discussion (vous vous souvenez que f n'a même pas besoin d'être définie en a), alors que pour la continuité il faut que $f(a)$ soit définie.

Si f admet une limite ℓ en un point a, alors on peut définir ou modifier f au point a, en prenant $f(a) = \ell$, de façon que f devienne continue en a.

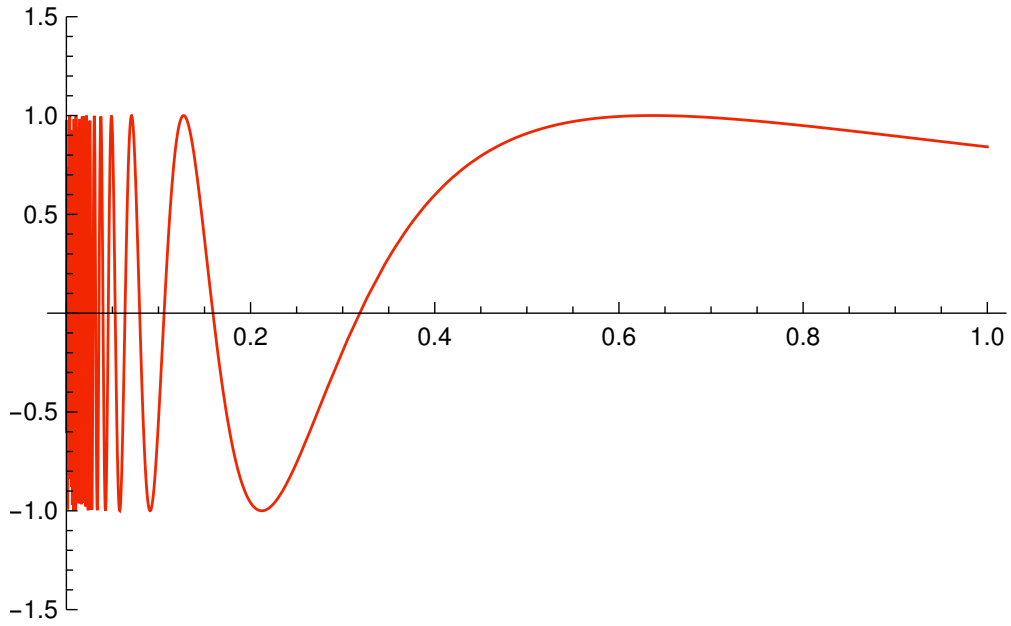

FIGURE 1.2 – Graphe de $x \mapsto \sin(1/x)$.

Définition 1.1.3: Fonction continue sur un domaine

Si pour tout $a \in I$, f est continue en a, alors f est *continue sur* I (ou continue tout court).

Remarque 1.1.5 Si $f : I \to \mathbb{R}$ est continue sur I, alors pour tout sous-ensemble $I' \subset I$, la restriction $f_{\restriction I'}$ est continue sur I'.

Exemple 1.1.4 Toutes les fonctions élémentaires usuelles (x^2, x^3, exp, ln, cos, sin, $1/x$,...) sont continues *sur leur domaine de définition*. Par exemple, ln est continue sur $]0, +\infty[$, et $1/x$ est continue sur $\mathbb{R}^* =]-\infty, 0[\cup]0, +\infty[$ (mais elle n'admet pas de limite en 0, donc elle ne peut pas être prolongée en une fonction continue sur $\mathbb{R}$).

A partir de quelques fonctions continues, on peut en construire beaucoup d'autres.

Proposition 1.1.1: Opérations sur les fonctions continues

1. Si f et g sont deux fonctions continues sur I, alors $f + g$ et fg sont continues sur I ;
2. Si f est continue sur I est λ est un réel quelconque, alors λf est continue sur I ;
3. Si $f : I \to J$ et $g : J \to \mathbb{R}$ sont continues, alors $g \circ f : I \to \mathbb{R}$ (composition de f par g) est continue sur I.

Attention Dans le dernier point (composition de fonctions), il est crucial que l'ensemble d'arrivée de f, noté ici J, coïncide avec (ou soit inclus dans) l'ensemble de départ de g, afin que la composition $g \circ f$ soit bien définie : on rappelle qu'elle est définie par la relation

$$\forall x \in I,\ (g \circ f)(x) = g(f(x)).$$

Pour pouvoir appliquer g au point $f(x)$, il faut bien que $f(x)$ appartienne à l'ensemble de départ de g.

Exemple 1.1.5 — La fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \cos(x) + \sin(x) \end{array} \right.$$

est continue sur $\mathbb{R}$. En effet, les fonctions élémentaires

$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \cos(x) \end{array} \right. , \quad h : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \sin(x) \end{array} \right.$$

le sont, et on a $f = g + h$ qui est donc continue en tant que somme de fonctions continues.

— La fonction
$$f : \left\{ \begin{array}{rcl} \mathbb{R}_+^* =]0, +\infty[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \cos(\ln(x)) \end{array} \right.$$
est continue sur $\mathbb{R}_+^*$, en tant que composition ($f = g \circ h$) des fonctions élémentaires
$$g : \left\{ \begin{array}{rcl} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \cos(y) \end{array} \right. , \quad h : \left\{ \begin{array}{rcl} \mathbb{R}_+^* & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \ln(x) \end{array} \right.$$
qui sont continues.

— La fonction
$$f : \left\{ \begin{array}{rcl} I & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{1}{\cos(x)} \end{array} \right.$$
est continue sur
$$I = \mathbb{R} \setminus \left\{ \frac{\pi}{2} + 2k\pi \ : \ k \in \mathbb{Z} \right\}.$$
En effet, pour tout $x \in I$, on a $\cos(x) \neq 0$ et donc les fonctions
$$g : \left\{ \begin{array}{rcl} I & \longrightarrow & \mathbb{R}^* \\ x & \longmapsto & \cos(x) \end{array} \right. , \quad h : \left\{ \begin{array}{rcl} \mathbb{R}^* & \longrightarrow & \mathbb{R} \\ y & \longmapsto & 1/y \end{array} \right. ,$$
sont bien définies, et continues en tant que fonctions élémentaires [1]. De plus, on a bien $f = h \circ g$ (on a bien fait attention à ce que l'ensemble de départ de h coïncide avec l'ensemble d'arrivée de g) et donc f est continue sur I en tant que composition de fonctions continues.

— La fonction
$$f : \left\{ \begin{array}{rcl}]-1, +\infty[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \ln(1+x) \end{array} \right.$$
est continue sur $]-1, +\infty[$. En effet, pour tout $x \in]-1, +\infty[$, on a $1 + x > 0$ et donc les fonctions
$$g : \left\{ \begin{array}{rcl}]-1, +\infty[& \longrightarrow &]0, +\infty[\\ x & \longmapsto & 1+x \end{array} \right. , \quad h : \left\{ \begin{array}{rcl}]0, +\infty[& \longrightarrow & \mathbb{R} \\ y & \longmapsto & \ln(y) \end{array} \right.$$
sont bien définies et continues, en tant que fonctions élémentaires [2]. On a de plus $f = h \circ g$ (là encore, on a bien fait attention à ce que l'ensemble d'arrivée de g coïncide avec l'ensemble de départ de h), donc f est continue en tant que composition de fonctions continues.

Remarque 1.1.6 Il sera important par la suite de décomposer proprement les fonctions en fonctions élémentaires comme on vient de le faire dans les exemples ci-dessus : cela permettra de ne pas se tromper dans le calcul des dérivées.

1. On a vu que la fonction cos est continue sur $\mathbb{R}$. Ici, on a restreint la fonction cos à I : d'après la Remarque 1.1.5, la restriction $cos_{\restriction I}$ reste continue.

2. La fonction g est élémentaire car c'est une fonction *polynomiale* (et même affine). Les fonctions polynomiales sont toutes continues.

1.1.2 Dérivabilité

Définition 1.1.6: Dérivée d'une fonction

Soit $f : I \to \mathbb{R}$ une fonction et $a \in I$. On dit que f est *dérivable en* a si la limite

$$\lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

existe. Dans ce cas, on appelle cette limite *dérivée de* f *en* a, que l'on note $f'(a)$.

Si f est dérivable en tout point de I, on dit que f est dérivable sur I (ou dérivable tout court).

Remarquons que la définition ci-dessus inclut le cas où a est au bord de l'intervalle I.

Remarque 1.1.7 La dérivée de f au point a correspond à la pente de la tangente au graphe de f au point $(a, f(a))$ (voir la Figure 1.3 et (1.10) pour le rappel de l'équation de la tangente).

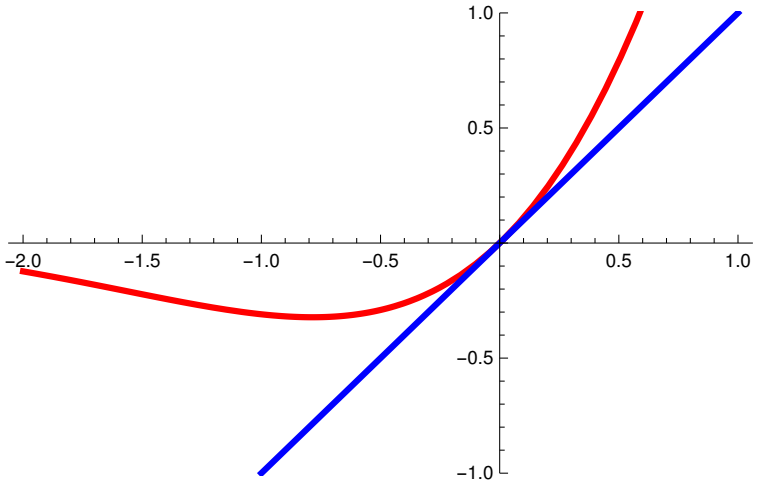

FIGURE 1.3 – Graphe de f (rouge) et sa tangente en l'origine (bleu).

Exemple 1.1.7 Les fonction usuelles (élémentaires) sont dérivables sur leur domaine de définition (sauf peut-être en un point). Par exemple, la fonction racine carrée est continue sur $[0, +\infty[$ mais dérivable uniquement sur $]0, +\infty[$. Pour se rappeler des dérivées des fonctions usuelles, on pourra consulter la Table 5.1 page 60 du poly de Math 101.

Proposition 1.1.2

Si f est dérivable en a, alors f est continue en a.

En particulier, si une fonction n'est pas continue en un point, elle n'a aucune chance d'être dérivable en ce point. La fonction $\mathbb{1}_{\mathbb{R}_+}$ définie dans l'Exemple 1.1.2 n'est donc pas dérivable en 0.

Exemple 1.1.8 la fonction racine carrée $f : \left\{ \begin{array}{ccc} \mathbb{R}_+ & \longrightarrow & \mathbb{R}_+ \\ x & \longmapsto & \sqrt{x} \end{array} \right.$ est un exemple de fonction continue en un point mais non dérivable en ce point (l'origine). Géométriquement, c'est dû au fait que le graphe de la fonction racine carrée admet une tangente verticale en $(0, 0)$, qui est donc de pente infinie.

La non-dérivabilité peut cependant être due à un autre phénomène. Si l'on considère la fonction valeur absolue

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & |x| \end{array} \right. ,$$

elle est continue en $x = 0$ mais on a

$$\lim_{\substack{x \to 0 \\ x>0}} \frac{f(x) - f(0)}{x - 0} = \lim_{\substack{x \to 0 \\ x>0}} \frac{x - 0}{x - 0} = 1, \quad \lim_{\substack{x \to 0 \\ x<0}} \frac{f(x) - f(0)}{x - 0} = \lim_{\substack{x \to 0 \\ x>0}} \frac{-x - 0}{x - 0} = -1,$$

ce qui montre que f n'est pas dérivable en 0. Ici, le phénomène est différent : on a deux limites différentes selon qu'on approche $x = 0$ par la gauche ou par la droite. Géométriquement, on a une "tangente à gauche" et une "tangente à droite" au graphe de f au point $(0, 0)$, dont les pentes sont différentes (voir la Figure 1.4).

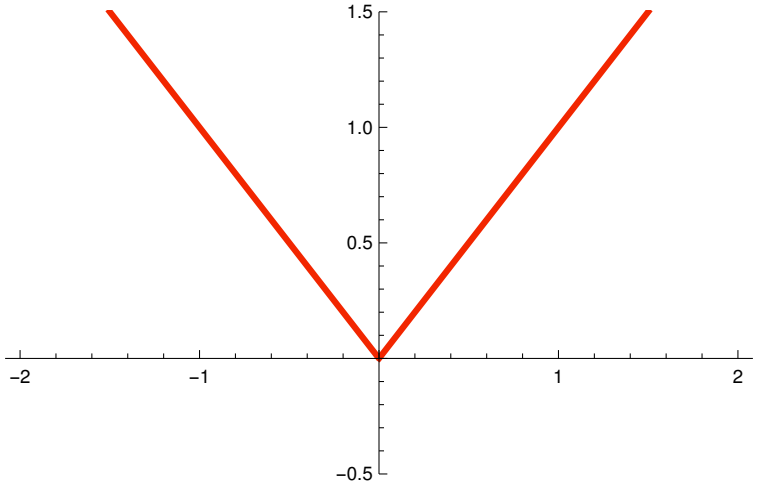

FIGURE 1.4 – Fonction admettant une tangente à gauche et à droite en l'origine

Comme pour les fonctions continues, à partir de fonctions dérivables on peut en construire beaucoup d'autres.

Proposition 1.1.3: Opérations sur les fonctions dérivables

1. Si f et g sont deux fonctions dérivables sur I, alors $f + g$ est dérivable sur I et

$$(f + g)' = f' + g'.$$

2. Si f est dérivable sur I est λ est un réel quelconque, alors λf est dérivable sur I et
$$(\lambda f)' = \lambda f'.$$
3. Si f et g sont deux fonctions dérivables sur I, alors fg est dérivable sur I et
$$\boxed{(fg)' = f'g + fg'} \quad \text{(Formule de Leibniz)}$$
4. Si $f : I \to J$ et $g : J \to \mathbb{R}$ sont dérivables, alors $g \circ f : I \to \mathbb{R}$ est dérivable sur I et
$$\boxed{(g \circ f)' = (g' \circ f) \times f'} \quad \text{(Règle de chaîne).}$$
Il est crucial ici que l'ensemble image de f soit égal, ou au moins inclus dans l'ensemble source de g (ici, l'ensemble J).

Les propriétés 1. et 2. montrent la *linéarité* de la dérivée par rapport aux combinaisons linéaires de fonctions.

Exemple 1.1.9 — Soit la fonction
$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 + \cos(x) \end{array} \right. .$$
On peut la décomposer à l'aide des fonctions
$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \end{array} \right. , \quad h : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \cos(x) \end{array} \right.$$
qui sont dérivables en tant que fonction élémentaires. On a de plus $f = g + h$, donc f est également dérivable en tant que somme de fonctions dérivables, et on a de plus pour tout $x \in \mathbb{R}$,
$$f'(x) = g'(x) + h'(x) = 2x - \sin(x).$$

— Soit la fonction
$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x^2 \cos(x) \end{array} \right. .$$
On a cette fois-ci $f = gh$, avec les même fonctions g et h que dans l'exemple précédent. Ainsi, f est dérivable en tant que produit de fonctions dérivables, et on a pour tout $x \in \mathbb{R}$,
$$f'(x) = g'(x)h(x) + g(x)h'(x) = 2x\cos(x) - x^2\sin(x).$$

— Soit la fonction
$$f : \left\{ \begin{array}{ccc}]0, +\infty[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \cos(\ln(x)) \end{array} \right. .$$
On peut la décomposer à l'aide des fonctions
$$g : \left\{ \begin{array}{ccc}]0, +\infty[& \longrightarrow & \mathbb{R} \\ x & \longmapsto & \ln(x) \end{array} \right. , \quad h : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \cos(y) \end{array} \right. ,$$

qui sont dérivables en tant que fonctions élémentaires. On a de plus $f = h \circ g$ donc f est également dérivable en tant que composition de fonctions dérivables et pour tout $x > 0$,

$$f'(x) = g'(x)h'(g(x)) = \frac{1}{x} \times (-\sin(\ln(x))) = -\frac{\sin(\ln(x))}{x}.$$

— Soit la fonction

$$f : \left\{ \begin{array}{ccc} I & \longrightarrow & \mathbb{R} \\ x & \longmapsto & \dfrac{1}{x^2+x^3} \end{array} \right. ,$$

où $I = \{x \in \mathbb{R} \ : \ x^2 + x^3 \neq 0\}$. On peut la décomposer à l'aide des fonctions

$$g : \left\{ \begin{array}{ccc} I & \longrightarrow & \mathbb{R}^* \\ x & \longmapsto & x^2+x^3 \end{array} \right. , \quad h : \left\{ \begin{array}{ccc} \mathbb{R}^* & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \dfrac{1}{y} \end{array} \right. ,$$

qui sont dérivables en tant que fonctions élémentaires. On a de plus $f = h \circ g$ (on a bien fait attention à ce que l'ensemble d'arrivée de g coïncide avec l'ensemble de départ de h), donc f est dérivable en tant que composition de fonctions dérivables, et pour tout $x \in I$,

$$f'(x) = g'(x) \times h'(g(x)) = (2x+3x^2) \times \left(-\frac{1}{(x^2+x^3)^2} \right) = -\frac{2x+3x^2}{(x^2+x^3)^2} = -\frac{2+3x}{x(x+x^2)^2}.$$

On peut préciser ce qu'est l'ensemble I : en effet, comme $x^2 + x^3 = x^2(1 + x)$, on a $x^2 + x^3 = 0$ si et seulement si $x^2 = 0$ ou $1 + x = 0$, c'est-à-dire si et seulement si $x = 0$ ou $x = -1$. Ainsi, $I = \mathbb{R} \setminus \{0, -1\}$.

Définition 1.1.10: Dérivées d'ordre supérieur

Soit f une fonction dérivable. Si f' est elle-même dérivable, on dit que f est *deux fois dérivable*, et on note $(f')' = f'' = f^{(2)}$ la dérivée seconde de f. De même, si f est trois fois dérivable, on note $(f'')' = f''' = f^{(3)}$ la dérivée troisième de f. Si f est k-fois dérivable, on note $f^{(k)}$ sa dérivée k-ième.

Les dérivées d'ordre supérieur apparaissent notamment dans la formule de Taylor, qui leur donne une interprétation que l'on détaille à présent.

1.1.3 Développements limités et formule de Taylor

Dans ce paragraphe on ne va traiter que le cas des ordres 1 et 2 pour les formules de Taylor, et plutôt se concentrer sur leur interprétation et leurs applications. Pour des rappels sur les formules de Taylor et les développements limités, on pourra consulter le chapitre 9 du poly de Math 101 et le chapitre 2 du poly de Math 104.

On commence par le cas de la formule de Taylor à l'ordre 1. On va voir qu'elle n'est qu'en fait une reformulation de la dérivabilité d'une fonction en un point. Soit donc $f : I \to \mathbb{R}$ une fonction et $a \in I$ un point quelconque du domaine de définition de f. On a vu que f est dérivable en a si et seulement si

$$\lim_{\substack{x\to a\\ x\neq a}} \frac{f(x)-f(a)}{x-a} \quad \text{existe, et vaut} \quad f'(a),$$

ce que l'on peut réécrire comme

$$\lim_{\substack{x\to a\\ x\neq a}} \left(\frac{f(x)-f(a)}{x-a} - f'(a) \right) = 0.$$

On définit alors la fonction

$$\epsilon_1(x) = \frac{f(x)-f(a)}{x-a} - f'(a), \tag{1.4}$$

qui est la fonction sous la limite (et qui n'est a priori définie que pour $x \neq a$, étant donné que l'on divise par $x-a$). La dérivabilité de f en a est donc équivalente à la propriété

$$\lim_{\substack{x\to a\\ x\neq a}} \epsilon_1(x) = 0, \tag{1.5}$$

c'est-à-dire que la fonction ϵ tend vers 0 quand x tend vers a : $\epsilon_1(x)$ devient très petit lorsque x tend vers a. En particulier, on peut étendre par continuité la fonction ϵ_1 au point a en définissant $\epsilon_1(a) := 0$, de sorte que la fonction ϵ_1 possède le même domaine de définition que f. Si l'on multiplie à présent la définition (1.4) de la fonction ϵ_1 par $x-a$, on obtient

$$(x-a)\epsilon_1(x) = f(x) - f(a) - f'(a)(x-a),$$

ou encore

$$f(x) = f(a) + f'(a)(x-a) + (x-a)\epsilon_1(x), \tag{1.6}$$

et ce pour tout $x \in I$. L'identité (1.6) avec la propriété (1.5) est donc une reformulation de la dérivabilité de f en a, on l'appelle *formule de Taylor à l'ordre* 1. Elle fournit un *développement limité de f en a à l'ordre* 1.

Proposition 1.1.4: Formule de Taylor à l'ordre 1

Soit $f : I \in \mathbb{R}$ et $a \in I$. Si f est dérivable en a, alors on a le *développement limité de f en a à l'ordre* 1 suivant : pour tout $x \in I$

$$f(x) = f(a) + f'(a)(x-a) + (x-a)\epsilon_1(x), \tag{1.7}$$

où $\epsilon_1 : I \to \mathbb{R}$ est une fonction continue vérifiant

$$\lim_{x\to a} \epsilon_1(x) = 0\,.$$

Exemple 1.1.11 Si f est la fonction exponentielle et $a = 0$, on a $f(0) = 1 = f'(0)$, donc on peut écrire pour tout $x \in \mathbb{R}$,

$$f(x) = 1 + x + x\epsilon_1(x),$$

pour une certaine fonction $\epsilon_1 : \mathbb{R} \to \mathbb{R}$ vérifiant $\epsilon_1(x) \to 0$ lorsque $x \to 0$.

Remarque 1.1.8 Même si l'équation (1.7) est valide pour tout x dans le domaine de définition de f, la seule véritable information qu'elle donne est pour x proche de a. En effet, la seule information dont on dispose est que la fonction ϵ_1 tend vers 0 lorsque x tend vers a : le comportement de la fonction ϵ_1 pour x "loin" de a est inconnu. La formule de Taylor donne donc une information locale au voisinage de a. Cette information s'interprète de la façon suivante : supposons que les coefficients $f(a)$ et $f'(a)$ ne sont pas nuls. Alors, pour x proche de a, $f'(a)(x-a)$ est beaucoup plus petit que $f(a)$ lorsque $x \to a$. En effet, $f'(a)(x-a)$ tend vers 0 lorsque $x \to a$ alors que $f(a)$ est constante et non nulle. Ainsi, dans l'expression

$$f(a) + f'(a)(x-a),$$

le premier terme $f(a)$ peut être vu comme le terme "dominant" et le second terme $f'(a)(x-a)$ comme une "correction" au terme dominant puisqu'il est beaucoup plus petit que celui-ci. Ce raisonnement s'applique de manière similaire avec les deux termes suivants dans (1.7) : lorsque x tend vers a, $\epsilon_1(x)(x-a)$ est beaucoup plus petit que $f'(a)(x-a)$: dans le premier terme, le facteur de $(x-a)$ est $\epsilon_1(x)$ qui tend vers 0 et donc est très petit, alors que le facteur de $(x-a)$ dans le second terme est $f'(a)$ qui est non-nul et constant. Ainsi, dans le terme de droite de (1.7), chaque terme est beaucoup plus petit que le terme à sa gauche lorsque x tend vers a. Un développement limité apporte donc une description de plus en plus précise de la valeur de $f(x)$ lorsque x est proche de a.

L'expression approchant $f(x)$ n'est de plus pas quelconque. Si on néglige le terme d'erreur, on trouve que

$$f(x) \simeq f(a) + f'(a)(x-a)$$

lorsque $x \to a$. Le terme de droite est une fonction affine, dont le graphe est donc une droite. Le développement limité à l'ordre 1 permet donc d'approcher une fonction f au voisinage de a par une fonction affine. Géométriquement, on approche le graphe de f au voisinage du point $(a, f(a))$ par une droite : c'est la **tangente au graphe en ce point**, dont l'équation cartésienne est

$$y = f(a) + f'(a)(x-a). \tag{1.8}$$

Les formules de Taylor d'ordres supérieurs fournissent une description de plus en plus précise (à mesure que l'ordre augmente) du comportement de f au voisinage du point a. On énonce ci-dessous ce qui se passe à l'ordre 2.

Proposition 1.1.5: Formule de Taylor à l'ordre 2

Soit $f : I \to \mathbb{R}$ et $a \in I$. On suppose que f est 2 fois dérivable en a. Alors, f admet le *développement limité à l'ordre 2 au point a* suivant : pour tout $x \in I$,

$$f(x) = f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2 + (x-a)^2\epsilon_2(x), \tag{1.9}$$

où $\epsilon_2 : I \to \mathbb{R}$ est une fonction continue satisfaisant

$$\lim_{x\to a} \epsilon_2(x) = 0.$$

Exemple 1.1.12 Si f est encore la fonction exponentielle, on a $f''(0) = 1$ donc pour tout $x \in \mathbb{R}$,

$$f(x) = 1 + x + \frac{1}{2}x^2 + x^2\epsilon_2(x),$$

où $\epsilon_2 : \mathbb{R} \to \mathbb{R}$ est une fonction continue satisfaisant $\epsilon_2(x) \to 0$ lorsque $x \to 0$.

Remarque 1.1.9 Là encore, il faut réaliser que l'expression (1.9) contient des termes de plus en plus petits les uns que les autres lorsque $x \to a$. Par exemple, si $f''(a) \neq 0$, le dernier terme $(x-a)^2\epsilon_2(x)$ est beaucoup plus petit que le terme $\frac{1}{2}f''(a)(x-a)^2$ quand $x \to a$ puisque dans le premier terme, le facteur de $(x-a)^2$ est $\epsilon_2(x)$ qui tend vers 0 donc est très petit, alors que dans le deuxième terme le facteur de $(x-a)^2$ est $\frac{1}{2}f''(a)$ qui est une constante non nulle. De la même manière, si $f'(a) \neq 0$, le terme $\frac{1}{2}f''(a)(x-a)^2$ est beaucoup plus petit que le terme $f'(a)(x-a)$ quand $x \to a$: en effet, la fonction $y \mapsto y^2$ tend beaucoup plus vite vers 0 que la fonction $y \mapsto y$ lorsque $y \to 0$ (et ici, $y = x - a$ tend bien vers 0). Ainsi, la formule (1.9) contient bien des termes successifs de plus en plus petits.

Ce développement à l'ordre 2 fournit une description plus précise que le développement limité à l'ordre 1 : le terme de reste dans le DL d'ordre 1 (1.7), $(x-a)\epsilon_1(x)$, peut se décomposer en

$$(x-a)\epsilon_1(x) = \frac{1}{2}f''(a)(x-a)^2 + (x-a)^2\epsilon_2(x)\,,$$

ce qui décrit la fonction f plus précisément lorsque $x \to a$. Si l'on néglige le terme d'erreur dans (1.9), l'expression approchant $f(x)$ est

$$f(x) \simeq f(a) + f'(a)(x-a) + \frac{1}{2}f''(a)(x-a)^2$$

lorsque $x \to a$. Le terme de droite est un polynôme de degré 2 en x, dont le graphe est une parabole. Le développement limité à l'ordre 2 approche donc $f(x)$ pour x proche de a par un polynôme de degré 2 en x (alors que le développement limité à l'ordre 1 approche $f(x)$ par une fonction affine en x – autrement dit un polynôme de degré 1 en x). Géométriquement, on approche le graphe de f autour du point $(a, f(a))$ par une parabole (alors qu'à l'ordre 1 on l'approche par une droite), d'équation

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

comme illustré en Figure 1.5. Cette parabole est unique : non seulement elle est tangente au graphe de f au point $(a, f(a))$, mais elle lui est aussi tangente « au 2e ordre », en ce sens qu'elle a la même courbure que le graphe de f au point a (on dit que cette parabole est *osculatrice* au graphe de f en a).

Plus généralement, la formule de Taylor à l'ordre k consiste à approcher $f(x)$ au voisinage de a par un polynôme de degré k (comme illustré en Figure 1.6), dont les coefficients s'expriment à l'aide des dérivées successives de f en a jusqu'à l'ordre k. Ceci fournit donc une interprétation des dérivées d'ordre supérieur de f comme outils dans son approximation par des polynômes. Pour l'expression des formules de Taylor aux ordres supérieurs, on pourra consulter les polys de L1 cités en début de paragraphe.

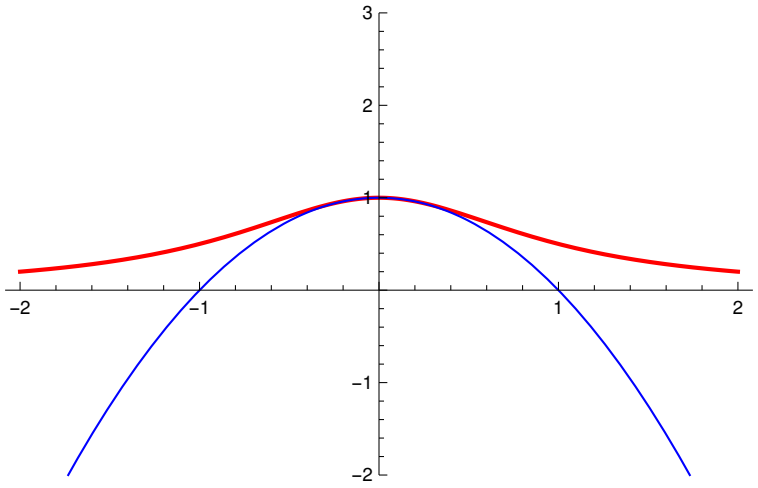

FIGURE 1.5 – Graphe d'une fonction (rouge) et de la parabole donnée par son développement limité à l'ordre 2 en l'origine (bleu)

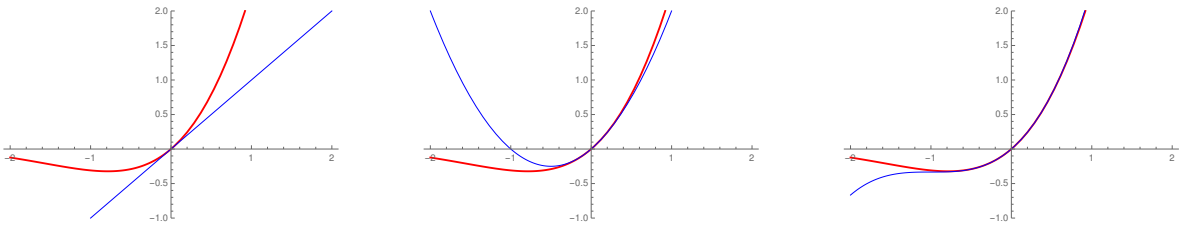

FIGURE 1.6 – Graphe d'une fonction (rouge) et de son développement de Taylor à l'origine (bleu) aux ordres 1 (gauche), 2 (milieu), et 3 (droite).

Nous allons maintenant voir deux applications des développements limités.

Application A : Position du graphe par rapport à la tangente

Les développements limités permettent de déterminer la position relative du graphe d'une fonction et de sa tangente au voisinage d'un point. Soit $f : I \to \mathbb{R}$ et $a \in I$. Si f est dérivable

en a, alors la tangente au graphe de f au point $(a, f(a))$ a pour équation

$$y = f(a) + f'(a)(x - a).$$

Si de plus f est deux fois dérivable en a, on peut comparer localement le graphe de f et sa tangente. En effet, par le développement limité à l'ordre 2 de f, on a

$$f(x) - (f(a) + f'(a)(x - a)) \simeq \frac{1}{2} f''(a)(x - a)^2$$

pour x proche de a. Comme $(x - a)^2 \geq 0$, on en déduit que si $f''(a) > 0$, alors

$$f(x) - (f(a) + f'(a)(x - a)) \simeq \frac{1}{2} f''(a)(x - a)^2 > 0$$

pour x proche de a, $x \neq a$, autrement dit

$$f(x) > f(a) + f'(a)(x - a).$$

Cette inégalité montre que le graphe de f est au-dessus de sa tangente en a, pour x proche de a. Inversement, si $f''(a) < 0$, on a

$$f(x) - (f(a) + f'(a)(x - a)) \simeq \frac{1}{2} f''(a)(x - a)^2 < 0$$

pour x proche de a, $x \neq a$, autrement dit

$$f(x) < f(a) + f'(a)(x - a).$$

Cela implique que le graphe de f est en-dessous de sa tangente en a, pour x proche de a.

Proposition 1.1.6: Position relative du graphe par rapport à sa tangente

Soit $f : I \to \mathbb{R}$ une fonction et $a \in I$. On suppose que f est deux fois dérivable en a. Alors :

- Si $f''(a) > 0$, le graphe de f est situé au-dessus de sa tangente au voisinage du point $(a, f(a))$.
- Si $f''(a) < 0$, le graphe de f est situé en-dessous de sa tangente au voisinage du point $(a, f(a))$.

Exemple 1.1.13 Si f est la fonction exponentielle et $a = 0$, on a $f(0) = 1 = f'(0)$ et donc l'équation de la tangente au graphe de f au point $(0, f(0)) = (0, 1)$ est

$$y = x + 1.$$

Comme $f''(0) = 1 > 0$, le graphe de l'exponentielle est situé au-dessus de sa tangente au voisinage de $(0, 1)$, comme on le voit sur la Figure 1.7.

Remarque : Dans ce cas particulier de la fonction exponentielle, le graphe de f est *partout* au-dessus de la tangente, pas seulement au voisinage du point $(0, 1)$. Cela est dû à une propriété particulière de la fonction exponentielle : sa *convexité*, qui vient du fait que $f''(x) > 0$ pour tout x. Pour montrer que le graphe de f est au-dessus de sa tangente d'équation $y = x + 1$, ou de manière équivalente que $f(x) \geq x + 1$ pour tout $x \in \mathbb{R}$, le plus simple est de faire quelques études de fonctions, comme suit. Puisque $f''(x) > 0$ partout, la fonction f' est strictement croissante sur $\mathbb{R}$ et donc (puisque $f'(0) = 1$), on trouve que $f'(x) < 1$ pour $x < 0$ et $f'(x) > 1$ pour $x > 0$. Ensuite, comme la fonction $x \mapsto f(x) - x - 1$ s'annule en 0 et a une dérivée qui est du signe de x, on peut en déduire qu'elle est strictement positive, à la fois pour $x < 0$ et pour $x > 0$.

En général, la Proposition 1.1.6 ne donne qu'une information au voisinage du point a, comme l'illustre la Figure 1.8, où le graphe est situé au dessous de sa tangente uniquement au voisinage de $(a, f(a))$: plus loin, le graphe peut repasser au dessus de la tangente.

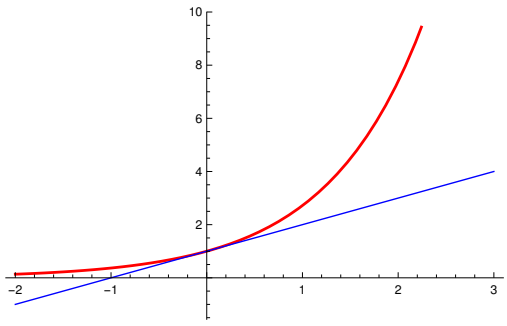

FIGURE 1.7 – Graphe de l'exponentielle (rouge) et de sa tangente en $(0, 1)$ (bleu).

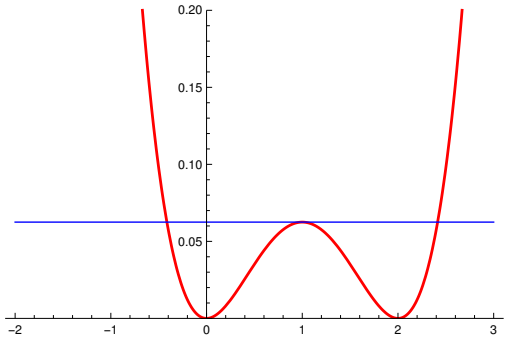

FIGURE 1.8 – Graphe d'une fonction (rouge) et de sa tangente en $(1, 1/16)$ (bleu).

Application B : Formes indéterminées

Les développements limités permettent également de résoudre des formes indéterminées, comme on le montre sur un exemple. On s'intéressera ici au point $a = 0$. Supposons que l'on veuille calculer la limite

$$\lim_{\substack{x \to 0 \\ x \neq 0}} \frac{\sin(x)}{x} \tag{1.10}$$

Lorsque $x \to 0$, à la fois le numérateur et le dénominateur tendent vers 0 : on dit qu'on a une *forme indéterminée*, car on ne peut calculer la limite du quotient en formant le quotient des limites. A priori, une limite du type "0/0" peut avoir différents comportements possibles. Montrons que le développement limité à l'ordre 1 du numérateur $\sin(x)$ au voisinage de $x = 0$ permet de lever cette indétermination :

$$\sin(x) = x + x\epsilon_1(x),$$

où $\epsilon_1 : \mathbb{R} \to \mathbb{R}$ est une fonction satisfaisant $\epsilon_1(x) \to 0$ lorsque $x \to 0$. En divisant cette relation par x, on obtient : pour tout $x \neq 0$,

$$\frac{\sin(x)}{x} = 1 + \epsilon_1(x),$$

qui tend donc vers 1 lorsque $x \to 0$. On a donc levé l'indétermination à l'aide d'un développement limité.

Il peut arriver que l'on doive utiliser un développement limité à des ordres supérieurs, comme on le verra en TD. L'idée principale est que l'on se ramène à des fonctions polynomiales au numérateur et au dénominateur, pour lesquelles il est facile de lever des indéterminations : une limite du type

$$\lim_{\substack{x \to 0 \\ x \neq 0}} \frac{x^m}{x^n}$$

est a priori une forme indéterminée, puisque numérateur et dénominateur tendent vers 0. On lève simplement cette indétermination en remarquant que

$$\frac{x^m}{x^n} = x^{m-n},$$

dont la limite peut prendre différentes valeurs :

$$\lim_{x \to 0} x^{m-n} = \begin{cases} 0 & \text{si } m > n, \\ 1 & \text{si } m = n, \\ \pm\infty & \text{si } m < n. \end{cases}$$

Les développements limités permettent de savoir comme quelle puissance de x se comportent numérateur et dénominateur. Dans notre exemple, $\sin(x)$ se comporte comme $x = x^1$ et le dénominateur est déjà un polynôme qui se comporte comme $x = x^1$, donc on est dans le cas $m = n = 1$.

1.1.4 Fonctions de 1 variable à valeurs vectorielles

Avant de parler de fonctions de plusieurs variables (la partie intéressante du cours), disons deux mots des fonctions **à valeurs dans** un espace vectoriel $\mathbb{R}^n$ (au lieu de $\mathbb{R}$). Par exemple, $f : [0, 1] \to \mathbb{R}^2$. On parle alors de fonction vectorielle, puisque pour tout x, l'image $f(x) = (f_1(x), f_2(x))$ est un vecteur dans $\mathbb{R}^2$, où chaque composante f_1, $f_2 : [0, 1] \to \mathbb{R}$ est une fonction à valeur réelle.

Dans ce cas, il est d'usage d'étudier chacune des composantes de f séparément, et rien de vraiment compliqué n'arrive. Par exemple, on dit que f est dérivable quand f_1 et f_2 sont toutes les deux dérivables. Et on note $f'(x)$ le vecteur dont les deux coordonnées sont $f_1'(x)$ et $f_2'(x)$. En gros, tout ce qu'on a dit ci-dessus se généralise facilement à des fonctions à valeurs dans $\mathbb{R}^2$, ou plus généralement $\mathbb{R}^n$. On en reparlera en section 1.5, mais en tout cas ne confondez pas cette question des fonctions vectorielles avec les fonctions de plusieurs variables (pour lesquelles $\mathbb{R}^n$ est l'espace source).

1.2 Limite et continuité de fonctions de plusieurs variables

On considère dans ce paragraphe des fonctions plusieurs variables réelles $f(x_1, \ldots, x_n)$, où chaque x_j est une variable réelle. On se contentera souvent du cas $n = 2$ (deux variables réelles), et l'extension au cas de $n \geq 3$ variables sera claire. On considérera donc des fonctions $f : \Omega \to \mathbb{R}$, où Ω est un sous-ensemble de $\mathbb{R}^2$ (le domaine de définition de f). Par exemple $\Omega = \mathbb{R}^2$ tout entier, $\Omega = I \times J$ (un rectangle) avec I et J des intervalles réels, ou un produit plus compliqué lorsque I et J ne sont pas des intervalles (essayer d'en dessiner), voire des Ω qui ne sont pas des produits (par exemple si Ω est un disque).

Notation Une des difficultés du formalisme des fonctions à plusieurs variables est dans le choix des notations. Une grande partie du cours se déroulera en dimension $n = 2$, pour laquelle les coordonnées cartésiennes d'un point sont usuellement notées $(x, y) \in \mathbb{R}^2$. En dimensions $n = 3$, le point standard est noté $(x, y, z) \in \mathbb{R}^3$. Mais lorsqu'on s'intéresse à des dimensions supérieures $n > 3$, ou même en dimensions $2, 3$ pour obtenir des expressions plus courtes, il est pratique de noter le point de $\mathbb{R}^n$ par $X = (x_1, \ldots, x_n)$, où $x_1, x_2, \ldots x_n$ sont les coordonnées euclidiennes de X. En interprétant le point X comme le vecteur $\vec{OX} = X - 0$ de $\mathbb{R}^n$, on notera

$$||X|| = \left(x_1^2 + x_2^2 + \ldots x_n^2\right)^{1/2} = \sqrt{x_1^2 + x_2^2 + \ldots x_n^2} \tag{1.11}$$

sa **norme euclidienne** (c'est-à-dire la longueur de ce vecteur). Puis, pour X, Y deux points de $\mathbb{R}^n$, on notera

$$\text{dist}\,(X, Y) = ||X - Y|| = \Big(\sum_{i=1}^{n} |x_i - y_i|^2\Big)^{1/2} \tag{1.12}$$

la **distance** (euclidienne) entre les points X et Y (on a bien sûr noté $Y = (y_1, \ldots, y_n)$ les coordonnées de Y).

Dans tous les cas, les lettres minuscules indiqueront des coordonnées réelles, tandis que les lettres majuscules indiqueront un point dans $\mathbb{R}^n$.

Pour $X \in \mathbb{R}^n$ et $r > 0$, je noterai

$$B(X, r) = \big\{Y \in \mathbb{R}^n \,;\, \text{dist}\,(Y, X) < r\big\} = \big\{Y \in \mathbb{R}^n \,;\, ||Y - X|| < r\big\} \tag{1.13}$$

la **boule ouverte de centre X et de rayon r.** Quand $n = 1$, c'est juste l'intervalle de longueur $2r$ centré en X ; quand $n = 2$ c'est un disque ouvert (sans le cercle de rayon r), et quand $n = 3$ c'est une boule ouverte (sans la sphère au bord).

Le graphe (je vais le noter G_f) de la fonction $f : \Omega \to \mathbb{R}$ est l'ensemble

$$G_f = \big\{(x_1, x_2, \ldots, x_n, f(x_1, x_2, \ldots, x_n)) \,;\, (x_1, x_2, \ldots, x_n) \in \Omega\big\} \subset \mathbb{R}^{n+1} \tag{1.14}$$

composé des points dont les n premières coordonnées donnent un point $X = (x_1, \ldots, x_n) \in \Omega$, et la dernière coordonnée prend la valeur$f(X) = f(x_1, \ldots, x_n)$. De façon plus condensée, le

graphe de f est donné par :

$$G_f = \big\{(X, f(X)) \in \mathbb{R}^n \times \mathbb{R}\,;\, X \in \Omega\big\}. \tag{1.15}$$

Notez bien que G_f est est un sous-ensemble de $\mathbb{R}^{n+1}$. Donc pour une fonction de deux variables, le graphe G_f est un sous-ensemble de $\mathbb{R}^3$, typiquement une surface vivant dans $\mathbb{R}^3$, comme l'illustre la Figure 1.9, et dont la projection sur le plan horizontal est le domaine de définition Ω.

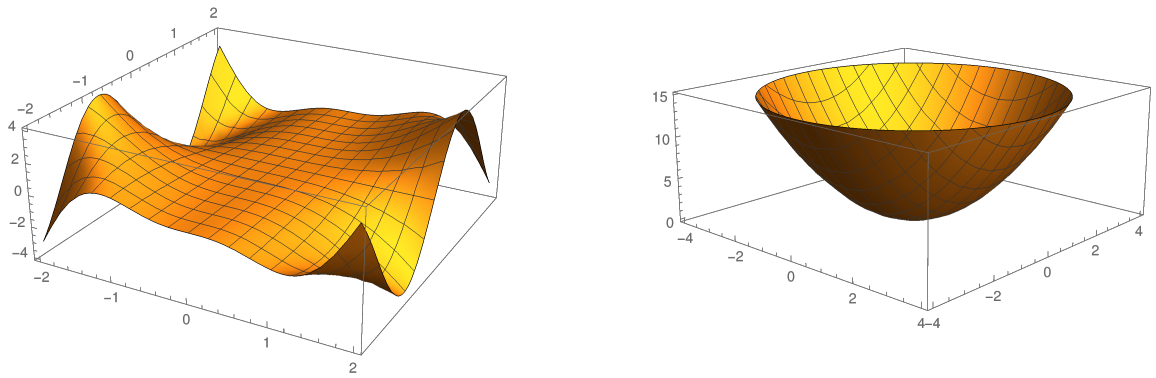

FIGURE 1.9 – Exemples de graphes de fonctions de 2 variables

1.2.1 Limites

La limite d'une fonction de plusieurs variables est le même concept que celui de limite de fonction d'une variable. Cependant, on verra qu'il comporte quelques subtilités qu'on illustrera sur des exemples.

Définition 1.2.1: Limite d'une fonction de 2 variables

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction. Soit $A = (a_1, a_2)$ un point quelconque de Ω. On dit que f admet une limite $\ell \in \mathbb{R}$ lorsque $X = (x_1, x_2)$ tend vers $A = (a_1, a_2)$ si $f(x_1, x_2)$ se rapproche de ℓ lorsque x_1 se rapproche de a_1 et x_2 se rapproche de a_2, tout en assurant que $(x_1, x_2) \neq (a_1, a_2)$. On note

$$\lim_{\substack{X \to A \\ X \neq A}} f(X) = \ell, \quad \text{ou aussi} \quad \lim_{(x_1,x_2)\to(a_1,a_2)} f(x_1, x_2) = \ell\,.$$

Remarque : la définition « officielle » est que lorsque $\Omega \subset \mathbb{R}^n$, $f : \Omega \to \mathbb{R}$ et $A \in \Omega$, on dit que $\lim_{\substack{X \to A \\ X \neq A}} f(X) = \ell$ quand

$$\begin{gathered} \text{pour tout } \varepsilon > 0, \text{ il existe } \delta > 0 \text{ tel que} \\ |f(X) - \ell| < \varepsilon \text{ pour tout } x \in \Omega \setminus \{A\} \text{ tel que } ||X - A|| < \delta. \end{gathered} \tag{1.16}$$

Bref, on fait pareil qu'en (1.2), mais on utilise la distance euclidienne $||X - A||$ pour dire que X est assez proche de A.

Remarque importante. Dans cette histoire, il est important de noter qu'on a besoin de la notion de "X tend vers A" (où A est un point donné de $\mathbb{R}^n$ et X est un autre point de $\mathbb{R}^n$). On essaie souvent de dire à la place quelque chose comme "X se rapproche indéfiniment de A," mais en tout cas on parle de la même notion. Ceci signifie que "la distance entre X et A tend vers 0", c-à-d. que $||X - A||$ tend vers 0. Ce sur quoi je veux insister est le fait que

$$||X - A|| \text{ tend vers 0 si et seulement si } |x_i - a_i| \text{ tend vers 0 pour chaque } i \in \{1, 2, \ldots, n\}, \tag{1.17}$$

où bien entendu j'ai noté $X = (x_1, \ldots, x_n)$ et $A = (a_1, \ldots, a_n)$. Autrement dit, chaque coordonnée de X doit converger vers la coordonnée correspondante de A. Et c'est facile à vérifier : $||X - A||$ tend vers 0 SSI (mon abréviation de si et seulement si) $||X - A||^2$ tend vers 0, SSI $\sum_i |x_i - a_i|^2$ tend vers 0, SSI chaque $|x_i - a_i|^2$ tend vers 0 (car c'est une somme de carrés, donc par exemple chacun des termes est inférieur à la somme), SSI chaque $|x_i - a_i|$ tend vers 0.

Exemple 1.2.2 Soit f la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x_1, x_2) & \longmapsto & x_1^2 + x_2^3 \end{array} \right. .$$

Alors, on a

$$\lim_{(x_1,x_2)\to(1,2)} f(x_1, x_2) = 1^2 + 2^3 = 9.$$

En effet, si x_1 tend vers 1 alors x_1^2 tend vers 1^2, et si x_2 tend vers 2 alors x_2^3 tend vers 2^3.

Exemple 1.2.3 La subtilité de notion de limite multi-dimensionnelle réside dans le fait que $(x_1, x_2) \to (a_1, a_2)$ signifie que x_1 tend vers a_1 *et* que x_2 tend vers a_2, mais les façons dont x_1 et x_2 convergent respectivement vers a_1 et a_2 peuvent être complètement différentes : on doit considérer *toutes* les façons dont x_1 peut tendre vers a_1 et *toutes* les façons dont x_2 tend vers a_2, ce qui amène à une plus grande variété de comportements. Illustrons cela sur un exemple.

Considérons la fonction

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

Ici, on a utilisé (x, y) pour les coordonnées et non (x_1, x_2), pour alléger les notations. On chercher à déterminer

$$\lim_{(x,y)\to(0,0)} f(x, y).$$

D'abord, remarquons que si $x \to 0$ et $y \to 0$, alors $xy \to 0$ et $x^2 + y^2 \to 0$: cette limite est a priori une forme indéterminée. On va montrer que cette limite n'existe pas en considérant trois façons différentes d'approcher l'origine.

— Pour tout $x \neq 0$, on a $f(x,0) = x \times 0/(x^2+0^2) = 0$. Ainsi,

$$\lim_{\substack{x\to 0\\ x\neq 0}} f(x,0) = 0.$$

Lorsque $x \to 0$, le point $(x,0)$ tend bien vers $(0,0)$: il s'agit d'une première façon d'approcher l'origine, en forçant la seconde coordonnée y à rester nulle et en faisant tendre la première coordonnée vers 0. Géométriquement, le point $(x,0)$ se déplace sur l'axe des abscisses lorsque x varie, donc on approche l'origine selon une droite horizontale.

— On approche cette fois-ci l'origine verticalement. Pour tout $y \neq 0$, on a $f(0,y) = 0 \times y/(0^2+y^2) = 0$. Ainsi,

$$\lim_{\substack{y\to 0\\ y\neq 0}} f(y,0) = 0.$$

Lorsque $y \to 0$, le point $(0,y)$ tend encore bien vers $(0,0)$: il s'agit d'une seconde façon d'approcher l'origine, en forçant la première coordonnée x à rester nulle et en faisant tendre la seconde coordonnée vers 0. Géométriquement, le point $(0,y)$ se déplace sur l'axe des ordonnées lorsque y varie, donc on approche l'origine selon une droite verticale.

— Enfin, on va approcher l'origine selon une droite de pente 1. Pour tout $x \neq 0$, on a

$$f(x,x) = \frac{x \times x}{x^2+x^2} = \frac{x^2}{2x^2} = \frac{1}{2}.$$

Ainsi,

$$\lim_{\substack{x\to 0\\ x\neq 0}} f(x,x) = \frac{1}{2}.$$

Lorsque $x \to 0$, le point (x,x) approche bien $(0,0)$. Géométriquement, on approche l'origine selon la diagonale (droite de pente 1, puisque abscisse et ordonnée sont les même).

Selon ces trois façons d'approcher l'origine, on trouve des limites différentes : on trouve 0 si l'on s'approche de l'origine horizontalement ou verticalement, mais on trouve $\frac{1}{2}$ si l'on s'en approche en diagonale. Ainsi, la fonction f n'admet pas de limite en $(0,0)$. On a tracé le graphe de f dans la Figure 1.10 pour remarquer le comportement spécial en $(0,0)$.

Remarque : cet exemple est moins tordu qu'il n'y paraît si on utilise les coordonnées polaires sur $\mathbb{R}^2$: En prenant $x = r\cos(\theta)$ et $y = r\sin\theta$, avec par exemple $0 \leq r < +\infty$ et $\theta \in [0, 2\pi]$, alors $f(x,y) = \frac{xy}{x^2+y^2} = \frac{r^2\cos(\theta)\sin(\theta)}{r^2} = \cos(\theta)\sin(\theta)$. Donc on s'est débrouillé pour que f soit constante sur chaque droite passant par l'origine, avec des valeurs différentes sur des droites différentes. D'où les limites différentes quand (x,y) tend vers $(0,0)$ le long de droites différentes. Essayez de vous convaincre que le graphe ci-dessous est bien obtenu en prenant la droite d'angle θ qui passe par l'origine, en la soulevant verticalement de $\cos(\theta)\sin(\theta)$, puis en

prenant l'union de ces droites qui montent et descendent en même temps qu'elles tournent. Un tel graphe G_f est appelé une *surface réglée*.

Si vous aimez la définition (1.16), vérifiez que pour tout choix de ℓ, la propriété (1.16) est prise en défaut avec $\varepsilon = 1/8$, soit parce que $|\ell| \geq 1/4$, et alors dans toute boule B de centre $(0,0)$ et de rayon δ (aussi petit que vous voulez), il y a encore plein de points (x,y) (ici, sur l'axe horiontal $\{y = 0\}$) tels que $|f(x,y) - \ell| > \varepsilon$, soit parce que $|\ell| < 1/4$, et alors dans toute boule B de centre $(0,0)$ et de rayon δ (aussi petit que vous voulez), il y a encore plein de points (x,y) (maintenant sur la première diagonale $\{x = y\}$) tels que $|f(x,y) - \ell| > \varepsilon$ car $f(x,y) = 1/2$.

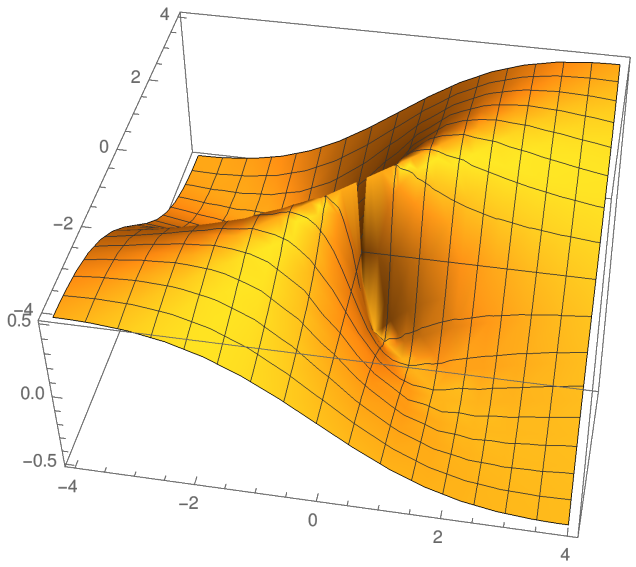

FIGURE 1.10 – Graphe de la fonction f

Exemple 1.2.4 Dans l'exemple précédent, on a approché un point selon trois droites. On pourrait approcher ce point selon des droites de pentes différentes (ce qui donne autant de limites à vérifier), mais aussi selon des trajectoires plus compliquées. Par exemple, pour la fonction

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

définie sur le domaine $\{x \in \mathbb{R},\ y \in \mathbb{R}_+,\ (x,y) \neq (0,0)\}$, on a également $f(x,0) = 0$ pour tout $x \neq 0$, d'où

$$\lim_{\substack{x \to 0 \\ x \neq 0}} f(x,0) = 0.$$

Pour tout $\alpha > 0$, en s'approchant de l'origine le long de la demi-droite de pente α située dans le demi-espace $\{y \geq 0\}$, on observe de même

$$\lim_{\substack{x \to 0 \\ x > 0}} f(x,ax) = 0$$

On trouve le même comportement si on s'approche de $(0,0)$ le long de toute demi-droite de pente $\alpha < 0$. On pourrait donc penser que f admet bien 0 comme limite.

Ce n'est cependant pas le cas. Pour tout $x \neq 0$, on remarque que $f(x,x^2) = \frac{1}{2}$, d'où

$$\lim_{\substack{x\to 0\\ x\neq 0}} f(x,x^2) = \frac{1}{2}.$$

Lorsque $x \to 0$, le point (x,x^2) approche bien $(0,0)$. Géométriquement, il l'approche selon la parabole d'équation $y = x^2$. Comme on ne trouve pas la même limite selon que l'on s'approche de l'origine en suivant n'importe quelle droite ou une parabole, la fonction f n'admet pas de limite en l'origine. *Il ne suffit donc pas de regarder ce qui se passe lorsqu'on s'approche de l'origine de manière rectiligne : il faut aussi considérer toutes les trajectoires courbées !*

Exemple 1.2.5 L'exemple précédent paraît désespérant : on ne peut pas vérifier que toutes les trajectoires possibles donnent la même limite : on ne peut pas toutes les imaginer. Il ne faut cependant pas voir cet exemple comme cela : il permet plutôt d'offrir une variété de choix pour démontrer qu'une fonction n'admet *pas* de limite en un point, en essayant d'approcher ce point de plusieurs manières différentes (en pratique, on ne le fait pas aléatoirement, la forme précise de la fonction permet de déterminer les choix judicieux). Lorsque l'on veut montrer qu'une fonction *admet* une limite en un point A, on considère juste l'information minimale dont on dispose, à savoir que $x_1 \to a_1$ et $x_2 \to a_2$. Par exemple, considérons la fonction

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

On va montrer qu'elle admet une limite nulle en l'origine. Pour cela, on va recycler l'idée utilisée dans le paragraphe 1.1.3, Application B des développements limités : pour résoudre une forme indéterminée du type "0/0", on doit pouvoir dire lequel du numérateur ou du dénominateur tend le plus vite vers 0. Pour cela, il faut pouvoir comparer numérateur et dénominateur. Dans notre cas, on peut comparer le numérateur xy au dénominateur $\sqrt{x^2+y^2}$ de la façon suivante : pour tout $(x,y) \in \mathbb{R}^2$, on a $0 \leq x^2$ et $0 \leq y^2$ donc

$$|x| = \sqrt{x^2} \leq \sqrt{x^2+y^2}, \quad |y| = \sqrt{y^2} \leq \sqrt{x^2+y^2}.$$

Ainsi [3],

$$|xy| = |x| \times |y| \leq \sqrt{x^2+y^2} \times \sqrt{x^2+y^2} = x^2+y^2,$$

ce qui est la comparaison voulue : $|xy|$ est le numérateur (ou plutôt sa valeur absolue), et on a montré qu'il était plus petit que x^2+y^2 qui est le carré du dénominateur.

3. En fait, en remarquant que $(x+y)^2$ et $(x-y)^2$ sont tous les deux positifs (car ce sont des carrés) et en développant les carrés, on montre l'inégalité $2|xy| \leq x^2+y^2$, qui est « deux fois meilleure » que l'inégalité ci-dessus.

Comme le dénominateur tend vers 0, son carré tend beaucoup plus vite vers 0 que lui, et donc le numérateur tend vers 0 beaucoup plus vite que le dénominateur : c'est le numérateur qui "gagne" et la limite du quotient doit être 0. On peut formaliser cela de la façon suivante : on a montré que pour tout $(x, y) \in \mathbb{R}^2$,

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

De plus, $|f(x, y)|$ est toujours une quantité positive. Pour tout $(x, y) \in \mathbb{R}^2$, on a donc

$$0 \leq |f(x,y)| \leq \sqrt{x^2+y^2} = \operatorname{dist}\left((x,y);(0,0)\right).$$

Les deux côtés de l'inégalité tendent vers 0 lorsque $(x, y) \to (0, 0)$ donc par le théorème des gendarmes $|f(x, y)| \to 0$ lorsque $(x, y) \to (0, 0)$, ce qui est la même chose que de dire que $f(x, y) \to 0$ lorsque $(x, y) \to (0, 0)$.

Remarque 1.2.1 Comme on l'a déjà vu, une méthode utile pour étudier certaines formes indéterminées est d'introduire les coordonnées polaires

$$\begin{cases} x = r\cos\theta, \\ y = r\sin\theta, \end{cases}$$

avec $r = \sqrt{x^2+y^2} \geq 0$ et $\theta \in [0, 2\pi]$, qui sont utiles pour étudier les limites $(x, y) \to (0, 0)$. En effet, $(x, y) \to (0, 0)$ correspond à $r \to 0$: il n'y a donc plus qu'un seul paramètre qui tend vers 0. Le paramètre θ est cependant libre (et ne tend pas forcément vers 0), donc doit être traité avec attention. Voyons comment utiliser ces coordonnées dans les trois exemples que nous venons de voir.

Si

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

alors

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

Lorsque $r \to 0$, $r\cos\theta\sin\theta \to 0$ et ce indépendamment de la façon dont θ varie car $\cos\theta\sin\theta$ est bornée entre -1 et 1 : on dit que $f(r\cos\theta, r\sin\theta) \to 0$ lorsque $r \to 0$ *uniformément en* $\theta \in [0, 2\pi]$. Cela est suffisant pour déduire que $f(x, y) \to 0$ lorsque $(x, y) \to (0, 0)$.

Si l'on considère plutôt la fonction

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

alors

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

Lorsque $r \to 0$, on ne trouve pas de limite indépendamment de la façon dont θ varie : le comportement de $f(r\cos\theta, r\sin\theta)$ lorsque $r \to 0$ dépend fortement du comportement de θ.

Par exemple, si θ reste constant égal à 0, $f(r,0) = 0 \to 0$ lorsque $r \to 0$. Par contre, si θ reste constant égal à $\pi/4$ par exemple, alors $f(r/\sqrt{2}, r/\sqrt{2}) = 1/2 \to 1/2$ lorsque $r \to 0$. Ainsi, selon le comportement de θ, on trouve des limites différentes : la fonction f n'a donc pas de limite en $(0,0)$.

Dans le dernier exemple,

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

on a

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

si θ est tel que $\sin\theta \geq 0$ (donc $\theta \in [0,\pi]$). Ici, le numérateur $\sqrt{r}\cos\theta\sqrt{\sin\theta}$ tend vers 0 lorsque $r \to 0$ unformément en θ puisque $\sqrt{r} \to 0$ et $\cos\theta\sqrt{\sin\theta}$ reste borné. Il serait tentant de dire que le dénominateur $\sin\theta + r\cos\theta$ tend lui vers $\sin\theta$ lorsque $r \to 0$, et donc que le quotient tend vers 0 : on devrait avoir $f(x,y) \to 0$ lorsque $(x,y) \to (0,0)$. Or, on a vu que ce n'était pas le cas : il y a un problème quelque part. Ce problème se situe dans le fait que $\sin\theta + r\cos\theta \to \sin\theta$ lorsque $r \to 0$ uniquement lorsque θ est fixé ! Cependant, θ pourrait tout à fait varier en même temps que r : si $\theta \to 0$ lorsque $r \to 0$ par exemple, alors $\sin\theta + r\cos\theta \to 0$, et on a encore une forme indéterminée ! Ici, les coordonnées polaires n'ont pas levé la forme indéterminée, et ne sont d'aucun secours. On peut néanmoins s'en sortir en procédant à un autre changement de variables : posons

$$\begin{cases} x' = x, \\ y' = \sqrt{y}, \end{cases}$$

de sorte qu'on a $(x,y) \to (0,0) \iff (x',y') \to (0,0)$ et

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

On est ramené à une fonction de (x',y') que l'on peut étudier avec des coordonnées polaires.

Pour des fonctions de deux variables, les coordonnées polaires sont donc utiles dès que l'on peut écrire le dénominateur (après éventuellement un changement de variables) comme une fonction de $x^2 + y^2$.

1.2.2 Continuité

Une fois que la notion de limite est définie, on peut parler de continuité.

Définition 1.2.6: Continuité d'une fonction de 2 variables

Soit $f : \Omega \to \mathbb{R}$ une fonction de 2 variables et $A = (a_1, a_2) \in \Omega$. On dit que f est *continue au point* A si

$$\lim_{\substack{X \to A \\ X \neq A}} f(X) = f(A).$$

Si f est continue en tout point de Ω, on dit que f est *continue sur* Ω (ou continue tout court).

Comme pour les fonctions d'une variable, la somme, le produit, la multiplication par un scalaire, et la composition [4] sont des opérations qui donnent de nouvelles fonctions continues à partir de fonctions continues données.

Une autre façon importante de construire des fonctions continues de 2 variables est d'utiliser des fonctions continues à 1 variable :

Proposition 1.2.1

Soit $f : I \to \mathbb{R}$ une fonction continue à 1 variable ($I \subset \mathbb{R}$). Alors, les fonctions de 2 variables

$$f_1 : \left\{ \begin{array}{ccc} I \times \mathbb{R} & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f_1(x,y) = f(x) \end{array} \right. , \quad f_2 : \left\{ \begin{array}{ccc} \mathbb{R} \times I & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & f_2(x,y) = f(y) \end{array} \right.$$

sont continues sur leurs domaines de définition.

Exemple 1.2.7 Montrons que la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & x \sin(y) \end{array} \right.$$

est continue sur $\mathbb{R}^2$. Les fonctions

$$g : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & x \end{array} \right. , \quad h : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & \sin(y) \end{array} \right.$$

sont continues sur $\mathbb{R}$ (ce sont des fonction élémentaires), donc d'après la proposition précédénte, les fonctions

$$g_1 : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & g(x) \end{array} \right. , \quad h_2 : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & h(y) \end{array} \right.$$

sont continues sur $\mathbb{R}^2$. On a de plus $f = g_1 \times h_2$, donc f est également continue sur $\mathbb{R}^2$ en tant que produit de fonctions continues sur $\mathbb{R}^2$.

Exemple 1.2.8 Montrons que la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \left\{ \begin{array}{ll} \frac{xy}{\sqrt{x^2+y^2}} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{array} \right. \end{array} \right.$$

4. On fera attention à ce qu'on compose une fonction de 2 variables avec une fonction de 1 variable : si $f : \mathbb{R}^2 \to \mathbb{R}$ est une fonction de 2 variables, la composition $g \circ f$ a un sens si $g : \mathbb{R} \to \mathbb{R}$, donc si g est une fonction de 1 variable.

est continue sur $\mathbb{R}^2$. Ici, f s'exprime à l'aide de fonctions élémentaires uniquement en dehors de l'origine : on montrera donc la continuité en 2 étapes (i) en dehors de l'origine à l'aide d'opérations élémentaires et (ii) en l'origine en revenant à la définition de continuité. Commençons par montrer la continuité en dehors de l'origine. La fonction $x \mapsto x$ est continue sur $\mathbb{R}$ donc les fonctions $(x, y) \mapsto x$ et $(x, y) \mapsto y$ sont continues sur $\mathbb{R}^2$. Ainsi, la fonction $h : (x, y) \mapsto xy$ est continue sur $\mathbb{R}^2$ en tant que produit de fonctions continues sur $\mathbb{R}^2$. De même, la fonction $x \mapsto x^2$ est continue sur $\mathbb{R}$ donc les fonctions $(x, y) \mapsto x^2$ et $(x, y) \mapsto y^2$ sont continues sur $\mathbb{R}^2$. Ainsi, la fonction $(x, y) \mapsto x^2 + y^2$ est continue sur $\mathbb{R}^2$ en tant que somme de fonctions continues sur $\mathbb{R}^2$. De plus, cette fonction s'annule uniquement en l'origine, donc la fonction

$$g_1 : \left\{ \begin{array}{ccc} \mathbb{R}^2 \setminus \{(0,0)\} & \longrightarrow & \mathbb{R}^* \\ (x, y) & \longmapsto & x^2 + y^2 \end{array} \right.$$

est bien définie et continue. La fonction

$$g_2 : \left\{ \begin{array}{ccc} \mathbb{R}^* & \longrightarrow & \mathbb{R} \\ z & \longmapsto & \dfrac{1}{\sqrt{z}} \end{array} \right.$$

est continue en tant que fonction élémentaire, donc la composition $g_2 \circ g_1$ est bien définie, et continue sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Le produit $h \times g_2 \circ g_1$ est alors aussi continu sur $\mathbb{R}^2 \setminus \{(0, 0)\}$, et coïncide avec f sur cet ensemble. Ainsi, f est continue en dehors de l'origine. Montrons à présent la continuité en l'origine. On a vu dans l'Exemple 1.2.5 que

$$\lim_{(x,y)\to(0,0)} f(x, y) = 0,$$

et on a bien $f(0, 0) = 0$: f est bien continue en l'origine. En conclusion, f est continue sur $\mathbb{R}^2$.

1.3 Dérivées partielles

1.3.1 Définition

On introduit maintenant le concept de dérivées de fonctions de plusieurs variables, qui découle de la dérivée de fonction d'une variable. On vous prévient : ça aura l'air plus compliqué à cause des deux variables, mais on arrivera à faire pas mal de choses en fixant une variable (ou toutes les variables sauf une), en regardant f comme une fonction de la variable restante, et en utilisant les résultats valables pour les fonctions d'une seule variable. C'est l'idée des dérivées partielles ci-dessous.

Remarque sur les points de bord. Pour parler de dérivées de fonctions f d'une variable, on vous a demandé de vous placer sur un intervalle I, et on a seulement parlé de $f'(a)$ en un point intérieur de I, avec une extension, la notion de demi-dérivée à droite ou à gauche, dans le cas où a est une des extrémités de l'intervalle.

Ici c'est pareil : pour ne pas avoir d'ennui, le mieux est de supposer que l'ensemble de définition de f est un sous-ensemble *ouvert* Ω de $\mathbb{R}^n$. "Ouvert" signifie que pour tout point $A \in \Omega$, il existe une petite boule $B(A,r)$ centrée en A et qui est contenue dans Ω. En général, on dit que A est un *point intérieur* à Ω quand il existe $r > 0$ (qui peut être minuscule) tel que $B(A,r) \subset \Omega$. Donc Ω est ouvert si tous ses points sont des points intérieurs de Ω. Les exemples les plus simples sont $\Omega = \mathbb{R}^n$, ou $\Omega = B(X,r)$ (une boule ouverte), ou $\Omega =]a,b[\times]c,d[\subset \mathbb{R}^2$ (rectangle ouvert).

Mais on s'autorisera à utiliser les mêmes définitions (continuité, dérivées partielles, etc) même dans des cas un peu plus généraux. Par exemple, Ω pourra être la boule fermée $\overline{B}(X,r) = \left\{Y \in \mathbb{R}^n \,;\, ||Y - X|| \leq r\right\}$, et on s'autorisera à prendre A au bord, c'est-à-dire sur la sphère $S(X,r) = \left\{Y \in \mathbb{R}^n \,;\, ||Y - X|| = r\right\}$. L'essentiel de ce qu'on raconte ci-dessous restera vrai, avec des adaptations convenables, comme dans le cas où f est définie sur un intervalle I et on regarde ce qui se passe en une des extrémités de I.

Définition 1.3.1: Dérivées partielles

Prenons $\Omega \subset \mathbb{R}^2$ pour simplifier les notations.

Soit $f : \Omega \to \mathbb{R}$ et $A = (a_1, a_2) \in \Omega$.

— On dit que f est *dérivable par rapport à la* 1*-ère variable en* A si la fonction à une variable $f_1 : x_1 \mapsto f(x_1, a_2)$ est dérivable en $x_1 = a_1$. Dans ce cas, on note

$$\frac{\partial f}{\partial x_1}(A) = f_1'(a_1),$$

et on appelle cette quantité *la dérivée partielle par rapport à la* 1*-ère variable de* f *en* A.

— On dit que f est *dérivable par rapport à la* 2*-ème variable en* A si la fonction à une variable $f_2 : x_2 \mapsto f(a_1, x_2)$ est dérivable en a_2. Dans ce cas, on note

$$\frac{\partial f}{\partial x_2}(A) = f_2'(a_2).$$

On l'appelle *la dérivée partielle par rapport à la* 2*-ème variable de* f *en* A.

Notation Selon que l'on note une fonction de deux variables $f(x_1, x_2)$ ou $f(x,y)$, on peut noter les dérivées partielles $\frac{\partial f}{\partial x_1}$ ou $\frac{\partial f}{\partial x}$, respectivement $\frac{\partial f}{\partial x_2}$ ou $\frac{\partial f}{\partial y}$. Parfois on notera même ∂_1 au lieu de $\frac{\partial f}{\partial x_1}$. C'est plus court, et parfois c'est bien de ne pas dire le nom de la variable, juste que c'est la première variable de la fonction. On verra des exemples (typiquement quand on compose des fonctions) où cette notation est pratique.

Notez que quand A est un point intérieur de Ω, alors la fonction partielle $f_1(x) = f(x, a_2)$ est bien définie dans un petit intervalle centré en a_1, ce qui permet de parler de sa dérivée, et pareillement la seconde fonction partielle $f_2(y) = f(a_1, y)$ est bien définie dans un voisinage de a_2 et on peut penser à la dériver.

L'extension aux fonctions de 3 variables ou plus est immédiate. Par exemple, pour f définie sur $\Omega \subset \mathbb{R}^4$, la seconde dérivée partielle de f au point $A = (a_1, a_2, a_3, a_4)$, si elle existe, est la dérivée au point a_2 de la fonction à une variable $x \mapsto f(a_1, x, a_3, a_4)$, qui est définie au voisinage de a_2. Et on la note $\frac{\partial f}{\partial x_2}(A)$, ou carrément $\partial_2 f(A)$. Cette remarque vaut pour la suite : tout ce qui suit marche aussi bien quand $\Omega \subset \mathbb{R}^n$, $n \geq 3$.

Exemple 1.3.2 Soit la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & x^2 \sin(y) \end{array} \right. ,$$

et fixons $A = (a_1, a_2) \in \mathbb{R}^2$. On "gèle" alternativement les variables x et y pour définir les fonctions

$$f_1 : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \longmapsto & f(x, a_2) = x^2 \sin(a_2) \end{array} \right. , \quad f_2 : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ y & \longmapsto & f(a_1, y) = a_1^2 \sin(y) \end{array} \right.$$

qui sont dérivables (la première est un polynôme, la seconde un multiple de la fonction sinus. L'important est que les nombres a_1 et a_2 sont vus ici comme des constantes), avec pour tout $(x, y) \in \mathbb{R}^2$,

$$f_1'(x) = 2x \sin(a_2), \quad f_2'(y) = a_1 \cos(y).$$

Ainsi, f est dérivable par rapport aux deux variables en A, et on a

$$\frac{\partial f}{\partial x}(A) = 2a_1 \sin(a_2), \quad \frac{\partial f}{\partial y}(a) = a_1^2 \cos(a_2).$$

Comme la dérivation des fonctions de deux variables se ramène à la dérivation des fonctions d'une variable, la plupart des propriété de la dérivation persistent. Par exemple, la formule de Leibniz :

$$\frac{\partial (fg)}{\partial x_i}(X) = \frac{\partial f}{\partial x_i}(X) g(X) + f(X) \frac{\partial g}{\partial x_i}(X).$$

La seule exception viendra de la composition, parce qu'il faudra comme toujours faire attention aux domaines de définition et d'arrivée, qu'on expliquera dans la suite.

Remarque 1.3.1 Pour les fonctions d'une variable, dérivabilité en un point implique continuité en ce point. **Ce n'est malheureusement plus vrai pour les fonctions de deux variables** : la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & \begin{cases} \frac{xy}{x^2+y^2} & \text{si } (x,y) \neq (0,0), \\ 0 & \text{si } (x,y) = (0,0). \end{cases} \end{array} \right. \tag{1.18}$$

est dérivable par rapport aux deux variables en $A = (0, 0)$ puisque les fonctions $x \mapsto f(x, 0)$ et $y \mapsto f(0, y)$ sont identiquement nulles, donc dérivables et de dérivée nulle. Par contre, on a vu dans l'Exemple 1.2.3 que la fonction f n'est pas continue en $(0, 0)$.

Pour s'assurer qu'une fonction admettant des dérivées partielles soit continue, il nous faudra une propriété un peu plus forte, obtenue en demandant aussi que les dérivées partielles soient continues.

Définition 1.3.3: Fonction de classe C^1

Soit $f : \Omega \to \mathbb{R}$ une fonction. On dit que f est *de classe* C^1 *sur* Ω (ou C^1 tout court, et on écrira volontiers $f \in C^1(\Omega)$) si

1. f est dérivable par rapport aux 2 variables sur tout Ω ;
2. les application dérivées partielles $A \mapsto \frac{\partial f}{\partial x}(A)$ et $A \mapsto \frac{\partial f}{\partial y}(A)$ sont continues sur Ω.

Exercice : calculer les dérivées partielles en tout point $A \in \mathbb{R}^2$ de la fonction f définie en (1.18), et montrer que celles-ci ne sont pas continues en $A = (0, 0)$.

Remarque 1.3.2 On aime mieux C^1 parce qu'on peut montrer qu'une fonction C^1 est automatiquement continue, et en plus a automatiquement un développement limité d'ordre 1 au voisinage de chaque point (décrit ci-dessous). Contrairement, comme on l'a vu, au cas où on suppose seulement que f a des dérivées partielles en chaque point. Et ça ne coûtera pas bien cher de vérifier systématiquement la continuité des dérivées partielles.

Notation Si f est C^1 sur Ω, on appelle *gradient de* f l'application continue notée $\operatorname{grad} f$ ou ∇f :

$$\operatorname{grad} f : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R}^2 \\ A & \longmapsto & \left(\dfrac{\partial f}{\partial x}(A), \dfrac{\partial f}{\partial y}(A) \right) \end{array} \right. .$$

On fera bien attention au fait qu'il s'agit d'une application *vectorielle*, dans le sens où pour tout $A \in \Omega$, $\operatorname{grad} f(A)$ est un vecteur : $\operatorname{grad} f(A) \in \mathbb{R}^2$.

Comme dans le cas des fonctions d'une variable réelle, les dérivées partielles (maintenant continues !) permettent d'écrire un développement limité à l'ordre 1.

Proposition 1.3.1: Formule de Taylor à l'ordre 1 pour les fonctions de 2 variables de classe C^1

Soit $f : \Omega \to \mathbb{R}$ une fonction de classe C^1 et $A = (a_1, a_2) \in \Omega$ (un point intérieur). Alors, on a pour tout $X = (x_1, x_2) \in \Omega$,

$$f(X) = f(A) + \frac{\partial f}{\partial x_1}(A)(x_1 - a_1) + \frac{\partial f}{\partial x_2}(A)(x_2 - a_2) + \|X - A\|\epsilon(X), \tag{1.19}$$

où $\epsilon : \Omega \to \mathbb{R}$ est une fonction telle que

$$\lim_{X \to A} \epsilon(x) = 0.$$

On rappelle que $\|X - A\| = \sqrt{(x_1 - a_1)^2 + (x_2 - a_2)^2}$ désigne la *norme* du vecteur $X - A$.

Remarque 1.3.3 Comme dans le cas d'une variable réelle, la formule (1.19) s'interprète de deux manières : (i) elle donne une approximation de f avec des termes de plus en plus petits, en approchant f par une fonction polynomiale de degré 1 (une fonction affine) et (ii)

géométriquement, cela revient à approcher le graphe de f (qui est maintenant une surface dans $\mathbb{R}^3$) par un plan affine dans $\mathbb{R}^3$. Détaillons ces deux interprétations. D'abord, essayons de comprendre pourquoi les termes successifs du terme de droite de (1.19) sont de plus en plus petits lorsque $x \to a$. Auparavant, on va regrouper certains termes pour obtenir une écriture plus concise : on va écrire

$$\frac{\partial f}{\partial x_1}(A)(x_1 - a_1) + \frac{\partial f}{\partial x_2}(A)(x_2 - a_2) = \operatorname{grad} f(A) \cdot (X - A),$$

le *produit scalaire* des vecteurs $\operatorname{grad} f(A)$ et $X - A$ de $\mathbb{R}^2$ (voir le poly d'algèbre linéaire). Ce faisant, on a maintenant 3 termes dans le terme de droite de (1.19) :

$$f(x) = I + II + III,$$

avec

$$I = f(A), \quad II = \operatorname{grad} f(A) \cdot (X - A), \quad III = \|X - A\|\epsilon(X),$$

comme dans le cas des fonctions d'une variable. Il est alors clair que le terme II est beaucoup plus petit que le terme I lorsque $X \to A$ si $f(A) \neq 0$: en effet, le terme II tend vers 0 quand $X \to A$ alors que le terme I reste constant et non-nul. De même, le terme III est en général beaucoup plus petit que le terme II lorsque $X \to A$. Pour le voir, on se rappelle la formule

$$\operatorname{grad} f(A) \cdot (X - A) = \|X - A\| \| \operatorname{grad} f(A)\| \cos\theta,$$

où θ est l'angle entre les vecteurs $\operatorname{grad} f(A)$ et $X - A$. Ainsi, les deux termes II et III sont facteurs de $\|X - A\|$, avec en facteur $\| \operatorname{grad} f(A)\| \cos\theta$ pour le II et $\epsilon(X)$ pour III. Comme $\epsilon(X)$ tend vers 0 lorsque $X \to A$, il suffit alors que $\| \operatorname{grad} f(A)\| \cos\theta$ ne tende pas vers 0 pour que III soit beaucoup plus petit que II. C'est le cas si $\operatorname{grad} f(A) \neq (0, 0)$ et si $X - A$ n'est pas perpendiculaire à $\operatorname{grad} f(A)$ (il ne faut donc pas que X approche A de manière perpendiculaire à $\operatorname{grad} f(A)$). En général, III est donc beaucoup plus petit que II. Si l'on néglige le terme III, on a donc

$$f(X) \simeq f(A) + \operatorname{grad} f(A) \cdot (X - A)$$

lorsque $X \to A$. L'application

$$X \in \mathbb{R}^2 \mapsto f(A) + \operatorname{grad} f(A) \cdot (X - A)$$

est une application affine (c'est-à-dire de la forme $X \mapsto \alpha + \beta \cdot X$ avec $\alpha \in \mathbb{R}$ et $\beta \in \mathbb{R}^2$, ou de manière équivalente une application polynomiale de degré 1), donc (1.19) permet d'approcher f par une fonction "simple" au voisinage de A (comme pour le cas des fonctions d'une variable). Le graphe d'une application affine est un *plan affine dans* $\mathbb{R}^3$: dans notre cas il s'agit du *plan tangent* au graphe de f au point $(A, f(A))$, d'équation cartésienne

$$x_3 = f(A) + \frac{\partial f}{\partial x_1}(A)(x_1 - a_1) + \frac{\partial f}{\partial x_2}(A)(x_3 - a_2).$$

Ces calculs se généralisent en dimension $n \geq 3$, pour $\Omega \subset \mathbb{R}^n$ et $f : \Omega \to \mathbb{R}$. Dans ce cas, sous les mêmes hypothèses (existence de dérivées partielles continues), (1.19) devient

$$f(X) = f(A) + \sum_{i=0}^{n} \frac{\partial f}{\partial x_i}(A)(x_i - a_i) + ||X - A||\varepsilon(X), \tag{1.20}$$

ou de manière plus condensée

$$f(X) = f(A) + \nabla f(A) \cdot (X - A) + ||X - A||\varepsilon(X), \tag{1.21}$$

où A et X sont des points de Ω, et avec toujours $\lim_{X \to A} \varepsilon(X) = 0$.

Il est aussi d'usage (mais vous n'en aurez pas besoin) de donner un nom à l'existence du développement limité (1.19) (ou (1.20) ou (1.21)) : on dit que f est *différentiable au point* A. Et c'est la notion qui est utile, parce qu'elle permet de bien calculer, de composer, etc. On a donc montré qu'une fonction de classe $C^1(\Omega)$ est différentiable en tout point de Ω.

Exemple 1.3.4 Considérons la fonction f donnée par

$$f(x, y) = 1 + x^2 + y^3$$

pour tout $(x, y) \in \mathbb{R}^2$. Déterminons son plan tangent au point $(1, 1)$. L'application f est polynomiale donc de classe C^1 sur $\mathbb{R}^2$, et on a pour tout $(x, y) \in \mathbb{R}^2$,

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

et en particulier

$$\frac{\partial f}{\partial x}(1, 1) = 2, \quad \frac{\partial f}{\partial y}(1, 1) = 3.$$

Ainsi, l'équation cartésienne du plan tangent au graphe de f au point $(1, 1, f(1, 1)) = (1, 1, 3)$ est

$$z = f(1, 1) + \frac{\partial f}{\partial x}(1, 1)(x - 1) + \frac{\partial f}{\partial y}(1, 1)(y - 1) = 3 + 2 \times (x - 1) + 3 \times (y - 1) = 2x + 3y - 2,$$

où l'on a désigné les coordonnées cartésiennes de $\mathbb{R}^3$ par (x, y, z) plutôt que par (x_1, x_2, x_3). On l'a représenté graphiquement en Figure 1.11. On a également le développement limité de f au point $(1, 1)$ à l'ordre 1 en appliquant la Proposition 1.3.1 : pour tout $(x, y) \in \mathbb{R}^2$,

$$f(x, y) = 3 + 2 \times (x - 1) + 3 \times (y - 1) + \|(x - 1, y - 1)\|\epsilon(x, y),$$

pour une certaine fonction $\epsilon : \mathbb{R}^2 \to \mathbb{R}$ vérifiant

$$\lim_{(x,y) \to (1,1)} \epsilon(x, y) = 0.$$

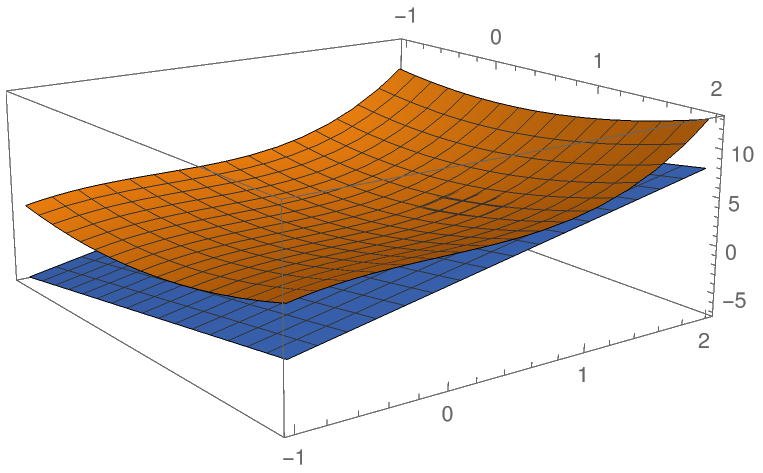

FIGURE 1.11 – Graphe de f (jaune) et son plan tangent en $(1, 1, 3)$ (bleu)

1.3.2 Généralisation aux applications vectorielles

On fait à présent un léger aparté sur l'extension des notions que l'on vient de voir aux fonctions *à valeurs vectorielles*, dont l'espace d'arrivée n'est pas $\mathbb{R}$ mais $\mathbb{R}^m$ pour un certain $m \geq 2$: pour tout $X \in \Omega$, $f(X)$ n'est pas un scalaire mais un vecteur de dimension m. Pour ce type de fonctions, on a une généralisation naturelle des notions déjà présentées, et tout se passe bien parce qu'on peut étudier chaque coordonnée de f séparément.

De telles fonctions apparaissent naturellement en physique : on peut penser au champ électrique ; au champ magnétique ; au champ de vitesses dans un fluide ; au champ de déplacements dans un milieu élastique.

ATTENTION ! Il faudra bien distinguer la dimension de l'espace de départ ($X \in \mathbb{R}^n$, ici on prendra $n = 2$ la plupart du temps), et celle de l'espace d'arrivée ($f(X) \in \mathbb{R}^m$).

Définition 1.3.5: Fonction vectorielle de classe C^1

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}^m$ pour un certain $m \geq 2$. Pour tout $(x, y) \in \Omega$, $f(x, y) \in \mathbb{R}^m$ et on note ses composantes $f_1(x, y), \ldots, f_m(x, y)$:

$$f(x, y) = (f_1(x, y), f_2(x, y), \ldots, f_m(x, y)).$$

On dit alors que f est de classe C^1 si chacune de ses fonctions composantes f_j (qui sont des fonctions scalaires) est de classe C^1.

Exemple 1.3.6 La fonction

$$f : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^3 \\ (x, y) & \longmapsto & (x^2, xy, y\sin(x)) \end{cases}$$

est de classe C^1 sur $\mathbb{R}^2$ puisque ses fonctions composantes

$$f_1 : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & x^2 \end{array} \right. , \quad f_2 : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & xy \end{array} \right. , \quad f_3 : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & y\sin(x) \end{array} \right.$$

sont de classe C^1.

Exercice : écrire le développement limité d'ordre 1 de f au point $A = (0,0)$. On calculera le DL de chaque composante avant de réunir les expressions.

Définition 1.3.7: Matrice jacobienne

Soit $f : \Omega \to \mathbb{R}^m$ de classe C^1. La matrice $2 \times m$

$$\begin{bmatrix} \dfrac{\partial f_1}{\partial x}(x,y) & \dfrac{\partial f_1}{\partial y}(x,y) \\ \dfrac{\partial f_2}{\partial x}(x,y) & \dfrac{\partial f_2}{\partial y}(x,y) \\ \vdots & \vdots \\ \dfrac{\partial f_m}{\partial x}(x,y) & \dfrac{\partial f_m}{\partial y}(x,y) \end{bmatrix} = \begin{bmatrix} {}^t\nabla f_1(x,y) \\ {}^t\nabla f_2(x,y) \\ \vdots \\ {}^t\nabla f_m(x,y) \end{bmatrix}$$

est appelée *matrice jacobienne* de f au point $(x,y) \in \Omega$. Elle est notée $\operatorname{Jac} f(x,y)$.

Ainsi, on a gardé l'écriture de la fonction à valeurs dans $\mathbb{R}^m$ comme une colonne de hauteur m, et on a mis côte-à-côte les deux vecteurs qui sont les deux dérivées partielles de f par rapport à x et y. Si f avait été une fonction vectorielle d'une seule variable, on aurait obtenu le vecteur colonne $f'(x)$. Et pour une fonction de n variables $X = (x_1, \ldots, x_n)$, $Jacf(X)$ est une matrice $n \times m$ composée de n vecteurs colonnes (dans $\mathbb{R}^m$), à savoir les n dérivées partielles de f.

C'est important de faire attention, parce que quand on composera deux fonctions, il faudra faire attention à ce que la matrice jacobienne de la composée soit le produit des deux matrices de dimensions compatibles.

Exemple 1.3.8 Pour la fonction f de l'Exemple 1.3.6, on a pour tout $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial f_1}{\partial x}(x,y) = 2x, \ \frac{\partial f_1}{\partial y}(x,y) = 0, \ \frac{\partial f_2}{\partial x}(x,y) = y, \tag{1.22}$$

$$\frac{\partial f_2}{\partial y}(x,y) = x, \ \frac{\partial f_3}{\partial x}(x,y) = y\cos(x), \ \frac{\partial f_3}{\partial y}(x,y) = \sin(x), \tag{1.23}$$

donc la matrice jacobienne de f en (x,y) vaut

$$\operatorname{Jac} f(x,y) = \begin{bmatrix} 2x & 0 \\ y & x \\ y\cos(x) & \sin(x) \end{bmatrix}.$$

On verra dans la prochaine section que les matrices jacobiennes sont utiles pour calculer la dérivée de la composée de deux fonctions.

1.3.3 Dérivée d'une composition

On explique à présent comment dériver la composée de deux fonctions de plusieurs variables. Pour ne pas surcharger la présentation, on va uniquement considérer dans cette partie des fonctions qui sont définies sur l'espace $\mathbb{R}^2$ tout entier. L'extension aux fonctions définies sur des sous-domaines Ω sera immédiate, à partir du moment où les compositions seront bien définies.

Jusqu'à présent, on avait majoritairement traité des fonctions du type $f : \mathbb{R}^2 \to \mathbb{R}$. Pour former la composition $g \circ f$, il faut donc que g soit une fonction d'une seule variable. La dérivation de $g \circ f$ fera donc intervenir à la fois des dérivées partielles, et des dérivées usuelles. On a le résultat suivant.

Proposition 1.3.2

Soient $f : \mathbb{R}^2 \to \mathbb{R}$ et $g : \mathbb{R} \to \mathbb{R}$ des fonctions de classe C^1. Alors, la composition $g \circ f : \mathbb{R}^2 \to \mathbb{R}$ est aussi de classe C^1, et on a pour tout $(x, y) \in \mathbb{R}^2$:

$$\frac{\partial(g \circ f)}{\partial x}(x, y) = g'(f(x, y))\frac{\partial f}{\partial x}(x, y), \quad \frac{\partial(g \circ f)}{\partial y}(x, y) = g'(f(x, y))\frac{\partial f}{\partial y}(x, y). \tag{1.24}$$

Remarque : un moyen mnémotechnique pour retrouver ces formules est de considérer, lorsqu'on écrit $g \circ f(x, y) = g(f(x, y))$, que la fonction g dépend de la "variable" f, et que cette variable f dépend elle-même de (x, y). De cette façon, on peut écrire **schématiquement** :

$$\frac{\partial(g \circ f)}{\partial x} = \frac{dg}{df}\frac{\partial f}{\partial x}, \tag{1.25}$$

avec sous-entendu que $\frac{dg}{df}$ est la dérivée de g (donc g') évaluée au point $f(x, y)$, tandis que la dérivée partielle $\frac{\partial f}{\partial x}$ est évaluée au point (x, y).
ATTENTION, l'écriture (1.25) est juste un moyen pour vous de retrouver la VRAIE formule (1.24), il ne faudra pas l'utiliser dans les devoirs.

Exemple 1.3.9 Si $f : \mathbb{R}^2 \to \mathbb{R}$ est de classe C^1 et h est la fonction

$$h : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & \sin(f(x, y)) \end{array} \right. ,$$

alors $h = g \circ f$ où $g : \mathbb{R} \to \mathbb{R}$ est la fonction sinus. De plus, pour tout $(x, y) \in \mathbb{R}^2$ on a

$$\frac{\partial h}{\partial x}(x, y) = \cos(f(x, y))\frac{\partial f}{\partial x}(x, y), \quad \frac{\partial h}{\partial y}(x, y) = \cos(f(x, y))\frac{\partial f}{\partial y}(x, y).$$

Dans le cas de dimensions quelconques, on a ce qu'on appelle la *règle de la chaîne*.

Proposition 1.3.3: Règle de la chaîne (ou règle de composition)

Soient

$$f : \left\{ \begin{array}{rcl} \mathbb{R}^p & \longrightarrow & \mathbb{R}^n \\ X = (x_1, \dots x_p) & \longmapsto & (f_1(X), \dots f_n(X)) \end{array} \right.$$

et

$$g : \left\{ \begin{array}{rcl} \mathbb{R}^n & \longrightarrow & \mathbb{R}^m \\ Y = (y_1, \dots y_n) & \longmapsto & (g_1(Y), \dots, g_m(Y)) \end{array} \right.$$

des fonctions de classe C^1. Alors $h = g \circ f : \mathbb{R}^p \to \mathbb{R}^m$ est de classe C^1.

Pour tout indice $j \in \{1, \dots, m\}$ de la composante de h, pour tout indice $\ell \in \{1, \dots, p\}$ de l'espace de départ, et pour tout $X = (x_1, \dots, x_p) \in \mathbb{R}^p$, on a la formule suivante pour les dérivées partielles de h en X :

$$\frac{\partial h_j}{\partial x_\ell}(X) = \frac{\partial g_j}{\partial y_1}(f(X))\frac{\partial f_1}{\partial x_\ell}(X) + \frac{\partial g_j}{\partial y_2}(f(X))\frac{\partial f_2}{\partial x_\ell}(X) + \cdots + \frac{\partial g_j}{\partial y_n}(f(X))\frac{\partial f_n}{\partial x_\ell}(X) \quad (1.26)$$

$$= \sum_{k=1}^{n} \frac{\partial g_j}{\partial y_k}(f(X))\frac{\partial f_k}{\partial x_\ell}(X). \quad (1.27)$$

Remarque 1.3.4 Cette terrifiante formule doit rappeler la formule définissant le produit matriciel vue dans le poly de Math 259. Ce n'est pas un hasard : la formule ci-dessus définit le (j, ℓ)-ième élément de la matrice jacobienne $\operatorname{Jac} h(X)$, tandis que $\frac{\partial g_j}{\partial y_k}$ et $\frac{\partial f_k}{\partial x_\ell}$ sont des éléments de $\operatorname{Jac} g$ et $\operatorname{Jac} f$. Elle montre que la matrice jacobienne de h est égale au produit matriciel de la matrice jacobienne de g par celle de f, évaluées aux points respectifs $Y = f(X)$ et X :

$$\begin{bmatrix} {}^t\nabla h_1(X) \\ {}^t\nabla h_2(X) \\ \vdots \\ {}^t\nabla h_m(X) \end{bmatrix} = \operatorname{Jac} h(X) = \operatorname{Jac} g(f(X)) \ \operatorname{Jac} f(X). \quad (1.28)$$

Quand on regarde bien, on voit qu'il suffit d'établir la formule quand g est à valeurs dans $\mathbb{R}$ ($m = 1$), puisqu'ensuite on peut écrire que les coordonnées de $h = g \circ f$ sont les $h_j = g_j \circ f$. Et donc se contenter de calculer les dérivées partielles des $g_j \circ f$ et les rassembler dans un vecteur colonne.

De même, on peut se contenter du cas où f est définie sur $\mathbb{R}$ (et pas sur $\mathbb{R}^p$, parce que de toute manière quand on calcule une dérivée partielle $\frac{\partial g \circ f}{\partial x_\ell}$, on ne regarde que la fonction partielle $t \mapsto g \circ f(x_1, \dots, t, x_{\ell+1}, \dots x_n)$ d'une seule variable.

Esquisse de preuve : Pour retrouver cette formule, on peut utiliser le fait que g_j admet un développement limité d'ordre 1 au voisinage de $Y = f(X)$, puisque g est C^1 :

$$g_j(Y') = g_j(Y) + \nabla g_j(Y) \cdot (Y' - Y) + ||Y' - Y|| \varepsilon'_\ell(Y'), \quad (1.29)$$

avec $\lim_{Y' \to Y} \varepsilon'_\ell(Y') = 0$. Maintenant, si on prend $Y' = f(X')$, cette expression devient

$$h_j(X') = h_j(X) + \nabla g_j(f(X)) \cdot (f(X') - f(X)) + ||f(X') - f(X)||\varepsilon'_j(f(X')) \tag{1.30}$$

$$= h_j(X) + \sum_{k=1}^{n} \frac{\partial g_j}{\partial y_k}(f(X))(f_k(X') - f_k(X)) + ||f(X') - f(X)||\varepsilon'_j(f(X')) \tag{1.31}$$

Pour traiter les termes de la somme, on utilise le DL de f_k au voisinage de X,

$$f_k(X') = f_k(X) + \nabla f_k(X) \cdot (X' - X) + ||X' - X||\varepsilon_k(X'), \tag{1.32}$$

avec $\lim_{X' \to X} \varepsilon_k(X') = 0$. Lorsqu'on remplace cette expression dans (1.31), on obtient

$$h_j(X') = h_j(X) + \sum_{k=1}^{n} \frac{\partial g_j}{\partial y_k}(f(X))\nabla f_k(X) \cdot (X' - X) + reste, \tag{1.33}$$

$$= h_j(X) + \sum_{k=1}^{n}\sum_{\ell=1}^{p} \frac{\partial g_j}{\partial y_k}(f(X))\frac{\partial f_k}{\partial x_\ell}(X)(x'_\ell - x_\ell) + reste \tag{1.34}$$

où le reste contient tous les termes contenant un ε'_j ou un ε_k. En travaillant un peu, on montre que ce reste est de la forme $\|X' - X\|\tilde{\varepsilon}_j(X')$, donc plus petit que le second terme. En identifiant le facteur devant chaque $(x'_\ell - x_\ell)$ dans le second terme, on retrouve la formule de la proposition.

Exemple 1.3.10 Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe C^1, et définissons

$$h : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x, y) & \longmapsto & g(xy, x^2 + y^3) \end{cases}.$$

Montrons que h est C^1, et calculons ses dérivées partielles en fonction de celles de g. En posant

$$f : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R}^2 \\ (x, y) & \longmapsto & (xy, x^2 + y^3) = (f_1(x, y), f_2(x, y)) \end{cases},$$

on voit que la fonction f est C^1 car ses fonctions composantes f_1 et f_2 le sont (elles sont polynomiales). De plus, on a $h = g \circ f$ donc h est également de classe C^1, et on utilise la Proposition 1.3.3 pour déterminer ses dérivées partielles en fonction de celles de g. Par rapport à l'énoncé de la proposition, on a ici $p = n = 2$ et $m = 1$.

Pour distinguer les variables de l'espace de départ de f de l'espace d'arrivée de f (qui est le même que l'espace de départ de g), on va noter $(x, y) \in \mathbb{R}^2$ pour les variables de l'espace de départ de f et $(u, v) \in \mathbb{R}^2$ les variables de l'espace d'arrivée de f, qui est l'espace de départ de g ; la fonction h peut donc s'écrire $g(u, v)$ pour $(u, v) = (f_1(x, y), f_2(x, y))$. On a ainsi pour tout $(x, y) \in \mathbb{R}^2$,

$$\nabla h(x, y) = \begin{cases} \dfrac{\partial h}{\partial x}(x, y) = \dfrac{\partial g}{\partial u}(f(x, y))\dfrac{\partial f_1}{\partial x}(x, y) + \dfrac{\partial g}{\partial v}(f(x, y))\dfrac{\partial f_2}{\partial x}(x, y), \\ \dfrac{\partial h}{\partial y}(x, y) = \dfrac{\partial g}{\partial u}(f(x, y))\dfrac{\partial f_1}{\partial y}(x, y) + \dfrac{\partial g}{\partial v}(f(x, y))\dfrac{\partial f_2}{\partial y}(x, y), \end{cases}$$

en sachant également que

$$\frac{\partial f_1}{\partial x}(x,y) = y, \quad \frac{\partial f_1}{\partial y}(x,y) = x, \quad \frac{\partial f_2}{\partial x}(x,y) = 2x, \quad \frac{\partial f_2}{\partial y}(x,y) = 3y^2.$$

Ainsi, pour tout $(x,y) \in \mathbb{R}^2$, on trouve que

$$\begin{cases} \dfrac{\partial h}{\partial x}(x,y) = y\dfrac{\partial g}{\partial u}(xy, x^2+y^3) + x\dfrac{\partial g}{\partial v}(xy, x^2+y^3), \\ \\ \dfrac{\partial h}{\partial y}(x,y) = 2x\dfrac{\partial g}{\partial u}(xy, x^2+y^3) + 3y^2\dfrac{\partial g}{\partial v}(xy, x^2+y^3). \end{cases}$$

Exemple 1.3.11 Soit $g : \mathbb{R}^2 \to \mathbb{R}$ une fonction de classe C^1. Définissons

$$h : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ t & \longmapsto & g(t+e^t, t^2-t) \end{array} \right. .$$

Montrons que h est dérivable et déterminons sa fonction dérivée en fonction des dérivées partielles de g. On introduit la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R}^2 \\ t & \longmapsto & (t+e^t, t^2-t) = (f_1(t), f_2(t)) \end{array} \right.$$

qui est de classe C^1 car ses fonctions composantes f_1, f_2 le sont (on peut les obtenir à l'aide d'opérations élémentaires sur des fonctions usuelles). On a de plus pour tout $t \in \mathbb{R}$,

$$f'(t) = (1+e^t, 2t-1).$$

Ainsi, $h = g \circ f$ est aussi de classe C^1 en tant que composition de fonctions de classe C^1, et on a pour $t \in \mathbb{R}$,

$$\begin{aligned} h'(t) &= \frac{\partial g}{\partial x}(f(t))f_1'(t) + \frac{\partial g}{\partial y}(f(t))f_2'(t) \\ &= (1+e^t)\frac{\partial g}{\partial x}(t+e^t, t^2-t) + (2t-1)\frac{\partial g}{\partial y}(t+e^t, t^2-t). \end{aligned}$$

Exemple 1.3.12 Enfin, le dernier exemple important est celui des coordonnées polaires et du calcul du gradient d'une fonction dans ces coordonnées. On considère donc une fonction $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^1, et on introduit la fonction

$$\tilde{f} : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times]0, 2\pi[& \longrightarrow & \mathbb{R} \\ (r, \theta) & \longmapsto & f(r\cos\theta, r\sin\theta) \end{array} \right.$$

qui représente "la fonction f en coordonnées polaires". Les dérivées partielles $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ représentent les dérivées partielles de f "en coordonnées cartésiennes" et les dérivées partielles $\frac{\partial \tilde{f}}{\partial r}$

et $\frac{\partial \tilde{f}}{\partial \theta}$ représentent "les dérivées partielles de f en coordonnées polaires". Les deux notions sont reliées par la Proposition 1.3.3. En introduisant la fonction

$$\varphi : \left\{ \begin{array}{ccc} \mathbb{R}_+^* \times]0, 2\pi[& \longrightarrow & \mathbb{R}^2 \\ (r, \theta) & \longmapsto & (r\cos\theta, r\sin\theta) = (\varphi_1(r,\theta), \varphi_2(r,\theta)) \end{array} \right. ,$$

on a en effet $\tilde{f} = f \circ \varphi$. Comme φ est C^1 car ses fonctions composantes φ_1 et φ_2 le sont (elles s'obtiennent à partir d'opérations élémentaires sur des fonctions usuelles), la fonction $\tilde{f}$ est C^1 en tant que composée de fonctions C^1. De plus, par la Proposition 1.3.3 on a pour tout $r > 0$ et tout $0 < \theta < 2\pi$,

$$\begin{aligned} \frac{\partial \tilde{f}}{\partial r}(r,\theta) &= \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial \varphi_1}{\partial r}(r,\theta) + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial \varphi_2}{\partial r}(r,\theta) \\ &= \cos\theta \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + \sin\theta \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \end{aligned}$$

ainsi que

$$\begin{aligned} \frac{\partial \tilde{f}}{\partial \theta}(r,\theta) &= \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta)\frac{\partial \varphi_1}{\partial \theta}(r,\theta) + \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta)\frac{\partial \varphi_2}{\partial \theta}(r,\theta) \\ &= -r\sin\theta \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) + r\cos\theta \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta). \end{aligned}$$

Ces relations peuvent d'écrire matriciellement

$$\begin{bmatrix} \frac{\partial \tilde{f}}{\partial r}(r,\theta) & \frac{\partial \tilde{f}}{\partial \theta}(r,\theta) \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) & \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \end{bmatrix} \begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix},$$

la dernière matrice apparaissant n'étant rien d'autre que la jacobienne de φ, conformément à la formule 1.28. En particulier, on peut inverser la relation en inversant la matrice : comme

$$\begin{bmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{bmatrix}^{-1} = \frac{1}{r}\begin{bmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$

on en déduit que

$$\begin{bmatrix} \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) & \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) \end{bmatrix} = \frac{1}{r}\begin{bmatrix} \frac{\partial \tilde{f}}{\partial r}(r,\theta) & \frac{\partial \tilde{f}}{\partial \theta}(r,\theta) \end{bmatrix} \begin{bmatrix} r\cos\theta & r\sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$

et donc que

$$\left\{ \begin{array}{ll} \frac{\partial f}{\partial x}(r\cos\theta, r\sin\theta) & = \cos\theta \frac{\partial \tilde{f}}{\partial r}(r,\theta) - \frac{\sin\theta}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta), \\ \frac{\partial f}{\partial y}(r\cos\theta, r\sin\theta) & = \sin\theta \frac{\partial \tilde{f}}{\partial r}(r,\theta) + \frac{\cos\theta}{r}\frac{\partial \tilde{f}}{\partial \theta}(r,\theta). \end{array} \right.$$

On peut voir ces dernières relations comme l'expression du gradient de f en coordonnées cartésiennes en fonction du gradient de f en coordonnées polaires.

1.3.4 Dérivées d'ordre supérieur

Comme pour les fonctions d'une variable, si les dérivées partielles d'une fonction sont elles-même dérivables par rapport à l'une ou l'autre des variables, on peut former la dérivée partielle de la dérivée partielle, et ainsi de suite. A chaque étape, on a plusieurs choix : on peut d'abord dériver par rapport à x puis par rapport à y, ou d'abord par rapport à y puis par rapport à x (ou même deux fois de suite par rapport à x, et deux fois de suite par rapport à y). Quand on dérive encore plus de fois, on voit qu'on obtient un nombre considérable de dérivées partielles successives possibles. En fait, le théorème de Schwarz montre qu'il y en a moins que prévu, comme on l'illustre ici sur un exemple. Considérons la fonction

$$f : \left\{ \begin{array}{ccc} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ (x,y) & \longmapsto & x\sin(x+y) \end{array} \right. .$$

Elle est manifestement C^1 (vérifier que l'on sait le montrer rigoureusement en la décomposant en fonctions usuelles), et on a pour tout $(x,y) \in \mathbb{R}^2$,

$$\frac{\partial f}{\partial x}(x,y) = \sin(x+y) + x\cos(x+y), \quad \frac{\partial f}{\partial y}(x,y) = x\cos(x+y).$$

Les fonctions $\frac{\partial f}{\partial x}$ et $\frac{\partial f}{\partial y}$ sont aussi manifestement C^1, et on a pour tout $(x,y) \in \mathbb{R}^2$,

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

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

Ces dérivées successives sont notées

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

On les appelles *dérivées partielles d'ordre* 2 de f. De la même manière, on a des dérivées partielles d'ordre k lorsque l'on dérive k fois f, successivement par rapport à toutes les combinaisons de variables possibles.

Définition 1.3.13: Fonction C^k à plusieurs variables

Soit $f : \Omega \to \mathbb{R}$ une fonction de 2 variables. On dit que f est C^k *sur* Ω si f admet des dérivées partielles d'ordre k et si ces dérivées partielles sont continues sur Ω. Si f est C^k pour tout k, on dit que f est C^∞ sur Ω.

Remarque 1.3.5 Les fonctions usuelles sont de classe C^∞ sur leur domaine de définition (en faisant attention au cas des fonctions du type $\sqrt{x}$ comme dans le cas d'une variable). De plus, on peut additionner, multiplier, composer, multiplier par un scalaire des fonctions de classe C^k pour obtenir d'autres fonctions de classe C^k.

L'exemple considéré ci-dessus possède la propriété que $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$, autrement dit on obtient la même fonction en dérivant d'abord par rapport à x et ensuite par rapport à y ou l'inverse. Il s'agit en fait d'une propriété générale.

Proposition 1.3.4: Théorème de Schwarz

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction de classe C^2. Alors on a en tout point $X \in \mathbb{R}^2$:

$$\frac{\partial^2 f}{\partial x \partial y}(X) = \frac{\partial^2 f}{\partial y \partial x}(X).$$

Remarque 1.3.6 Cette relation reste vraie en dimension supérieure : on peut permuter deux coordonnées et obtenir les même dérivées partielles. Par exemple en 3D, si $f : \mathbb{R}^3 \to \mathbb{R}$ est C^2, on a en tout point

$$\frac{\partial^2 f}{\partial x \partial z}(X) = \frac{\partial^2 f}{\partial z \partial x}(X).$$

GD : La démonstration n'est pas fondamentalement difficile, mais on la passe. Pour un monôme $(x, y) \mapsto ax^k y^\ell$, la formule est facile à vérifier. Puis on peut en déduire qu'elle est vraie pour tout polynôme (somme de monômes). Le cas général de f de classe C^2 s'obtient par exemple en utilisant le théorème des accroissements finis, appliqué en utilisant successivement deux dérivées partielles, pour trouver deux développements limités de la même quantité, puis dire que les coefficients doivent être égaux. C'est un peu technique alors on passe.

Notation Cette propriété se propage aux dérivées d'ordre supérieur, par exemple pour une fonction $f : \mathbb{R}^3 \to \mathbb{R}$ de classe C^3, on a

$$\frac{\partial^3 f}{\partial y \partial z \partial x} = \frac{\partial^3 f}{\partial x \partial y \partial z}.$$

En particulier, cela motive la notation suivante dans le cas général d'un nombre quelconque de variables : pour une fonction $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^k, les dérivées partielles d'ordre k sont les dérivées partielles de la forme

$$\frac{\partial^k f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \cdots \partial x_n^{\alpha_n}}$$

avec $\alpha_1 + \alpha_2 + \cdots + \alpha_n = n$. Ici, on a dérivé α_1 fois par rapport à x_1, α_2 fois par rapport à x_2, etc. jusqu'à α_n fois par rapport à x_n, et ce dans n'importe quel ordre.

Définition 1.3.14

Soit $\Omega \subset \mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction C^2. Pour $(x, y) \in \Omega$, on appelle *matrice hessienne*

de f au point (x,y) la matrice

$$\operatorname{Hess} f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(x,y) & \frac{\partial^2 f}{\partial x \partial y}(x,y) \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) & \frac{\partial^2 f}{\partial y^2}(x,y) \end{bmatrix}.$$

Remarque 1.3.7 Par le théorème de Schwarz, la matrice hessienne en tout point est une matrice symétrique.

Remarque 1.3.8 En dimension n quelconque, la matrice hessienne d'une fonction $f : \mathbb{R}^n \to \mathbb{R}$ de classe C^2 au point $X = (x_1, \ldots, x_n) \in \mathbb{R}^n$ est la matrice

$$\left[\frac{\partial^2 f}{\partial x_j \partial x_k}(X)\right]_{\substack{1 \le j \le n \\ 1 \le k \le n}}.$$

Il s'agit d'une matrice carrée symétrique de taille $n \times n$.

De la même manière que pour une seule variable, les dérivées partielles d'ordre 2 permettent d'obtenir un développement limité à l'ordre 2.

Proposition 1.3.5: Formule de Taylor à l'ordre 2 pour des fonctions de 2 variables

Soit $\Omega \subset \mathbb{R}^2$, $f : \Omega \to \mathbb{R}$ une fonction C^2 et $A \in \Omega$. Alors, on a le développement limité à l'ordre 2 suivant : pour tout $X \in \Omega$,

$$f(X) = f(A) + \operatorname{grad} f(A) \cdot (X-A) + \frac{1}{2}(X-A) \cdot \operatorname{Hess} f(A)(X-A) + \|X-A\|^2 \epsilon(X), \quad (1.35)$$

où $\epsilon : \Omega \to \mathbb{R}$ est une fonction satisfaisant

$$\lim_{X \to A} \epsilon(X) = 0.$$

Remarque 1.3.9 Par rapport à la formule de Taylor à l'ordre 1 donnée à la Proposition 1.3.1, il y a le nouveau terme

$$\frac{1}{2}(X-A) \cdot \operatorname{Hess} f(A)(X-A)$$

dans lequel $\operatorname{Hess} f(A)(X-A)$ désigne le produit matriciel entre la matrice carrée $\operatorname{Hess} f(A)$ de taille 2×2 et $X - A$ qu'on identifie ici à un vecteur colonne de taille 2. Le produit $\operatorname{Hess} f(A)(X-A)$ s'identifie donc à un vecteur colonne de taille 2, dont on prend ensuite le produit scalaire avec le vecteur $X - A$ pour obtenir $(X-A) \cdot \operatorname{Hess} f(A)(X-A)$.

Plus explicitement, si on note (y_1, y_2) les coordonnées du vecteur $X - A$, et les dérivées secondes de f

$$b = \frac{\partial^2 f}{\partial x^2}(A), \quad c = \frac{\partial^2 f}{\partial x \partial y}(A), \quad d = \frac{\partial^2 f}{\partial y^2}(A),$$

alors ce terme prend la forme

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \cdot \begin{bmatrix} b & c \\ c & d \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \cdot \begin{bmatrix} by_1 + cy_2 \\ cy_1 + dy_2 \end{bmatrix} = y_1(by_1 + cy_2) + y_2(cy_1 + dy_2) = by_1^2 + 2cy_1y_2 + dy_2^2.$$

Remarque 1.3.10 Comme pour les fonctions d'une variable, les termes successifs de la formule de Taylor à l'ordre 2 sont de plus en plus petits lorsque $X \to A$, on interprète donc encore cette formule comme une description de plus en plus précise du comportement de f au voisinage de A. Là encore, cette description est donnée à partir de fonctions simples, ici le terme dominant correspond à la fonction

$$X \mapsto f(A) + \operatorname{grad} f(A) \cdot (X - A) + \frac{1}{2}(X - A) \cdot \operatorname{Hess} f(A)(X - A)$$

qui est polynomiale de degré 2 en les composantes (x, y) de X, alors que la formule de Taylor à l'ordre 1 est polynomiale de degré 1 (c'est-à-dire affine). Le graphe d'un tel polynôme de degré 2 peut avoir différentes formes selon ses coefficients. On en donne plusieurs exemples en Figure 1.12

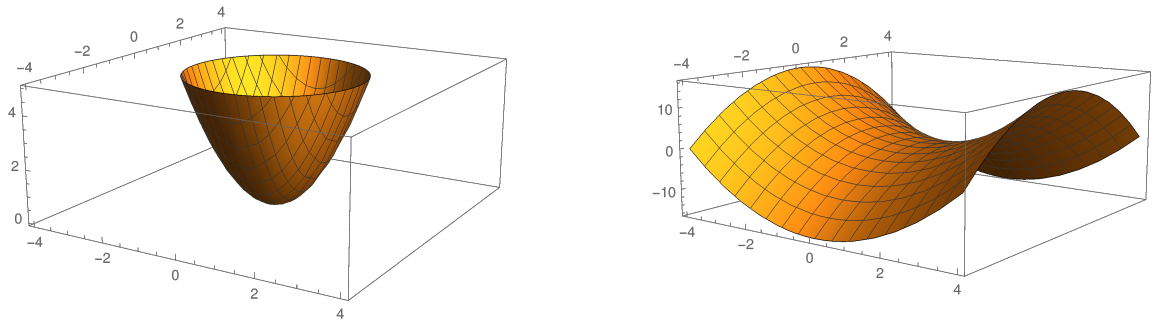

FIGURE 1.12 – Graphes de $(x, y) \mapsto x^2 + y^2$ (paraboloïde elliptique) et de $(x, y) \mapsto x^2 - y^2$ (paraboloïde hyperbolique).

En particulier, la formule de Taylor à l'ordre 2 signifie que l'on approche le graphe de f au voisinage de $(A, f(A))$ par un graphe de polynôme de degré 2, comme illustré en Figure 1.13.

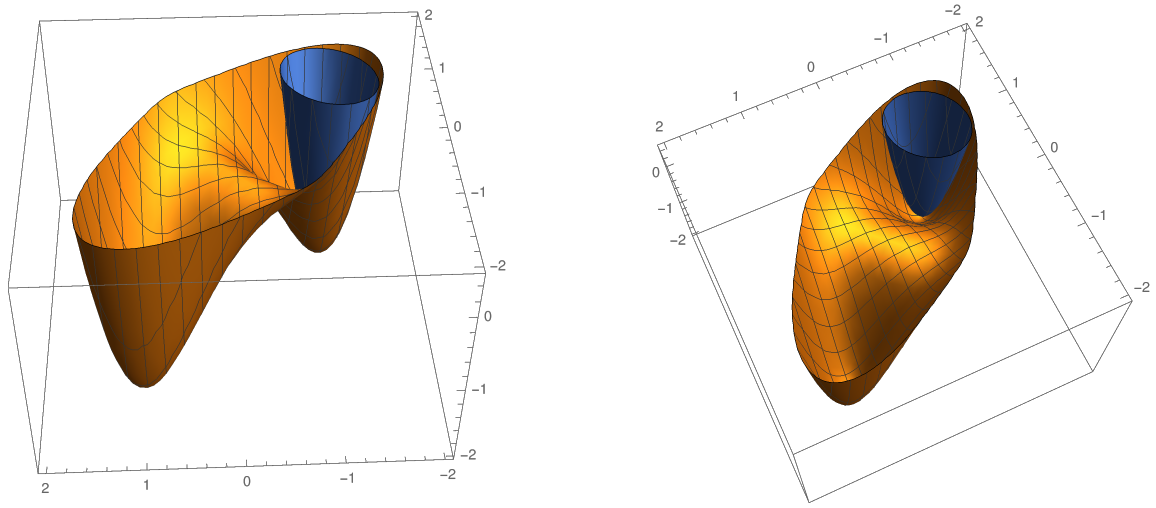

FIGURE 1.13 – Graphe de f (jaune) approché par un graphe de polynôme de degré 2 (bleu)

Pour comprendre pourquoi les termes successifs de la formule (1.35) sont de plus en plus petits quand $X \to A$, on peut écrire X sous la forme $X = A + tV$, où $V \in \mathbb{R}^2$ est un vecteur fixe et $t \in \mathbb{R}$ est un paramètre que l'on fait varier. La limite $X \to A$ correspond donc à $t \to 0$. La formule de Taylor nous donne

$$f(A+tV) = f(A) + t \operatorname{grad} f(A) \cdot V + \frac{t^2}{2} V \cdot \operatorname{Hess} f(A) V + t^2 \|v\|^2 \epsilon(A+tV),$$

et on voit apparaître des puissances successives de t. Comme $t^2 \ll t \ll 1 = t^0$ quand $t \to 0$, on voit que les trois premiers sont ordonnés (en supposant que les facteurs des t^k sont tous non nuls). Le dernier terme est encore plus petit que les autres, car $\epsilon(A+tV) \to 0$ quand $t \to 0$, donc $\epsilon(A+tV)t^2 \ll t^2$ lorsque $t \to 0$.

Dans le cas des fonctions d'une variable, on avait appliqué la formule de Taylor à l'ordre 2 pour savoir si le graphe d'une fonction était localement au dessus ou en dessous de sa tangente. Pour cela, on avait utilisé le fait que le terme $\frac{1}{2}f''(a)(x-a)^2$ a un signe (négatif ou positif selon le signe de $f''(a)$). Dans le cas des fonctions de plusieurs variables, le terme correspondant est

$$\frac{1}{2}(X-A) \cdot \operatorname{Hess} f(A)(X-A)$$

et il n'est pas forcément évident de savoir quand ce terme a un signe ou pas. Dans la section suivante nous allons voir comment déterminer ce signe sur un problème relié : l'étude des extrema locaux de f.

1.4 Application à l'étude des extrema locaux

Définition 1.4.1: Extremum local

Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction. On dit que f possède un *maximum local en* $x_0 \in \mathbb{R}^n$ si $f(x) \leq f(x_0)$ pour tout x suffisamment proche de x_0, autrement dit, s'il existe un voisinage U_{x_0} de x_0 tel que $f(x) \leq f(x_0)$ pour tout x dans ce voisinage.

Si on a de plus $f(x) < f(x_0)$ pour tout x suffisamment proche de x_0 mais distinct de x_0 (par exemple, pour $x \in U_{x_0} \setminus \{x_0\}$), on dit que x_0 est un maximum local *strict* de f.

Si on a plutôt $f(x) \geq f(x_0)$ (resp. $f(x) > f(x_0)$) dans la définition précédente, on parle de *minimum local* (resp. minimum local *strict*). Un extremum local est un maximum ou un minimum local.

Remarque 1.4.1 Le mot *local* est ici important : cela signifie (pour un maximum par exemple) que l'inégalité $f(x) \leq f(x_0)$ n'est pas nécessairement vraie pour tous les $x \in \mathbb{R}$ (auquel cas on parle de maximum *global*), mais seulement pour des x dans un voisinage assez petit de x_0. Par exemple, pour la fonction

$$f(x) = x^2(1-x)^2,$$

le point $x = 1/2$ est un maximum local strict (on pourra s'en convaincre en traçant son graphe), mais pas un maximum global : l'inégalité $f(x) \leq f(1/2)$ n'est vrai que pour $1/2 - \sqrt{2} \leq x \leq 1/2 + \sqrt{2}$, comme le montre la Figure 1.14.

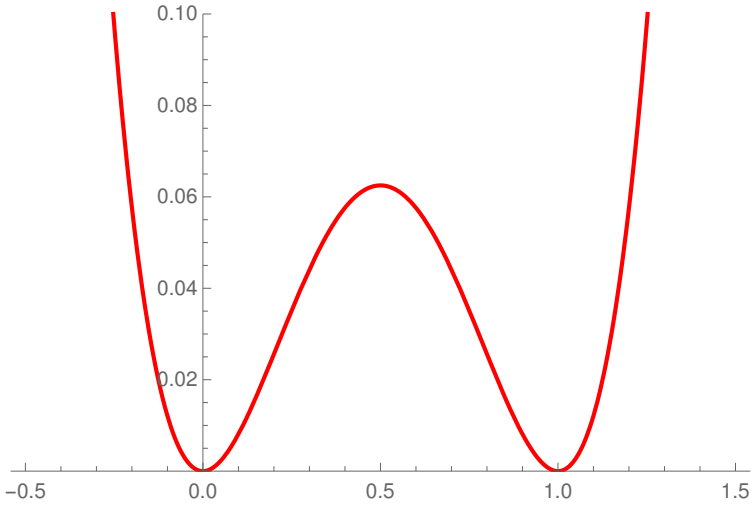

FIGURE 1.14 – Graphe de $x \mapsto x^2(1-x)^2$

1.4.1 Cas des fonctions d'une variable

Proposition 1.4.1: La dérivée s'annule en un extremum local

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction. Si x_0 est un extremum local de f et que f est dérivable en x_0, alors $f'(x_0) = 0$.

Démonstration. Cette propriété est évidente si l'on esquisse le graphe de f au voisinage de x_0. Pour la démontrer, supposons que x_0 est un minimum local par exemple. On se rappelle que

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

Comme x_0 est un minimum local, on a

$$f(x) - f(x_0) \geq 0$$

pour tout x assez proche de x_0. Si maintenant x est proche de x_0 avec $x > x_0$, on a donc

$$\frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

et ainsi

$$\lim_{\substack{x \to x_0 \\ x > x_0}} \frac{f(x) - f(x_0)}{x - x_0} \geq 0,$$

donc $f'(x_0) \geq 0$. De la même manière, si x est proche de x_0 mais avec $x < x_0$, on a alors

$$\frac{f(x) - f(x_0)}{x - x_0} \leq 0$$

et ainsi

$$\lim_{\substack{x \to x_0 \\ x > x_0}} \frac{f(x) - f(x_0)}{x - x_0} \leq 0,$$

donc $f'(x_0) \leq 0$. Ainsi, $f'(x_0)$ est à la fois positif et négatif, donc $f'(x_0) = 0$. La preuve fonctionne de la même façon lorsque x_0 est un maximum local. □

On voit donc qu'en un extremum local, on a une information sur la dérivée de la fonction. On a même mieux que ça : on a aussi une information sur la dérivée seconde.

Proposition 1.4.2: Dérivée seconde en un extremum local

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction deux fois dérivable en $x_0 \in \mathbb{R}$.

1. Si f admet un maximum local en x_0, alors $f''(x_0) \leq 0$.
2. Si f admet un minimum local en x_0, alors $f''(x_0) \geq 0$.

Démonstration. On esquisse juste le principe de la preuve. Par la formule de Taylor à l'ordre 2, on a pour tout $x \in \mathbb{R}$,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2\epsilon(x),$$

avec $\epsilon(x) \to 0$ lorsque $x \to x_0$. Comme x_0 est un extremum local, on vient de voir que $f'(x_0) = 0$. Le développement limité se simplifie donc en

$$f(x) = f(x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2\epsilon(x).$$

Supposons maintenant que f admet un maximum local en x_0. Alors $f(x) \leq f(x_0)$ pour tout x assez proche de x_0. Par le développement limité, on a donc

$$\frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2\epsilon(x) \leq 0$$

pour tout x assez proche de x_0. Si jamais $f''(x_0) > 0$, on aurait

$$\frac{1}{2}f''(x_0)(x - x_0)^2 > 0$$

pour x proche de x_0, $x \neq 0$. Comme le terme $(x - x_0)^2\epsilon(x)$ est beaucoup plus petit que $\frac{1}{2}f''(x_0)(x - x_0)^2$ quand $x \to x_0$, on aurait toujours

$$\frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2\epsilon(x) = \big(\frac{1}{2}f''(x_0) + \epsilon(x)\big)(x - x_0)^2 > 0$$

pour x proche de x_0, $x \neq x_0$, ce qui est absurde. Ainsi, on a bien $f''(x_0) \leq 0$. Par le même raisonnement, on montre que $f''(x_0) \geq 0$ en un minimum local. □

On a un énoncé réciproque.

Proposition 1.4.3: Condition suffisante pour un extremum local strict

Soit $f : \mathbb{R} \to \mathbb{R}$ une fonction 2 fois dérivable en $x_0 \in \mathbb{R}$.

1. Si $f'(x_0) = 0$ et $f''(x_0) < 0$, alors f admet un maximum local strict en x_0.
2. Si $f'(x_0) = 0$ et $f''(x_0) > 0$, alors f admet un minimum local strict en x_0.

Démonstration. La preuve repose encore sur la formule de Taylor à l'ordre 2, puisqu'on a

$$f(x) = f(x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2\epsilon(x)$$

avec $\epsilon(x) \to 0$ quand $x \to x_0$. Par le même raisonnement que dans la preuve précédente, si $f''(x_0) < 0$, alors

$$\frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2\epsilon(x) = \big(\frac{1}{2}f''(x_0) + \epsilon(x)\big)(x - x_0)^2 < 0$$

pour tout x assez proche de x_0, $x \neq x_0$. Ainsi, $f(x) < f(x_0)$ pour tout x assez proche de x_0, $x \neq x_0$: c'est bien la définition de maximum local strict. De même dans le cas $f''(x_0) > 0$. □

Remarque 1.4.2 On remarque que les conditions nécessaires et suffisantes ne coïncident pas : dans la Proposition 1.4.2 on a des inégalités large ($f''(x_0) \leq 0$) alors que dans la Proposition 1.4.3 on a des inégalités strictes ($f'(x_0) < 0$). Ce n'est pas un hasard. En effet, il est possible que x_0 soit un maximum local strict et que pourtant $f''(x_0) = 0$: c'est le cas par exemple pour $f(x) = -x^4$ en $x_0 = 0$. Réciproquement, il est possible que $f'(x_0) = 0$ et que $f''(x_0) \leq 0$ sans que x_0 ne soit un maximum local de f : par exemple avec $f(x) = x^4$ et $x_0 = 0$.

Remarque 1.4.3 Dans le cas où $f'(x_0) = 0$ et $f''(x_0) = 0$, on ne peut en général rien dire. Tous les cas de figure peuvent se présenter (maximum local, minimum local, ni l'un ni l'autre), par exemple choisissant $f(x) = x^4$, $f(x) = -x^4$, ou $f(x) = x^3$.

Conclusion : On voit ainsi que pour des fonctions d'une variable, la formule de Taylor à l'ordre 2 nous permet d'avoir des informations sur les extrema locaux d'une fonctions. Nous généralisons cela à présent aux fonctions de 2 variables.

1.4.2 Cas des fonctions de deux (ou n) variables

On commence encore par les conditions nécessaires. La proposition ci-dessous marche aussi pour une fonction $f : \mathbb{R}^n \to \mathbb{R}$, avec presque la même démonstration.

Proposition 1.4.4: Condition nécessaire pour un extremum local en $2d$

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 et $X_0 \in \mathbb{R}^2$. Si X_0 est un extremum local pour f, on a alors

$$\text{grad}\, f(X_0) = \vec{0}.$$

De plus,

1. si X_0 est un maximum local, alors pour tout vecteur $V \in \mathbb{R}^2$, on a
$$V \cdot \text{Hess}\, f(X_0)V \leq 0,$$
2. si X_0 est un minimum local, alors pour tout vecteur $V \in \mathbb{R}^2$, on a
$$V \cdot \text{Hess}\, f(X_0)V \geq 0.$$

Remarque 1.4.4 La condition $\text{grad}\, f(X_0) = \vec{0}$ correspond à la propriété $f'(X_0) = 0$ en $1d$. La condition $V \cdot \text{Hess}\, f(X_0)V \leq 0$ (resp. ≥ 0) pour tout $V \in \mathbb{R}^2$ correspond à la propriété $f''(X_0) \leq 0$ (resp. ≥ 0) en $1d$. Une matrice 2×2 symétrique H telle que $V \cdot HV \leq 0$ (resp. ≥ 0) est appelée matrice symétrique négative (resp. matrice symétrique positive). Cela ne signifie pas que tous ses éléments sont négatifs (resp. positifs) !

Remarque 1.4.5 Un point X_0 tel que $\text{grad}\, f(X_0) = \vec{0}$ s'appelle *point critique* de f.

Démonstration. On va déduire les propriétés de leur analogue $1d$. Pour le fait que le gradient soit nul, on remarque que si $X_0 = (a_1, a_2)$, alors les fonctions d'une variable
$$x_1 \mapsto f(x_1, a_2), \quad x_2 \mapsto f(a_1, x_2)$$
possèdent un extremum local, en $x_1 = a_1$ pour la première et en $x_2 = a_2$ pour la seconde. Par la Proposition 1.4.1, leurs dérivées sont donc nulles en ces points, ce qui veut dire que les dérivées partielles de f par rapport aux deux variables en X_0 sont nulles : le gradient de f en X_0 est donc nul.

Pour les propriétés de la hessienne qui vont suivre, on se ramène également au cas d'une variable. Soit donc $V \in \mathbb{R}^2$. La fonction d'une variable
$$\varphi : t \mapsto f(X_0 + tV)$$
admet un extremum local en $t = 0$. Si X_0 est un maximum local pour f, $t = 0$ est donc un maximum local pour φ, et donc $\varphi''(0) \leq 0$. Or un calcul montre que
$$\varphi''(0) = V \cdot \text{Hess}\, f(X_0)V,$$
ce qui démontre le point 1. Le point 2. se démontre de la même manière. □

Généralisation à la dimension n : La matrice Hessienne $H = \text{Hess}\, f(X_0)$ définit une forme quadratique Q sur $\mathbb{R}^n$:
$$Q(V) = V \cdot HV = \sum_{i=1}^{n} \sum_{j=1}^{n} H_{i,j} v_i v_j \,,$$
où $H_{i,j}$ est l'élément (i, j) (à la ligne i et la colonne j) de la matrice H, donc $H_{i,j} = \frac{\partial^2 f}{\partial x_i \partial x_j}(X_0)$. La formule de Taylor à l'ordre 2 fait apparaître cette forme quadratique :
$$f(X_0 + tV) = f(X_0) + t\, \text{grad}\, f(X_0) \cdot V + \frac{t^2}{2} Q(V) + t^2 \|V\|^2 \epsilon(tV)\,.$$

Si X_0 est un maximum, on doit avoir $\operatorname{grad} f(X_0)\cdot V = 0$ pour tout $V \in \mathbb{R}^2$, donc $\operatorname{grad} f(X_0) = 0$. On a alors

$$f(X_0 + tV) = f(X_0) + \frac{t^2}{2}Q(V) + t^2\|V\|^2\epsilon(tV)\,.$$

On peut prendre les vecteurs V de norme 1, quitte à modifier t. Donc

$$f(X_0 + tV) - f(X_0) = t^2\big(\frac{1}{2}Q(V) + \epsilon(tV)\big)\,,$$

avec $\epsilon(tV) \to 0$ lorsque $t \to 0$. Choisissons V arbitraire de norme 1. Si X_0 est un minimum local, le membre de gauche est ≥ 0 pour t assez petit, il faut alors que $Q(V) \geq 0$. Si c'est vérifié pour tout V de norme 1, alors la forme quadratique Q doit être positive.

On a également un énoncé réciproque.

Proposition 1.4.5: Condition suffisante pour un extremum local en $2d$

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe de C^2 et $X_0 \in \mathbb{R}^2$. On suppose que $\operatorname{grad} f(X_0) = \vec{0}$.

1. S'il existe $c > 0$ tel que pour tout $V \in \mathbb{R}^2$, $V \cdot \operatorname{Hess} f(X_0)V \leq -c\|V\|^2$, alors X_0 est un maximum local strict pour f.
2. S'il existe $c > 0$ tel pour tout $V \in \mathbb{R}^2$, $V \cdot \operatorname{Hess} f(X_0)V \geq c\|V\|^2$, alors X_0 est un minimum local strict pour f.

Ici aussi le résultat reste valable en dimension $n > 2$; en général, la forme quadratique $V \mapsto Q(V) = V \cdot \operatorname{Hess} f(X_0)V$ sera plus difficile à étudier.

Démonstration. On suppose la condition 1., la preuve pour la condition 2. se faisant de la même manière. Par la formule de Taylor à l'ordre 2, et en utilisant que $\operatorname{grad} f(X_0) = \vec{0}$, on a pour tout $X \in \mathbb{R}^2$

$$f(X) = f(X_0) + \frac{1}{2}(X - X_0) \cdot \operatorname{Hess} f(X_0)(X - X_0) + \|X - X_0\|^2\epsilon(X)$$

avec $\epsilon(X) \to 0$ quand $X \to X_0$. En utilisant l'hypothèse faite sur $\operatorname{Hess} f(X_0)$, on en déduit

$$f(X) - f(X_0) \leq (-c + \epsilon(X))\|X - X_0\|^2.$$

Comme $\|X - X_0\|^2 > 0$ pour $X \neq X_0$, que $-c < 0$ et que $\epsilon(X) \to 0$ quand $X \to X_0$, on en déduit que $-c + \epsilon(X) < 0$ pour X suffisamment proche de X_0 et donc

$$f(X) - f(X_0) < 0$$

pour X proche de X_0, $X \neq X_0$, ce qui montre bien que X_0 est un maximum local strict de f. □

Là encore, les conditions nécessaires et suffisantes diffèrent en ce qui concerne les dérivées secondes. Pour la condition nécessaire d'un maximum local par exemple, on a

$$\forall V \in \mathbb{R}^2, \quad V \cdot \operatorname{Hess} f(X_0) V \leq 0,$$

alors que la condition suffisante est

$$\exists c > 0, \quad \forall V \in \mathbb{R}^2, \quad V \cdot \operatorname{Hess} f(X_0) V \leq -c\|V\|^2.$$

Une matrice symétrique $H = \operatorname{Hess} f(X_0)$ satisfaisant la propriété ci-dessus est dite *définie négative*.

Il n'est pas évident de traduire ces propriétés en termes des coefficients de la matrice $\operatorname{Hess} f(X_0)$ (qui sont les dérivées partielles secondes de f). C'est le but du lemme suivant. Cette fois, on va utiliser le fait que en dimension 2, les calculs sont plus simples (et on a moins de possibilités).

Lemme 1.4.6 *Soit A une matrice carrée symétrique de taille 2, dont on note les coefficients*

$$A = \begin{bmatrix} r & s \\ s & t \end{bmatrix}$$

avec $r, s, t \in \mathbb{R}$.

1. *Il existe $c > 0$ telle que pour tout $V \in \mathbb{R}^2$, on ait $V \cdot AV \geq c\|V\|^2$ si et seulement si $\det(A) = rt - s^2 > 0$ et $r > 0$.*
2. *Il existe $c > 0$ telle que pour tout $V \in \mathbb{R}^2$, on ait $V \cdot AV \leq -c\|V\|^2$ si et seulement si $\det(A) = rt - s^2 > 0$ et $r < 0$.*

Démonstration. On montre le point 1., le point 2. se démontrant de la même façon. Supposons dans un premier temps que $rt - s^2 > 0$ et $r > 0$. Ceci implique $t > 0$, car sinon on aurait $rt - s^2 \leq 0$. En particulier, le polynôme

$$p(\lambda) := r\lambda^2 + 2s\lambda + t, \quad \lambda \in \mathbb{R}$$

a comme discriminant $\operatorname{disc}(p) = 4s^2 - 4rt$ qui est strictement négatif, donc $p(\lambda)$ ne s'annule pas sur $\mathbb{R}$; comme son coefficient dominant r est strictement positif, ce polynôme est strictement positif sur $\mathbb{R}$. Ainsi, il existe $c_1 > 0$ tel que

$$\forall \lambda \in \mathbb{R}^2, \quad p(\lambda) = r\lambda^2 + 2s\lambda + t \geq c_1.$$

De même, comme le polynôme $\lambda^2 p(1/\lambda) = t\lambda^2 + 2s\lambda + r$ est également strictement positif sur $\mathbb{R}$, il est supérieur à une constante $c_2 > 0$.

Montrons à présent qu'il existe $c > 0$ tel que pour tout $V \in \mathbb{R}^2$ on a $V \cdot AV \geq c\|V\|^2$. Soit donc $V = (x, y) \in \mathbb{R}^2$. On peut supposer que $V \neq 0$, car la relation que l'on veut montrer est vraie pour $V = 0$, et ce pour n'importe quelle valeur de c. On a de plus

$$V \cdot AV = rx^2 + 2sxy + ty^2.$$

On distingue alors deux cas. Si $y^2 \geq x^2$, alors $V \neq 0$ implique que $y^2 \neq 0$ et alors

$$V \cdot AV = y^2\big(r(x/y)^2 + 2s(x/y) + t\big) = y^2 p(x/y) \geq c_1 y^2.$$

Or en utilisant $y^2 \geq x^2$ on en déduit que

$$c_1 y^2 = \frac{c_1}{2}y^2 + \frac{c_1}{2}y^2 \geq \frac{c_1}{2}(x^2+y^2) = \frac{c_1}{2}\|v\|^2.$$

Si maintenant $x^2 \geq y^2$, on en déduit par la même méthode (en factorisant cette fois-ci par x^2) que

$$V \cdot AV \geq \frac{c_2}{2}\|V\|^2.$$

En posant $c = \min(c_1/2, c_2/2) > 0$, on a bien dans tous les cas

$$V \cdot AV \geq c\|V\|^2,$$

ce qui est la propriété voulue.

Réciproquement, si on suppose cette propriété, alors pour tout $x \in \mathbb{R}$, en choisissant $V = (x, 1)$, on a

$$V \cdot AV = p(x) \geq c\|V\|^2 = c(x^2+1) \geq c > 0,$$

et en particulier le polynôme $p(x)$ ne s'annule pas sur $\mathbb{R}$ et reste strictement positif. Cela implique que son discriminant est strictement négatif et que son coefficient dominant est strictement positif, autrement dit $rt - s^2 > 0$ et $r > 0$. □

En combinant la Proposition 1.4.5 et le Lemme 1.4.6, on aboutit au résultat suivant.

Proposition 1.4.7

Soit $f : \mathbb{R}^2 \to \mathbb{R}$ de classe C^2 et $x_0 \in \mathbb{R}$. On suppose que $\operatorname{grad} f(x_0) = \vec{0}$, et on pose

$$r = \frac{\partial^2 f}{\partial x^2}(x_0), \quad s = \frac{\partial^2 f}{\partial x \partial y}(x_0), \quad t = \frac{\partial^2 f}{\partial y^2}(x_0).$$

Si on a l'inégalité stricte

$$rt - s^2 > 0,$$

alors on a les deux possibilités suivantes :

1. Si $r > 0$, alors x_0 est un minimum local strict pour f.
2. Si $r < 0$, alors x_0 est un maximum local strict pour f.

Remarque 1.4.6 Les conditions $\operatorname{grad} f(x_0) = \vec{0}$, $rt - s^2 > 0$, et $r > 0$ ou $r < 0$ sont simples à vérifier en pratique, puisqu'il suffit de calculer les dérivées partielles de f jusqu'à l'ordre 2.

Remarque 1.4.7 Dans le cas où $rt - s^2 < 0$, on peut montrer que x_0 n'est ni un maximum local, ni un minimum local : on dit que x_0 est un *point selle* ou un *point col*. La forme quadratique $V \mapsto V \cdot AV$ prend alors les deux signes.

Remarque 1.4.8 Quand $\det(A) = rt - s^2 = 0$ (forme quadratique *dégénérée*, et matrice A de rang 1 ou 0), on peut avoir tous les cas de figure ; il faut étudier les dérivées de f d'ordres supérieurs pour voir si f admet un maximum, un minimum ou un point selle.

Exemple 1.4.2 Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

$$f(x, y) = x^4 + y^4 - (x - y)^2.$$

Déterminons ses points critiques, et leurs natures (maximum local, minimum local, ou point selle). On a pour tout $(x, y) \in \mathbb{R}^2$,

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

donc $\operatorname{grad} f(x, y) = \vec{0}$ si et seulement si

$$2x^3 = x - y, \quad 2y^3 = y - x.$$

En particulier, on a $y^3 = -x^3$ et donc $y = -x$. On trouve donc $2x^3 = x - y = 2x$ donc $x^3 = x$, ce qui est vrai uniquement si $x = 0$, $x = 1$, ou $x = -1$. On trouve donc 3 points critiques de f

$$(0, 0), \quad (1, -1), \quad (-1, 1).$$

Pour déterminer leur nature, on se sert de la Proposition 1.4.7 et donc on a besoin de connaître les dérivées partielles secondes de f en ces points. On a pour tout $(x, y) \in \mathbb{R}^2$,

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

Ainsi, aux points $(1, -1)$ et $(-1, 1)$, on a

$$r = 10, \quad s = 2, \quad t = 10,$$

donc $rt - s^2 = 96 > 0$ et $r = 10 > 0$ donc $(1, -1)$ et $(-1, 1)$ sont des minima locaux stricts pour f. Au point $(0, 0)$, on a

$$r = -2, \quad s = 2, \quad t = -2,$$

donc $rs - t^2 = 0$, ce qui n'est pas un cas couvert par la Proposition 1.4.7. Néanmoins, à partir de la forme de f, on peut quand même déterminer sa nature. En effet, on a $f(x, x) = 2x^4 > 0 = f(0, 0)$ et $f(x, -x) = 2x^4 - 4x^2 < 0 = f(0, 0)$ pour $x \neq 0$ petit. Ainsi, $(0, 0)$ n'est ni un maximum local, ni un minimum local pour f : c'est un point selle. On a représenté le graphe de f en Figure 1.15 : on y voit les deux minima locaux ainsi que le point selle entre les deux.

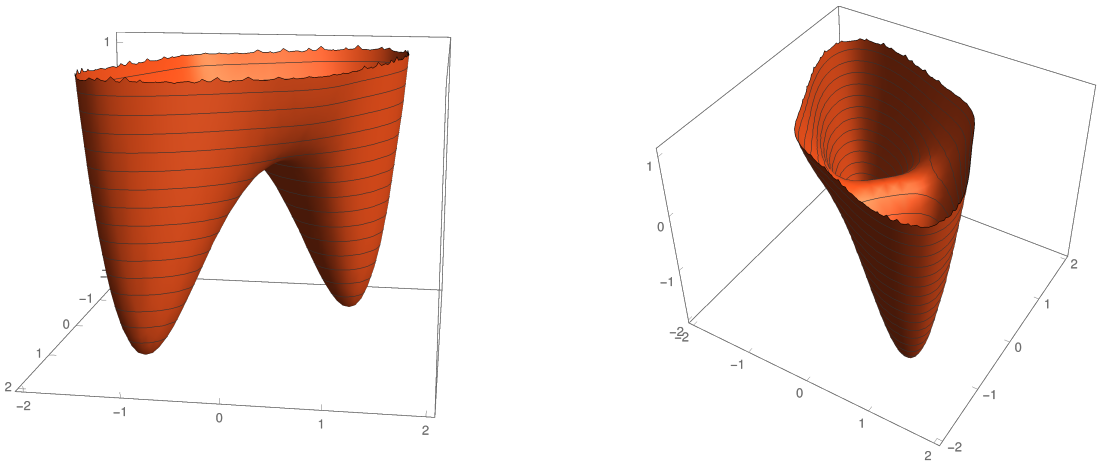

FIGURE 1.15 – Graphe de $(x, y) \mapsto x^4 + y^4(x - y)^2$.

Exemple 1.4.3 Soit $f : \mathbb{R}^2 \to \mathbb{R}$ la fonction définie par

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

pour tout $(x, y) \in \mathbb{R}^2$. On montre facilement que $(0, 0)$ est le seul point critique de f, et qu'en ce point

$$r = 2, \quad s = 0, \quad t = -2$$

donc $rt - s^2 < 0$: dans ce cas, on peut immédiatement conclure que $(0, 0)$ est un point selle pour f. Cela se voit très bien graphiquement en traçant le graphe de f, comme en Figure 1.12 à droite.

1.5 Analyse vectorielle

Enfin, on termine le premier chapitre par l'introduction de certaines notions d'analyse vectorielle (gradient, divergence, rotationnel) et leurs relations. Dans cette section on notera le point de $\mathbb{R}^n$ par la lettre M plutôt que par le vecteur de coordonnées $X = (x_1, \ldots, x_n)$, afin de se rapprocher des notations utilisées en physique. En notant O le point origine des coordonnées dans $\mathbb{R}^n$, on a donc l'identification entre vecteurs $\vec{OM} = X$.

Définition 1.5.1: Champ scalaire, champ vectoriel

On se placera en dimension $n \geq 2$. Soit $\Omega \subset \mathbb{R}^n$ un domaine de $\mathbb{R}^n$.

— Un *champ scalaire* sur Ω est une fonction $f : \Omega \to \mathbb{R}$: à chaque point $M \in \Omega$, on associe un scalaire $f(M) \in \mathbb{R}$.

— Un *champ vectoriel* (ou champ de vecteurs) sur Ω est une fonction $F : \Omega \to \mathbb{R}^n$: à chaque point M on associe un vecteur $F(M) \in \mathbb{R}^n$, qu'on représente par une

flèche attachée au point M. Pour tout $M \in \mathbb{R}^n$, on note

$$F(M) = (F_1(M), \ldots, F_n(M)) \in \mathbb{R}^n.$$

Chaque fonction composante F_j de F est donc un champ scalaire.

Notation On notera souvent les champs scalaires avec des lettres minuscules (f, g, h,...), et les champs vectoriels avec des lettres majuscules (F, G, H,...).

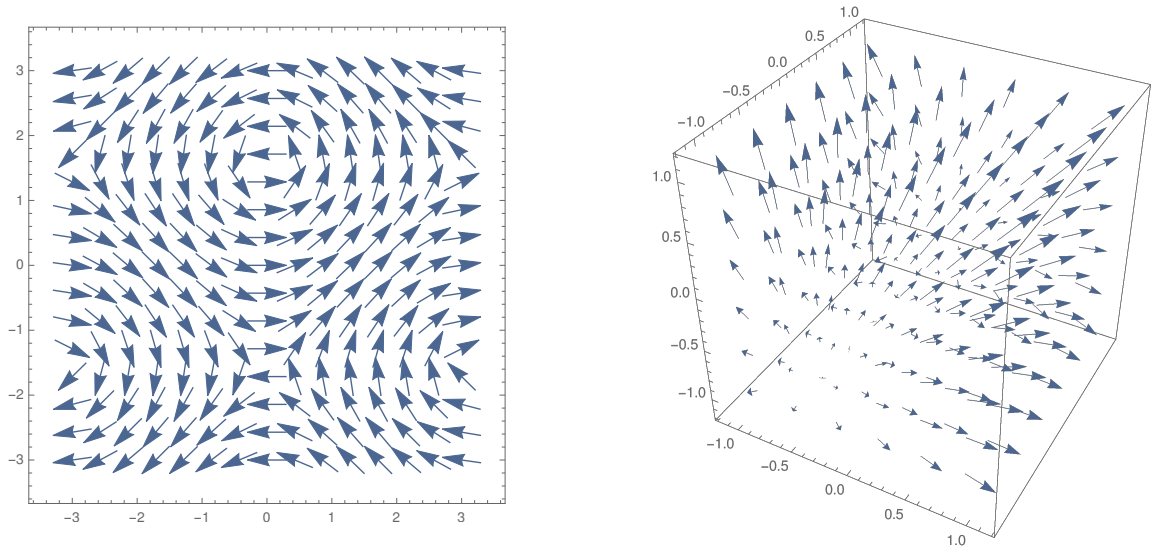

FIGURE 1.16 – Exemples de champs de vecteurs sur $\mathbb{R}^2$ et sur $\mathbb{R}^3$.

Exemple 1.5.2 — Si f est un champ scalaire C^1 sur Ω, alors son gradient

$$\operatorname{grad} f : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R}^n \\ M & \longmapsto & \left(\dfrac{\partial f}{\partial x_1}(M), \ldots, \dfrac{\partial f}{\partial x_n}(M) \right) \end{array} \right.$$

est un champ vectoriel (continu) sur Ω.

— Si f est un champ scalaire C^2, son *laplacien*

$$\Delta f : \left\{ \begin{array}{ccc} \Omega & \longrightarrow & \mathbb{R} \\ M & \longmapsto & \displaystyle\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}(M) = \frac{\partial^2 f}{\partial x_1^2}(M) + \frac{\partial^2 f}{\partial x_2^2}(M) + \cdots + \frac{\partial^2 f}{\partial x_n^2}(M) \end{array} \right.$$

est un autre champ scalaire (continu) sur Ω.

— Si $F(M)$ est la force coulombienne s'exerçant sur une particule ponctuelle de charge q_1 située en un point $M \in \mathbb{R}^3$ par une particule ponctuelle de charge q_2 située en un point $O \in \mathbb{R}^3$, on a

$$F(M) = \frac{q_1 q_2}{4\pi\varepsilon_0} \frac{\overrightarrow{OM}}{\|\overrightarrow{OM}\|^3} \in \mathbb{R}^3,$$

où ε_0 est la constante diélectrique du vide. En particulier, on peut voir F comme un champ vectoriel

$$F : \begin{cases} \mathbb{R}^3 \setminus \{0\} & \longrightarrow & \mathbb{R}^3 \\ M & \longmapsto & F(M) = \dfrac{q_1 q_2}{4\pi\varepsilon_0} \dfrac{\overrightarrow{OM}}{\|\overrightarrow{OM}\|^3} \end{cases},$$

que l'on a représenté en Figure 1.17. Ce champ vectoriel *dérive d'un champ scalaire*, dans le sens où l'on a

$$F = -\operatorname{grad} f,$$

où f est le potentiel coulombien

$$f : \begin{cases} \mathbb{R}^3 \setminus \{0\} & \longrightarrow & \mathbb{R} \\ M & \longmapsto & f(M) = \dfrac{q_1 q_2}{4\pi\epsilon_0} \dfrac{1}{\|\overrightarrow{OM}\|} \end{cases}$$

qui est lui un champ scalaire.

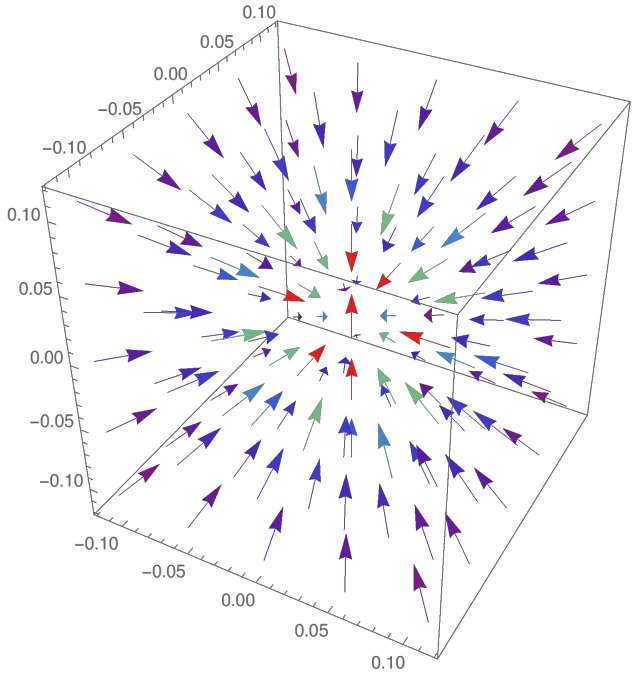

FIGURE 1.17 – Représentation du champ coulombien F dans le cas de charges opposées $q_1 q_2 < 0$ (la couleur du vecteur représente sa norme : plus la couleur est rouge plus la norme est importante).

Définition 1.5.3: Divergence, rotationnel

— Soit $F : \mathbb{R}^n \to \mathbb{R}^n$ un champ vectoriel C^1. On appelle *divergence de f* le champ scalaire

$$\operatorname{div} F : \begin{cases} \mathbb{R}^n & \longrightarrow & \mathbb{R} \\ M & \longmapsto & \displaystyle\sum_{j=1}^{n} \frac{\partial F_j}{\partial x_j}(M) = \frac{\partial F_1}{\partial x_1}(M) + \frac{\partial F_2}{\partial x_2}(M) + \cdots + \frac{\partial F_n}{\partial x_n}(M) \end{cases}.$$

— En dimension $n = 2$: soit $F : \mathbb{R}^2 \to \mathbb{R}^2$ un champ vectoriel C^1. On appelle *rotationnel de F* le champ scalaire

$$\operatorname{rot} F : \begin{cases} \mathbb{R}^2 & \longrightarrow & \mathbb{R} \\ M & \longmapsto & \displaystyle\frac{\partial F_2}{\partial x_1}(M) - \frac{\partial F_1}{\partial x_2}(M) \end{cases}.$$

— En dimension $n = 3$: soit $F : \mathbb{R}^3 \to \mathbb{R}^3$ un champ vectoriel C^1. On appelle *rotationnel de F* le champ vectoriel

$$\vec{\operatorname{rot}} F : \begin{cases} \mathbb{R}^3 & \longrightarrow & \mathbb{R}^3 \\ M & \longmapsto & \left(\dfrac{\partial F_3}{\partial x_2}(M) - \dfrac{\partial F_2}{\partial x_3}(M), \dfrac{\partial F_1}{\partial x_3}(M) - \dfrac{\partial F_3}{\partial x_1}(M), \dfrac{\partial F_2}{\partial x_1}(M) - \dfrac{\partial F_1}{\partial x_2}(M)\right). \end{cases}$$

Attention La divergence d'un champ vectoriel est défini en tout dimension, alors qu'on a défini le rotationnel de champs vectoriels uniquement en dimension 2 et 3. En dimension 2, il s'agit d'un champ scalaire alors qu'en dimension 3 il s'agit d'un champ vectoriel. J'essaierai de systématiquement mettre une flèche sur le rotationnel dans $\mathbb{R}^3$, mais parfois j'oublierai la flèche, il faudra tout de même se souvenir qu'il s'agit d'un champ vectoriel.

Remarque 1.5.1 En notant formellement le gradient comme un vecteur colonne

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

on a alors

$$\operatorname{div} F = \nabla \cdot F, \qquad \vec{\operatorname{rot}} F = \nabla \wedge F,$$

où $\wedge$ désigne le produit vectoriel de deux vecteurs. C'est un moyen mnémotechnique intéressant de se souvenir des formules précédentes.

Rappelons qu'en dimension 3, le produit vectoriel $X \wedge Y$ est donné par les formules $e_1 \wedge e_2 = e_3$, $e_2 \wedge e_3 = e_1$, $e_3 \wedge e_1 = e_2$ (où e_1, e_2, e_3 sont les vecteurs de base de $\mathbb{R}^3$), et noter que dans le produit on suit l'orientation), plus le fait que $X \wedge Y = -Y \wedge X$, plus le fait que $X \wedge Y$ est une fonction linéaire de X à Y fixé et une fonction linéaire de Y à X fixé.

En dimension 2 c'est plus simple, $X \wedge Y$ est un scalaire, et on prend $(x_1, x_2) \wedge (y_1, y_2) = x_1y_2 - y_1x_2$. Autrement dit, $e_1 \wedge e_2 = 1$ et comme plus haut $X \wedge Y = -Y \wedge X$, et le produit est une fonction linéaire de X à Y fixé et une fonction linéaire de Y à X fixé.

Remarque 1.5.2 Il est facile de montrer que le laplacien peut s'écrire ainsi :

$$\Delta f = \operatorname{div} \nabla f = \nabla \cdot \nabla f.$$

Proposition 1.5.1

— Si $f : \mathbb{R}^3 \to \mathbb{R}$ est un champ scalaire C^2, alors pour tout $M \in \mathbb{R}^3$ on a

$$\vec{\operatorname{rot}}(\operatorname{grad} f)(M) = \vec{0}.$$

— Si $F : \mathbb{R}^3 \to \mathbb{R}^3$ est un champ vectoriel C^2, alors pour tout $M \in \mathbb{R}^3$,

$$\operatorname{div}(\vec{\operatorname{rot}} F)(M) = 0.$$

— Mêmes énoncés dans $\mathbb{R}^2$, mais remplacez aussi $\vec{\operatorname{rot}}$ par rot (respectivement $\vec{0}$ par 0) dans la première équation.

Démonstration. Calcul direct. □

De la même manière que l'on a la formule de Leibniz $(fg)' = f'g + fg'$ en dimension 1, on a les analogues suivants en dimension supérieure.

Proposition 1.5.2

— Si f et g sont des champs scalaires C^1, alors

$$\operatorname{grad}(fg) = f \operatorname{grad} g + g \operatorname{grad} f.$$

— Si f est un champ scalaire C^1 et F un champ vectoriel C^1, on a

$$\operatorname{div}(fF) = f \operatorname{div} F + \operatorname{grad} f \cdot F.$$

— Soit $n = 2$ ou $n = 3$. Si $f : \mathbb{R}^n \to \mathbb{R}$ est un champ scalaire C^1 et $F : \mathbb{R}^n \to \mathbb{R}^n$ est un champ vectoriel C^1, on a

$$\operatorname{rot}(fF) = f \operatorname{rot} F + \operatorname{grad} f \wedge F.$$

— Toujours pour $n = 2$ ou $n = 3$, si le champ vectoriel F ci-dessus est donné par le gradient d'un champ scalaire C^2 $g : \mathbb{R}^n \to \mathbb{R}$, alors on trouve

$$\operatorname{rot}(\operatorname{grad} g) = 0, \quad \text{et par conséquent} \quad \operatorname{rot}(f \operatorname{grad} g) = \operatorname{grad} f \wedge \operatorname{grad} g.$$

Démonstration. Calcul direct. □

Définition 1.5.4: Champ de gradient

Soit $F : \mathbb{R}^n \to \mathbb{R}^n$ un champ vectoriel continu. On dit que F *dérive d'un potentiel scalaire* f si f est C^1 et que $F = \operatorname{grad} f$. On dit également que F est un *champ de gradient*. Un tel f est appelé *potentiel scalaire* associé à F.

Remarque 1.5.3 Si F dérive d'un potentiel scalaire f, ce dernier n'est pas unique : pour n'importe quelle constante C, le potentiel scalaire $f + C$ admet aussi pour gradient F.

D'après la Proposition 1.5.1, si F est un champ de gradient en dimension 2 ou 3 alors $\operatorname{rot} F \equiv 0$. On a un résultat réciproque qui permet de reconnaître un champ de gradient.

Théorème 1.5.3 *Soit $n = 2$ ou $n = 3$. Soit $\Omega \subset \mathbb{R}^n$ un domaine de $\mathbb{R}^n$ et $F : \Omega \to \mathbb{R}^n$ un champ vectoriel C^1 sur Ω. Si Ω est* étoilé *(voir la définition ci-dessous) et si* $\operatorname{rot} F \equiv 0$, *alors F est un champ de gradient.*

En fait, il s'agit d'un cas particulier du résultat suivant, qui a l'avantage d'être vrai dans tout $\mathbb{R}^n$ et de ne pas reposer sur un miracle de basse dimension.

Théorème 1.5.4 *Soit $\Omega \subset \mathbb{R}^n$ un domaine (ouvert) de $\mathbb{R}^n$ et $F : \Omega \to \mathbb{R}^n$ un champ vectoriel de classe C^1 sur Ω. Si F est un champ de gradient, alors*

$$\frac{\partial F_i}{\partial x_j}(X) = \frac{\partial F_j}{\partial x_i}(X) \quad \textit{pour tout choix de } i, j \in \{1, 2, \ldots, n\} \textit{ et tout } X \in \Omega. \tag{1.36}$$

Réciproquement, si Ω est étoilé, *et si $F : \Omega \to \mathbb{R}^n$ un champ vectoriel de classe C^1 sur Ω qui vérifie les identités* (1.36), *alors F est le gradient d'un certain champ scalaire C^2 $f : \Omega \to \mathbb{R}$.*

Le fait que (1.36) est une condition nécessaire est facile, elle provient de l'identité de Schwarz sur les dérivées croisées. En effet, si F est un champ de gradient C^1, donc si $F = \nabla f$ pour un certain champ scalaire f, alors clairement f est de classe C^2, et $\frac{\partial F_i}{\partial x_j} = \frac{\partial}{\partial x_j}(\frac{\partial f}{\partial x_i}) = \frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial F_j}{\partial x_i}$. Cet argument fournit donc en particulier la condition nécessaire dans le théorème 1.5.3 (l'utilisation du symbole rotationnel est juste un raccourci).

Pour la réciproque (condition suffisante à être un gradient), on va montrer le théorème pour $\Omega = \mathbb{R}^2$; ce serait pareil sur $\mathbb{R}^n$, avec juste plus d'étapes. Pour un domaine Ω étoilé, la démonstration est plus délicate : on peut trouver la solution f seulement près d'un point comme ci-dessous, mais ensuite il faut vérifier que cette solution peut être étendue à tout Ω de manière cohérente. En fait pour ceci on a besoin d'une propriété "topologique" de Ω, qu'on appelle la *simple connexité*, qui dit que Ω n'a pas de trous ; cette propriété est par exemple satisfaite quand Ω est étoilé par rapport à un de ses points, comme ci-dessous.

Rappel : On dit qu'un domaine Ω est *étoilé* s'il existe un point $M_0 \in \Omega$ tel que pour tout autre point $M \in \Omega$, le segment $[M_0, M]$ est inclus dans Ω. Par exemple, tout domaine convexe (une boule par exemple) est étoilé (avec M_0 n'importe quel point). Une étoile est un domaine étoilé (avec M_0 le centre de l'étoile). Un exemple de domaine non-étoilé (qui n'est

étoilé par rapport à aucun de ses points M_0) est $\mathbb{R}^n \setminus \{0\}$: pour tout choix de $M_0 \in \mathbb{R}^n \setminus \{0\}$, le segment $[-M_0, M_0]$ passe par l'origine et donc n'est pas inclus tout entier dans $\mathbb{R}^n \setminus \{0\}$. Ou peut-être, de manière encore plus visible, un anneau comme $B(0,2) \setminus B(0,1)$ (prendre à nouveau deux points opposés). Par contre, les boules, les ellipsoides, les domaines convexes, sont étoilés. Et je rappelle pour votre culture que l'ensemble $E \subset \mathbb{R}^n$ est convexe quand pour tout choix de $X, Y \in \Omega$, le segment de droite $[X, Y]$ qui va de X à Y est contenu dans E.

Exemple 1.5.5

1. Soit $F : \mathbb{R}^2 \to \mathbb{R}^2$ le champ de vecteurs défini par la relation : pour tout $(x, y) \in \mathbb{R}^2$,

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

Montrons que F est un champ de gradient, et trouvons un potentiel scalaire f dont F dérive. On calcule d'abord le rotationnel de F : pour tout $(x, y) \in \mathbb{R}^2$,

$$\operatorname{rot} F(x, y) = \frac{\partial F_2}{\partial x}(x, y) - \frac{\partial F_1}{\partial y}(x, y) = 1 - 1 = 0.$$

Comme F est de rotationnel nul et défini sur le domaine étoilé $\Omega = \mathbb{R}^2$, le Théorème 1.5.3 nous assure que F est un champ de gradient, qui dérive d'un certain potentiel f. Pour identifier un choix de f possible, on utilise la relation $F = \operatorname{grad} f$. Elle signifie que pour tout $(x, y) \in \mathbb{R}^2$,

$$F_1(x, y) = x + y = \frac{\partial f}{\partial x}(x, y), \quad F_2(x, y) = x + 2y = \frac{\partial f}{\partial y}(x, y).$$

En intégrant la première relation par rapport à x à y fixé, on obtient qu'il existe une constante $C(y) \in \mathbb{R}$ telle que pour tout $(x, y) \in \mathbb{R}$,

$$f(x, y) = \frac{1}{2}x^2 + xy + C(y).$$

On identifie à présent la fonction C. Comme f est C^1, on en déduit que C est C^1. En dérivant cette dernière relation par rapport à y, on obtient pour tout $(x, y) \in \mathbb{R}^2$,

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

Or, on sait que $\frac{\partial f}{\partial y} = F_2$, d'où pour tout $(x, y) \in \mathbb{R}^2$,

$$x + C'(y) = x + 2y,$$

d'où $C(y) = y^2 + c$ pour une certaine constante $c \in \mathbb{R}$. Ainsi, on trouve que pour tout $(x, y) \in \mathbb{R}^2$,

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

pour une certaine constante $c \in \mathbb{R}$.

2. Soit $F : \mathbb{R}^3 \to \mathbb{R}^3$ le champ de vecteurs défini par : pour tout $(x, y, z) \in \mathbb{R}^3$,

$$F(x, y, z) = (xy, y, z).$$

Déterminons si F est un champ de gradient ou non. Pour cela, on calcule son rotationnel : on trouve que pour tout $(x, y, z) \in \mathbb{R}^3$,

$$\operatorname{rot} F(x, y, z) = (0, 0, -x)$$

et donc $\operatorname{rot} F$ n'est pas identiquement le vecteur nul. Par la Proposition 1.5.1, F n'est donc pas un champ de gradient.

Remarque 1.5.4 En modifiant à peine l'argument ci-dessus, on peut démontrer les théorèmes 1.5.3 et 1.5.4 quand Ω est un rectangle, ou plus généralement un produit d'intervalles.

La condition d'être un domaine étoilé (en fait, d'être simplement connexe) est nécessaire. Dans le cas d'un anneau, par exemple $\Omega = B(0, 10) \setminus \overline{B(0, 1)}$ (ou bien de $\mathbb{R}^n \setminus \{0\}$), le théorème est faux. Dans $\mathbb{R}^2 \setminus \{0\}$, voici un exemple de champ de vecteur F dont le rotationnel est nul, et qui pourtant n'est pas un champ de gradient dans $\mathbb{R}^2 \setminus \{0\}$: Prenons

$$F(x, y) = \Big(-\frac{y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \Big) \quad \text{pour tout} \quad X = (x, y) \neq (0, 0).$$

Je pourrais laisse faire le calcul du rotationnel, juste pour vérifier qu'il s'annule. Mais il est plus simple d'expliquer comment on a trouvé cette fonction. En fait au voisinage de n'importe quel point $X_0 = (x_0, y_0) \neq (0, 0)$, il est possible de définir une fonction $\theta(X)$ qui donne (à un multiple de 2π près) l'angle entre la direction de l'axe des x avec la direction de X. Un petit calcul donne que le gradient de cette fonction est exactement la foncton F ci-dessus. Notez au passage qu'ajouter 2π à la fonction θ ne change pas son gradient. Donc localement, il n'y a pas de problème, on peut définir $\theta(X)$, et son gradient est bien $F(X)$. De plus, le même genre de calcul que ci-dessus nous dit qu'on n'a pas trop le choix : une fois que vous avez choisi la valeur de θ en un point, vous la connaissez partout (la valeur en un point détermine les constantes dans l'argument ci-dessus).

Le problème vient du fait que la fonction θ doit être définie et continue sur tout l'anneau Ω. Mais quand on tourne une fois autour de l'origine et qu'on suit la valeur de θ, celle-ci augmente de 2π. Alors, au point $(2, 0)$, il faudrait que θ soit à la fois égale à 0, par exemple, mais aussi 2π (ce qu'on trouve forcément après avoir fait un tour). Pour résumer, il est impossible de trouver un potentiel bien défini sur l'anneau, et dont le gradient soit donné par F.

On retrouvera ce problème, avec sans doute une démonstration plus claire, quand on verra que la circulation d'un champ de vecteurs F le long d'un cercle n'est pas toujours égale à 0, comme elle devrait l'être si F dérivait d'un potentiel.

Chapitre 2

Intégrales doubles, triples

2.1 Rappels sur les intégrales simples

On commence par rappeler la notion d'intégrale d'une fonction d'une seule variable réelle et leurs méthodes de calcul.

2.1.1 Définition et interprétation

On considère $[a, b]$ un segment de l'axe réel avec $a < b$ et $f : [a, b] \to \mathbb{R}$ une fonction *continue*. On rappelle que *l'intégrale de f sur* $[a, b]$, notée

$$\int_a^b f(x)\, dx,$$

représente l'aire "algébrique" de la région située entre la courbe représentative de f (le graphe de f, $\Gamma_f = \{(x, f(x)), x \in [a, b]\}$) et l'axe des abscisses.

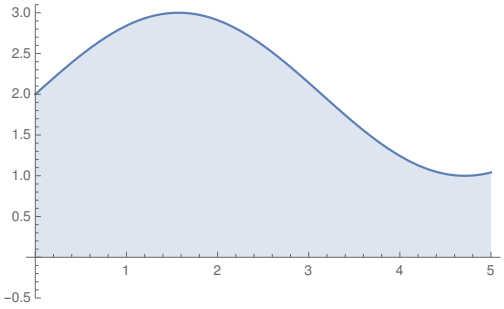

FIGURE 2.1 – L'intégrale représente l'aire de la région bleue sous la courbe représentant le graphe de f.

L'adjectif "algébrique" est là pour dire que dans les régions où f est négative (donc les régions où le graphe de f est situé sous l'axe des abscisses), l'aire est comptée négativement.

Cette notion d'aire sous la courbe est définie par un processus limite, de la manière suivante : on considère une subdivision

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

du segment $[a, b]$ en n sous-segments $[x_0, x_1], [x_1, x_2], \ldots, [x_{n-1}, x_n]$ de longueur (au plus) $\Delta x > 0$. Alors, la quantité

$$\sum_{j=0}^{n-1} (x_{j+1} - x_j) f(x_j)$$

possède une limite lorsque $\Delta x \to 0$, et cette limite est *indépendante* de la subdivision choisie. C'est cette limite que l'on appelle intégrale de f sur $[a, b]$. Géométriquement, la somme

$$\sum_{j=0}^{n-1} (x_{j+1} - x_j) f(x_j)$$

constitue une approximation raisonnable de l'aire sous la courbe représentative de f, puisque $(x_{j+1} - x_j) f(x_j)$ est l'aire d'un rectangle de base $[x_j, x_{j+1}]$ et de hauteur $f(x_j)$, qui approche bien l'aire sous la courbe représentative de f sur le segment $[x_j, x_{j+1}]$ si ce segment est suffisamment petit [1]. La construction de l'intégrale consiste donc à découper le domaine sous la courbe représentative de f en des tranches verticales de plus en plus fines, et d'approcher ces tranches verticales par des rectangles (voir Figure 2.2). Cette construction permet également de définir une notion d'intégrale en dimension supérieure, que l'on verra dans la suite.

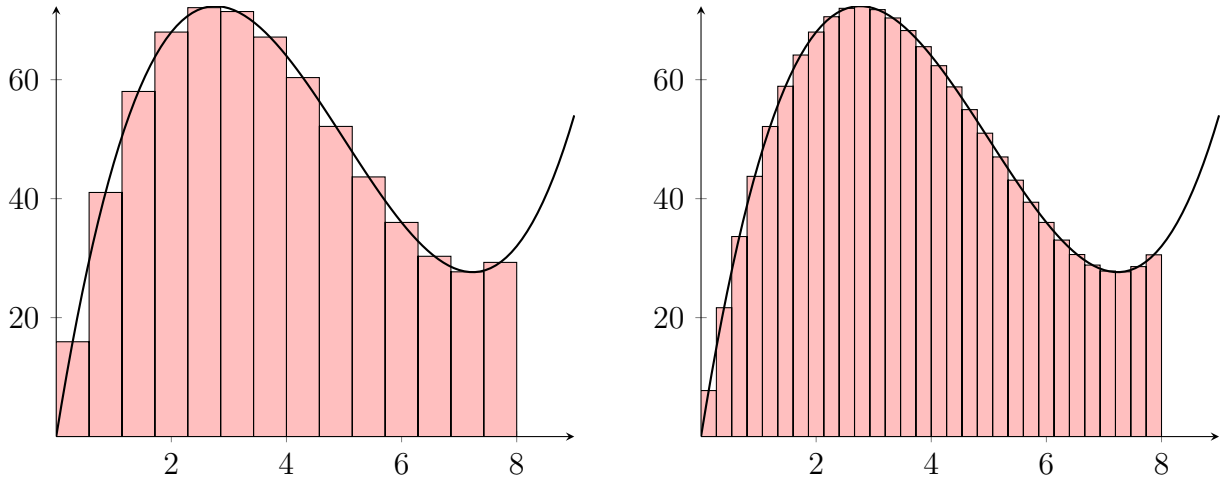

FIGURE 2.2 – Approximation de l'aire sous la courbe par une somme d'aires de rectangles

L'intégrale définie de cette manière possède les propriétés fondamentales suivantes.

1. Cela revient à dire que les valeurs de $f(x)$ sont proches de $f(x_j)$ lorsque $x \in [x_j, x_{j+1}]$. C'est le cas si $[x_j, x_{j+1}]$ est suffisamment petit (donc si Δx est petit), mais on se sert du fait que la fonction est continue ! Si la fonction n'est pas continue et oscille fortement sur des échelles arbitrairement petites (comme la fonction indicatrice des rationnels par exemple) cette construction d'intégrale ne fonctionne plus.

Proposition 2.1.1

1. Linéarité : si $f, g : [a, b] \to \mathbb{R}$ sont continues et si $\lambda, \mu \in \mathbb{R}$, alors
$$\int_a^b (\lambda f(x) + \mu g(x))\, dx = \lambda \int_a^b f(x)\, dx + \mu \int_a^b g(x)\, dx.$$
2. Positivité : si $f \geq 0$, alors $\int_a^b f(x)\, dx \geq 0$.
3. Relation de Chasles : si $a < c < b$ alors
$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx.$$

Remarque 2.1.1 La positivité implique la propriété de croissance par rapport à f de l'intégrale : si $f(x) \leq g(x)$ pour tout $x \in [a, b]$, alors $\int_a^b f(x)\, dx \leq \int_a^b g(x)\, dx$. En effet, il suffit d'appliquer la positivité à $g - f$.

Par convention, si $a < b$ on a
$$\int_b^a f(x)\, dx = -\int_a^b f(x)\, dx.$$
Ceci est l'un des aspects des intégrales simples qui ne passera pas aux intégrales multiples.

Ces propriétés algébriques de l'intégrale sont utiles pour les calculer, mais ne suffisent pas en pratique. On rappelle donc à présent les techniques élémentaires de calcul d'intégrales.

2.1.2 Calcul des intégrales simples

On rappelle ici trois outils basiques de calculs d'intégrales de fonctions d'une variable réelle :

1. l'utilisation des primitives,
2. l'intégration par parties,
3. le changement de variables.

Il existe d'autres techniques plus compliquées, mais on ne les développe pas ici. On pourra consulter le poly de Math 101 pour en voir certaines.

1. Utilisation des primitives

L'outil basique de calcul d'intégrales est le suivant :

Proposition 2.1.2: Théorème fondamental de l'analyse

Si $f : [a, b] \to \mathbb{R}$ est de classe C^1, alors

$$\int_a^b f'(x)\,dx = f(b) - f(a).$$

Cette proposition signifie que l'on sait calculer facilement l'intégrale d'une classe particulière de fonctions : celles qui sont les dérivées de fonctions de classe C^1. Pour calculer l'intégrale d'une fonction g, on cherche donc une fonction G dont elle est la dérivée ($G' = g$) : on appelle G une *primitive* de g. Notons que toute fonction continue admet des primitives C^1. Le calcul d'intégrale à l'aide de cette technique suppose donc que l'on connait une primitive explicite G.

Exemple 2.1.1 Calculons

$$\int_0^1 (x^3 + \cos(x))\,dx.$$

Pour cela, remarquons que la fonction $F : [0, 1] \to \mathbb{R}$ définie par : pour tout $x \in [0, 1]$,

$$F(x) = \frac{1}{4}x^4 + \sin(x),$$

vérifie $F'(x) = x^3 + \cos(x)$. Ainsi,

$$\int_0^1 (x^3 + \cos(x))\,dx = \int_0^1 F'(x)\,dx = F(1) - F(0) = \frac{1}{4} + \sin(1).$$

On trouvera une table de primitives de fonctions usuelles dans le poly de Math 101, page 92.

2. Intégration par parties

L'intégration par parties est liée au résultat précédent : elle repose sur le fait que si f et g sont C^1, alors le produit fg est C^1 et $(fg)' = f'g + fg'$. On en déduit le résultat suivant.

Proposition 2.1.3: Intégration par parties

Soient $f, g : [a, b] \to \mathbb{R}$ deux fonctions C^1. Alors, on a

$$\begin{aligned}\int_a^b f'(x)g(x)\,dx &= [f(x)g(x)]_{x=a}^{x=b} - \int_a^b f(x)g'(x)\,dx \\ &= f(b)g(b) - f(a)g(a) - \int_a^b f(x)g'(x)\,dx.\end{aligned}$$

Cette proposition permet de transformer le calcul de l'intégrale $\int_a^b f'(x)g(x)\,dx$ en le calcul de l'intégrale $\int_a^b f(x)g'(x)\,dx$, qui est potentiellement plus simple (si on connaît une primitive de

fg' par exemple). En pratique, l'utilisation de ce résultat nécessite de décomposer la fonction à intégrer en un produit $f'g$, puis d'intégrer fg'.

Ce calcul peut se faire en disposant les différentes fonctions dans un tableau :

$$\begin{array}{cc|ccc} + & g & & f' \\ & & \searrow & \\ - & g' & \Rightarrow & f, \end{array}$$

où la flèche oblique indique le terme de bord $[f(x)g(x)]_{x=a}^{x=b}$, la flèche horizontale l'intégrale $\int_a^b f(x)g'(x)\,dx$, les deux termes étant précédés des signes correspondants à la gauche des flèches. Dans la colonne de gauche on dérive en descendant, alors que dans la colonne de droite on intègre en descendant.

Exemple 2.1.2 Calculons

$$\int_0^1 x\cos(x)\,dx$$

par intégration par parties. Si $f, g : [0,1] \to \mathbb{R}$ sont définies pour tout $x \in [0,1]$ par $g(x) = x$ et $f(x) = \sin(x)$, alors $x\cos(x) = f'(x)g(x)$ et donc

$$\begin{aligned} \int_0^1 x\cos(x)\,dx &= [x\sin(x)]_{x=0}^{x=1} - \int_0^1 f(x)g'(x)\,dx \\ &= \sin(1) - \int_0^1 \sin(x) \\ &= \sin(1) - [-\cos(x)]_{x=0}^{x=1} \\ &= \sin(1) + \cos(1) - 1. \end{aligned}$$

La notation en tableau s'écrit alors :

$$\begin{array}{cc|ccc} + & x & & \cos(x) \\ & & \searrow & \\ - & 1 & \Rightarrow & \sin(x), \end{array}$$

Sur le tableau on lit $+[x\sin(x)]_0^1 - \int_0^1 \sin(x)\,dx$.

Cette notation en tableau permet aussi d'intégrer par parties plusieurs fois d'affilée, en dérivant à gauche, intégrant à droite, et en n'oubliant pas les signes. Ainsi, l'intégrale ci-dessus peut se faire au biais de 2 IPP :

$$\begin{array}{cc|ccc} + & x & & \cos(x) \\ & & \searrow & \\ - & 1 & & \sin(x) \\ & & \searrow & \\ + & 0 & \Rightarrow & -\cos(x) \end{array}$$

Les deux flèches obliques sont les termes de bord, alors que l'intégrale correspondant à la flèche horizontale est nulle, du fait de la nullité du facteur de gauche. La lecture de ce tableau donne donc $+[x\sin(x)]_0^1 - [-\cos(x)]_0^1 + \int_0^1 0\,dx$.

Voici un exemple un peu plus compliqué, pour lequel la notation en tableau est pratique.

Exemple 2.1.3 Calculons

$$\int_0^1 x^2 \sin(x)\, dx$$

en appliquant deux intégrations par parties successives. Si $f, g : [0,1] \to \mathbb{R}$ vérifient : pour tout $x \in [0,1]$, $f(x) = -\cos(x)$, $g(x) = x^2$, on en déduit que $x^2 \sin(x) = f'(x)g(x)$ et donc

$$\int_0^1 x^2 \sin(x)\, dx = [-x^2 \cos(x)]_0^1 + \int_0^1 2x \cos(x)\, dx = -\cos(1) + \int_0^1 2x \cos(x)\, dx.$$

La deuxième intégrale a été calculée par intégration par parties dans l'intégrale précédente. On a donc

$$\int_0^1 x^2 \sin(x)\, dx = -\cos(1) + 2(\sin(1) + \cos(1) - 1) = \cos(1) + 2\sin(1) - 2.$$

La notation en tableau permet d'effectuer 3 intégrations par parties d'un coup :

$$\begin{array}{c|ccc} + & x^2 & & \sin(x) \\ & & \searrow & \\ - & 2x & & -\cos(x) \\ & & \searrow & \\ + & 2 & & -\sin(x) \\ & & \searrow & \\ - & 0 & \Rightarrow & \cos(x) \end{array} \quad , \quad \text{qui se lit} + [-x^2\cos(x)]_0^1 - [-2x\sin(x)]_0^1 + [2\cos(x)]_0^1 + 0.$$

3. Changement de variables

Enfin, on rappelle le calcul d'intégrales par changement de variables.

Proposition 2.1.4: Changement de variables dans les intégrales simples

Soit $f : [a,b] \to \mathbb{R}$ une fonction continue et $\psi : [c,d] \to [a,b]$ une fonction de classe C^1 telle que $\psi(c) = a$ et $\psi(d) = b$. Alors,

$$\int_a^b f(x)\, dx = \int_c^d f(\psi(t))\psi'(t)\, dt.$$

Remarque 2.1.2 On a donc effectué le changement de variables $x = \psi(t)$, l'élément d'intégration devient $dx = \frac{dx}{dt}\, dt = \psi'(t)\, dt$. Ici on ne suppose pas que ψ est une bijection, et on fait attention au signe de $\psi'(t)$ (contrairement à ce qu'on fera en dimension supérieure).

Exemple 2.1.4 Pour calculer $\int_0^1 \sqrt{1-x^2}\, dx$, on peut se rendre compte que l'équation $y = \sqrt{1-x^2}$ est l'équation sur $x \in [0,1]$ représente l'équation d'un quart de cercle, qui peut être paramétré par $\{x = \cos(t),\ y = \sin(t),\ t \in [0, \pi/2]\}$. On a donc l'idée du changement de

variable $x = \cos(t)$. Posons $\psi : [0, \pi/2] \to [0, 1]$ la fonction telle que $\psi(t) = \cos(t)$ pour tout $t \in [0, \pi/2]$. On a bien ψ de classe C^1 et $\psi(0) = 1$, $\psi(\pi/2) = 0$. Ainsi,

$$\int_0^1 \sqrt{1-x^2}\,dx = \int_{\pi/2}^0 \sqrt{1-\cos^2(t)}(-\sin(t))\,dt = \int_0^{\pi/2} \sqrt{1-\cos^2(t)}\sin(t)\,dt.$$

Pour tout $t \in [0, \pi/2]$, $\sin(t) \geq 0$ donc $\sqrt{1-\cos^2(t)} = \sin(t)$. On a donc

$$\int_0^{\pi/2} \sqrt{1-\cos^2(t)}\sin(t)\,dt = \int_0^{\pi/2} \sin^2(t)\,dt.$$

Maintenant, on peut écrire que $\sin^2(t) = \frac{1}{2}(1-\cos(2t))$ pour avoir

$$\int_0^{\pi/2} \sin^2(t)\,dt = \int_0^{\pi/2} \frac{1-\cos(2t)}{2}\,dt = \frac{\pi}{4} - \left[\frac{\sin(2t)}{4}\right]_{t=0}^{t=\pi/2} = \frac{\pi}{4}.$$

Exemple 2.1.5 On peut aussi utiliser la formule de changement de variable "dans l'autre sens", c'est-à-dire lire l'intégrale qu'on nous donne comme étant déjà de la forme $\int f(\psi(t))\psi'(t)\,dt$. Il s'agit alors d'identifier les bonnes fonctions f et ψ, de façon à ce que l'intégrale $\int f(x)\,dx$ soit plus simple à calculer. Par exemple, calculons

$$\int_0^1 t^3\cos(t^2)\,dt.$$

Vu cette intégrale, il est naturel de poser $x = t^2 = \psi(t)$ (donc $\psi : [0, 1] \to [0, 1]$), c'est-à-dire qu'on veut écrire cette intégrale comme $\int_0^1 f(\psi(t))\psi'(t)\,dt$ pour une certaine fonction f à déterminer. Comme $\psi(t) = t^2$, on a $\psi'(t) = 2t$, donc on veut que pour tout $t \in [0, 1]$:

$$t^3\cos(t^2) = f(\psi(t))\psi'(t) = f(t^2)\,2t\,.$$

En simplifiant, on tombe alors sur

$$\frac{t^2}{2}\cos(t^2) = f(t^2), \qquad \text{d'où on tire} \quad \frac{x}{2}\cos(x) = f(x), \quad x \in [0, 1].$$

On a donc

$$\int_0^1 t^3\cos(t^2)\,dt = \frac{1}{2}\int_0^1 x\cos(x)\,dx = \sin(1) + \cos(1) - 1,$$

comme vu plus haut.

Notons que cette procédure est équivalente à opérer le changement de variables initial $t = \varphi(x)$, avec $\varphi(x) = \psi^{(-1)}(x) = \sqrt{x}$.

2.2 Intégrales doubles

On va maintenant voir le concept d'intégrale de fonction de deux variables.

2.2.1 Définition, propriétés

Soit $f : \Omega \to \mathbb{R}$ une fonction continue, où Ω est un domaine de $\mathbb{R}^2$. On veut définir l'intégrale de f sur Ω. Géométriquement, elle représente le *volume* de la région située entre le graphe de f (qui est une surface dans $\mathbb{R}^3$) et le plan (Oxy) dans $\mathbb{R}^3$, comme représenté en Figure 2.3. Pour définir mathématiquement ce volume, on procède de manière analogue aux

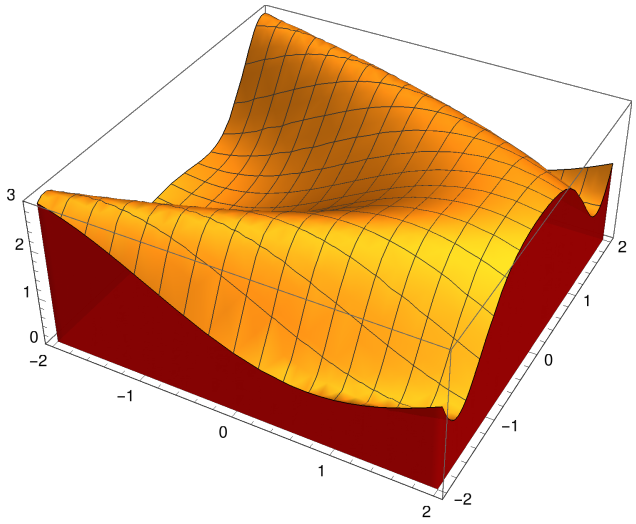

FIGURE 2.3 – L'intégrale en $2d$ représente le volume de la région en rouge sous le graphe

fonctions d'une variable : on découpe le domaine Ω en sous-domaines disjoints $\Omega_1, \ldots, \Omega_N$ d'aires plus petites qu'un petit nombre $\Delta\Omega > 0$. Ensuite, on choisit arbitrairement des points $M_1 \in \Omega_1, \ldots, M_2 \in \Omega_2, \ldots, M_N \in \Omega_N$. On montre alors que dans la limite $\Delta\Omega \to 0$ (et donc $N \to \infty$), la somme

$$\sum_{j=1}^{N} f(M_j)\,\mathrm{Aire}(\Omega_j)$$

possède une limite indépendante du choix du découpage de Ω en (Ω_j) et indépendante du choix des (M_j). On appelle cette limite *intégrale de f sur Ω*, notée

$$\iint_{\Omega} f(x,y)\,dx\,dy$$

ou aussi $\iint_\Omega f(x)\,d^2x$. Là encore, cette approximation est raisonnable car $f(M_j)\,\mathrm{Aire}(\Omega_j)$ représente l'aire d'un cylindre de base Ω_j et de hauteur $f(M_j)$ (c'est donc un parallélépipède rectangle si Ω_j est un rectangle). On approche donc le volume sous le graphe par une somme de volumes de petits cylindres (voir la Figure 2.4) qu'on sait calculer pourvu qu'on connaisse l'aire des Ω_j. Même si cette procédure paraît simple, elle est difficile à mettre en place mathématiquement car il faut définir l'aire des Ω_j, ce qu'on ne sait faire simplement que pour des Ω_j particuliers (des rectangles par exemples). On définit ainsi l'intégrale d'une fonction continue sur des domaines que l'on peut paver par des rectangles, puis on peut passer à des domaines plus généraux. Toute la difficulté réside donc dans le domaine d'intégration.

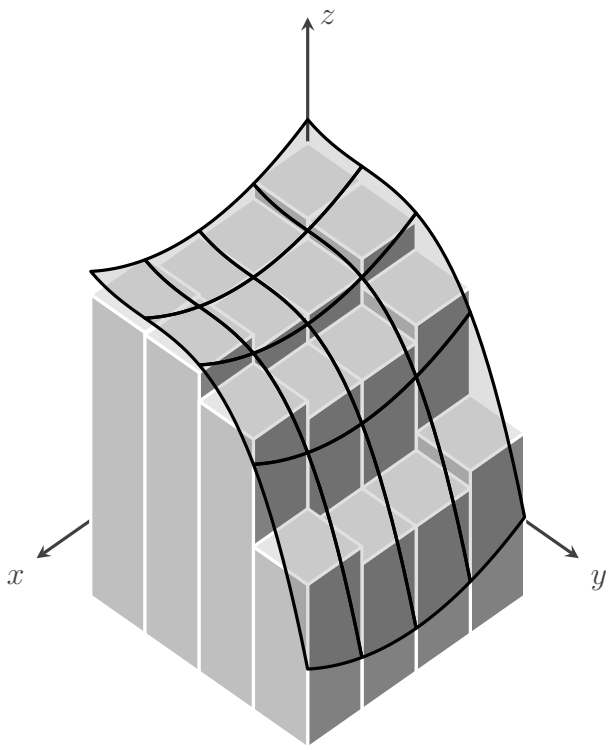

FIGURE 2.4 – Approximation du volume sous le graphe par une somme de volumes de parallélépipèdes.

On ne présente donc pas la définition rigoureuse d'intégrale de fonction de deux variables, mais on explique plutôt ses propriétés basiques ainsi que les méthodes de calcul associées. On commence par les propriétés "algébriques" de l'intégrale double.

Proposition 2.2.1

L'intégrale double possède les propriétés suivantes :

1. Linéarité : si $f, g : \Omega \to \mathbb{R}$ sont continues et $\lambda, \mu \in \mathbb{R}$, on a
$$\iint_\Omega (\lambda f(x,y) + \mu g(x,y))\, dx\, dy = \lambda \iint_\Omega f(x,y)\, dx\, dy + \mu \iint_\Omega g(x,y)\, dx\, dy.$$
2. Positivité : si $f \geq 0$ alors $\iint_\Omega f(x,y)\, dx\, dy \geq 0$;
3. Relation de Chasles : si $\Omega = \Omega_1 \cup \Omega_2$ et si $\Omega_1 \cap \Omega_2 = \emptyset$ (ou plus généralement si $\Omega_1 \cap \Omega_2$ est d'aire nulle), on a
$$\iint_\Omega f(x,y)\, dx\, dy = \iint_{\Omega_1} f(x,y)\, dx\, dy + \iint_{\Omega_2} f(x,y)\, dx\, dy.$$

Remarque 2.2.1 Si $f \equiv 1$ est la fonction constante égale à 1 sur un domaine Ω, alors

$$\iint_\Omega 1\, dx\, dy = \text{Aire}(\Omega)$$

représente l'aire du domaine Ω. Les intégrales doubles sont donc notamment un outil de calcul d'aires.

2.2.2 Calcul des intégrales doubles

Les propriétés que l'on a énoncées ne permettent pas réellement de calculer des intégrales doubles. On va donc à présent énoncer des propriétés de l'intégrale double qui permettent de ramener leur calcul à celui de deux intégrales simples successives. Comme on va le voir, cela va dépendre de la forme du domaine d'intégration.

Si le domaine d'intégration est un rectangle

Le premier exemple fondamental concerne le cas où Ω est un rectangle. Dans ce cas, on se ramène aisément au calcul d'intégrales simples.

Proposition 2.2.2: Théorème de Fubini $2d$

Soit $\Omega = [a, b] \times [c, d]$ un rectangle de $\mathbb{R}^2$ et $f : \Omega \to \mathbb{R}$ une fonction continue. Alors,

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

Remarque 2.2.2 Ce résultat signifie que l'on peut calculer l'intégrale sur un rectangle de deux manières (i) en gelant d'abord x et en intégrant d'abord sur y, puis en intégrant en x le résultat ou (ii) en gelant d'abord y et en intégrant sur x, puis en intégrant le résultat sur y. Dans tous les cas, on se ramène au calcul de deux intégrales $1d$ successives.

Remarque 2.2.3 Ce résultat est en fait valable pour des domaines plus généraux : une intégrale $2d$ peut se calculer en gelant l'une des variables, en intégrant par rapport à l'autre, puis en intégrant le résultat obtenu par rapport à la variable qui avait été gelée. La difficulté réside dans les domaines d'intégration à choisir. Si par exemple on gèle x d'abord, il faut déterminer l'ensemble des y pour lesquels $(x, y) \in \Omega$ (à x fixé). Cet ensemble peut-être très compliqué (et même pas forcément un segment ou une union de segments). Cette stratégie de calcul n'est donc pas forcément facile à mettre en place si le domaine d'intégration est compliqué. Dans le cas particulier d'un rectangle, l'ensemble des y pour lesquels $(x, y) \in \Omega$ à x fixé est très simple : c'est un segment indépendant de x ! C'est pourquoi ce cas est favorable.

Remarque 2.2.4 Dans le cas particulier où la fonction f est un produit de fonctions d'une variables : $f(x, y) = g(x)h(y)$ pour deux fonctions continues g et h, ce résultat implique l'égalité

$$\iint_\Omega g(x)h(y)\, dx\, dy = \left(\int_a^b g(x)\, dx \right) \left(\int_c^d h(y)\, dy \right).$$

Exemple 2.2.1 Calculons sur $\Omega = [0,1] \times [1,2]$ l'intégrale

$$\iint_\Omega (x + xy)\, dx\, dy.$$

Par le théorème de Fubini, on peut d'abord intégrer en y, puis en x :

$$\begin{aligned}
\iint_\Omega (x + xy)\, dx\, dy &= \int_0^1 \left(\int_1^2 (x + xy)\, dy \right) dx \\
&= \int_0^1 \left([xy + xy^2/2]_{y=1}^{y=2} \right) dx \\
&= \int_0^1 \left(x + \frac{3}{2}x \right) dx \\
&= \frac{5}{2}[x^2/2]_{x=0}^{x=1} \\
&= \frac{5}{4}.
\end{aligned}$$

On aurait pu aussi d'abord intégrer en x puis intégrer en y :

$$\begin{aligned}
\iint_\Omega (x + xy)\, dx\, dy &= \int_1^2 \left(\int_0^1 (x + xy)\, dx \right) dy \\
&= \int_1^2 \left([(1+y)x^2/2]_{x=0}^{x=1} \right) dy \\
&= \frac{1}{2} \int_1^2 (1 + y)\, dy \\
&= \frac{1}{2}[(1+y)^2/2]_{y=1}^{y=2} \\
&= \frac{5}{4}.
\end{aligned}$$

Exemple 2.2.2 Calculons sur $\Omega = [0, \pi/2]^2$ l'intégrale

$$\iint_\Omega \sin(x + y)\, dx\, dy\,.$$

Par le théorème de Fubini $2d$, on a

$$\begin{aligned}
\iint_\Omega \sin(x + y)\, dx\, dy &= \int_0^{\pi/2} \left(\int_0^{\pi/2} \sin(x + y)\, dx \right) dy \\
&= \int_0^{\pi/2} [-\cos(x + y)]_{x=0}^{x=\pi/2}\, dy \\
&= \int_0^{\pi/2} (\cos(y) - \cos(\pi/2 + y))\, dy \\
&= [\sin(y) - \sin(\pi/2 + y)]_{y=0}^{y=\pi/2} \\
&= 2.
\end{aligned}$$

Si le domaine d'intégration est compris entre deux graphes

Une variante du théorème de Fubini que l'on vient de voir concerne le cas plus général où x (ou y) est compris "entre deux graphes".

Proposition 2.2.3

1. Soit Ω un domaine de $\mathbb{R}^2$ de la forme

$$\Omega = \{(x,y) \in \mathbb{R}^2 \ : \ a \leq x \leq b, \ c(x) \leq y \leq d(x)\},$$

où $c, d : [a,b] \to \mathbb{R}$ sont des fonctions continues. Alors pour toute fonction $f : \Omega \to \mathbb{R}$ continue, on a

$$\iint_\Omega f(x,y)\,dx\,dy = \int_a^b \left(\int_{c(x)}^{d(x)} f(x,y)\,dy \right) dx.$$

2. Soit Ω un domaine de $\mathbb{R}^2$ de la forme

$$\Omega = \{(x,y) \in \mathbb{R}^2 \ : \ c \leq y \leq d, \ a(y) \leq x \leq b(y)\},$$

où $a, b : [c,d] \to \mathbb{R}$ sont des fonctions continues. Alors pour toute fonction $f : \Omega \to \mathbb{R}$ continue, on a

$$\iint_\Omega f(x,y)\,dx\,dy = \int_c^d \left(\int_{a(y)}^{b(y)} f(x,y)\,dx \right) dy.$$

Remarque 2.2.5 Dans le premier cas, la forme de Ω indique qu'il est naturel d'intégrer d'abord par rapport à y (dont le domaine d'intégration $[c(x), d(x)]$ varie en fonction de x). Pour une telle forme de Ω, il n'est pas évident d'échanger l'ordre des intégrales : à y fixé, l'ensemble des x pour lesquels $(x,y) \in \Omega$ n'est pas donné immédiatement, et pourrait être très compliqué. Il faut donc choisir astucieusement l'ordre dans lequel appliquer le théorème de Fubini.

GD : L'aire du domaine Ω, c'est aussi l'intégrale de la fonction 1 sur Ω. Plus haut on a dit que si $f : [a,b] \to \mathbb{R}$ est continue et positive, $\int_a^b f(x)dx$ est aussi l'aire comprise entre l'axe des x et le graphe de f (et aussi les droites verticales d'équations $x = a$ et $x = b$). Si on pense qu'on calcule l'aire de cette région Ω comme étant l'intégrale double de 1 sur Ω, on retrouve le même résultat, à savoir

$$\iint_\Omega 1\,dxdy = \int_a^b \int_0^f (x) 1\,dydx = \int_a^b f(x)dx,$$

en appliquant la formule de Fubini pour calculer l'intégrale double.

Exemple 2.2.3 Calculons sur le domaine $\Omega = \{(x,y) \in \mathbb{R}^2 \ : \ 0 \leq x \leq 1, \ 0 \leq y \leq x^2\}$, l'intégrale

$$\iint_\Omega x\,dx\,dy.$$

La région Ω est représentée sur la Figure 2.5. Par le résultat précédent (avec $[a,b]=[0,1]$ et

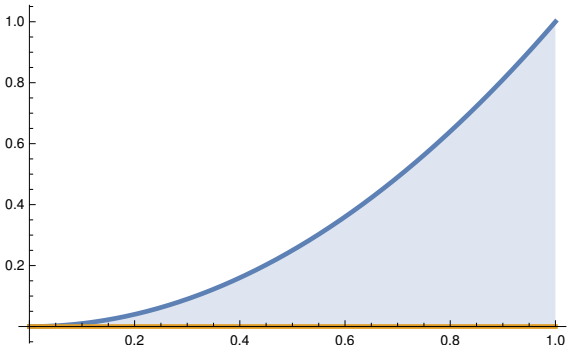

FIGURE 2.5 – Région Ω (bleu clair) entre les graphes de $x \mapsto x^2$ (bleu) et de $x \mapsto 0$ (jaune).

$c(x)=0$, $d(x)=x^2$), on peut utiliser Fubini comme suit :

$$\begin{aligned}\iint_\Omega x\,dx\,dy &= \int_0^1 \left(\int_0^{x^2} x\,dy\right) dx = \int_0^1 x\left(\int_0^{x^2} 1\,dy\right) dx \\ &= \int_0^1 x.x^2\,dx = \int_0^1 x^3\,dx = \frac{1}{4}.\end{aligned}$$

Dans ce cas, on aurait pu échanger les intégrales : en effet, $0 \le x \le 1$ et $0 \le y \le x^2$ est équivalent à $0 \le y \le 1$ et $\sqrt{y} \le x \le 1$, donc

$$\begin{aligned}\iint_\Omega x\,dx\,dy &= \int_0^1 \left(\int_{\sqrt{y}}^1 x\,dx\right) dy \\ &= \int_0^1 (1-y)/2\,dy \\ &= \frac{1}{4}.\end{aligned}$$

Exemple 2.2.4 Calculons

$$\iint_\Omega (x+y)\,dx\,dy$$

où Ω est l'intérieur du triangle ABC, avec $A=(0,0)$, $B=(1,2)$, $C=(1,-1)$, représenté en Figure 2.6.

Ici la difficulté réside dans le fait d'expliciter le triangle ABC en coordonnées cartésiennes. Le segment $[AB]$ correspond à la portion de la droite d'équation $y=2x$ pour $0 \le x \le 1$, et le segment $[AC]$ correspond à la portion de la droite d'équation $y=-x$ pour $0 \le x \le 1$. Ainsi, Ω peut se réécrire comme

$$\Omega = \{(x,y) \in \mathbb{R}^2 \ : \ 0 \le x \le 1,\ -x \le y \le 2x\}.$$

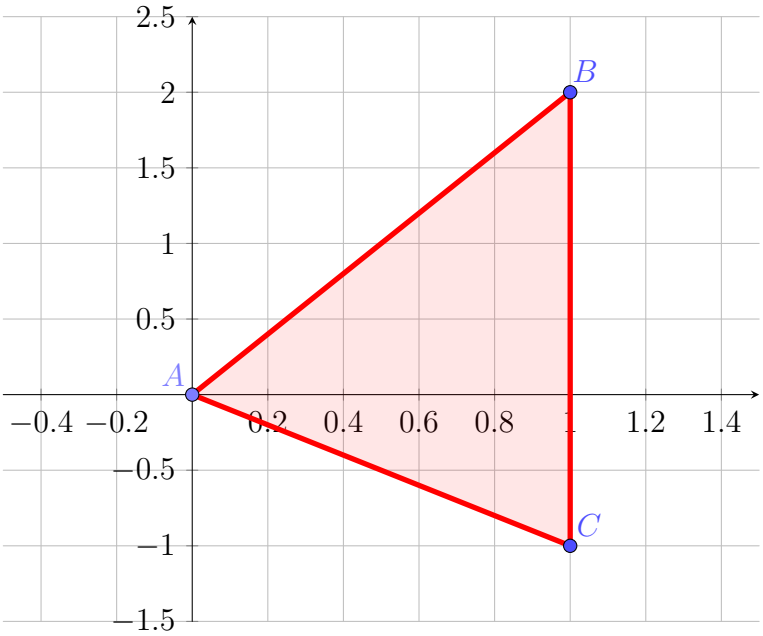

FIGURE 2.6 – Le triangle ABC

On peut donc calculer

$$\begin{aligned}\iint_\Omega (x+y)\,dx\,dy &= \int_0^1 \left(\int_{-x}^{2x} (x+y)\,dy\right) dx \\ &= \int_0^1 [(x+y)^2/2]_{y=-x}^{y=2x}\,dy \\ &= \frac{9}{2}\int_0^1 x^2\,dx \\ &= \frac{3}{2}.\end{aligned}$$

De manière générale, lorsqu'on paramètre ainsi un triangle, il est utile de placer l'un des sommets en l'origine.

Calcul par changement de variables

On peut aussi changer de variables dans les intégrales $2d$. L'intérêt d'un changement de variable dépend à la fois de la forme Ω (qui peut s'exprimer plus facilement dans les nouvelles variables, par exemple comme un rectangle) et de la fonction f.

Proposition 2.2.4: Changement de variables dans les intégrales $2d$

Soient Ω et U des domaines de $\mathbb{R}^2$, et $\psi : U_{u,v} \to \Omega_{x,y}$ une fonction de classe C^1, bijective.

Soit $f : \Omega \to \mathbb{R}$ une fonction continue. Alors, on a

$$\iint_\Omega f(x,y)\,dx\,dy = \iint_U f\big(\psi(u,v)\big)\,J_\psi(u,v)\,du\,dv,$$

où $J_\psi(u,v)$ désigne le *jacobien* de ψ au point $(u,v) \in U$, défini par

$$J_\psi(u,v) = |\det \mathrm{Jac}\,\psi(u,v)|,$$

la valeur absolue du déterminant de la matrice jacobienne de ψ au point (u,v).

Remarque 2.2.6 Ici, on a changé de variables en "posant $(x,y) = \psi(u,v)$".

GD : Il y a donc deux notions liées mais un peu différentes : la *matrice jacobienne*, qui est la matrice des dérivées partielles de ψ, et le *jacobien*, qui est la valeur absolue du déterminant de cette matrice.

Remarquons aussi qu'en 2d et 3d, on ne considérera que des changements de variables *bijectifs*. C'est plus simple, et cela évite de devoir de calculer des aires algébriques de domaines, comme lorsqu'on définissait $\int_b^a = -\int_a^b$ en 1D.

Exemple 2.2.5 Un exemple fondamental de changement de variables en $2d$ est donné par les coordonnées polaires

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

La fonction ψ est manifestement C^1, par contre pour qu'elle soit bijective il est nécessaire de la restreindre à $(r,\theta) \in]0,+\infty[\times[0,2\pi[$, ou à un sous-ensemble. Pour tout (r,θ), on a

$$\mathrm{Jac}\,\psi(r,\theta) = \begin{bmatrix}\cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta\end{bmatrix}, \quad \text{donc} J_\psi(r,\theta) = r(\cos^2\theta + \sin^2\theta) = r.$$

Si on veut calculer

$$\iint_\Omega x^2\,dx\,dy, \quad \text{avec } \Omega \text{ le disque centré en } (0,0) \text{ de rayon } R > 0,$$

on définit donc

$$U = \{(r,\theta)\ :\ 0 < r \le R,\ 0 \le \theta < 2\pi\} \quad \text{un rectangle dans les coordonnées polaires,}$$

et $\psi : U \to \Omega \setminus \{(0,0)\}$ est alors une bijection C^1. Par changement de variables $(x,y) = \psi(r,\theta)$, on a donc

$$\begin{aligned}\iint_\Omega x^2\,dx\,dy &= \int_0^R \int_0^{2\pi} (r\cos\theta)^2 r\,d\theta\,dr \\ &= \left(\int_0^R r^3\,dr\right)\left(\int_0^{2\pi}\cos^2\theta\,d\theta\right) \\ &= \frac{R^4}{4}\int_0^{2\pi}\frac{1+\cos 2\theta}{2}\,d\theta \\ &= \frac{\pi}{4}R^4.\end{aligned}$$

Exemple 2.2.6 Pour calculer l'aire de l'ellipse

$$\Omega = \{(x, y) \in \mathbb{R}^2 \ : \ \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 \leq 1\},$$

de paramètres $a, b > 0$ (représentée en Figure 2.7), il est naturel de définir une variante des

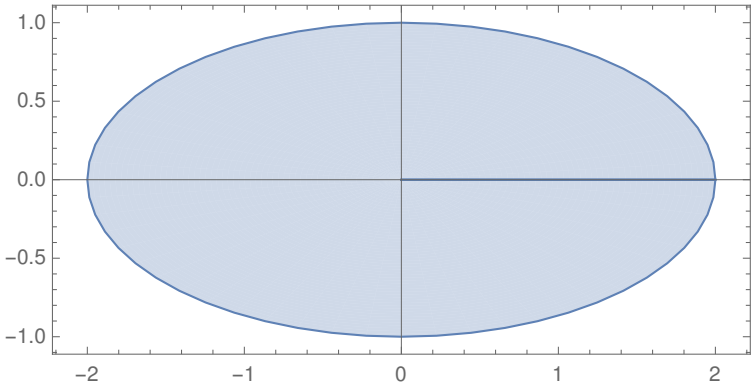

FIGURE 2.7 – L'ellipse Ω avec $a = 2$ et $b = 1$.

coordonnées polaires :

$$\psi(r, \theta) = (ar\cos\theta, br\sin\theta),$$

de telle manière à ce que $\psi :]0, 1] \times [0, 2\pi[\to \Omega \setminus \{(0, 0)\}$ soit une bijection C^1. Le jacobien de ψ vaut alors

$$J_\psi(r, \theta) = \left|\det \begin{bmatrix} a\cos\theta & -ar\sin\theta \\ b\sin\theta & br\cos\theta \end{bmatrix}\right| = |abr(\cos^2\theta + \sin^2\theta)| = abr.$$

Ainsi,

$$\begin{aligned} \text{Aire}(\Omega) &= \iint_\Omega dx\, dy \\ &= \iint_U J_\psi(r, \theta)\, dr\, d\theta \\ &= \int_0^{2\pi} \int_0^1 abr\, dr\, d\theta \\ &= \pi ab. \end{aligned}$$

Dans le cas où $a = b = R$, on retrouve l'aire d'un disque de rayon R : πR^2.

On aurait bien sûr pu calculer cette aire en utilisant la représentation $\Omega = \{|y| \leq b\sqrt{1 - \left(\frac{x}{a}\right)^2}\}$ de surface comprise entre deux graphes ; le calcul d'intégrale aurait alors été un peu plus compliqué, et aurait nécessité de toute façon un changement de variables.

Exemple 2.2.7 Comme pour les intégrales $1d$, on peut vouloir appliquer la formule de changement de variables de "droite à gauche" plutôt que de "gauche à droite". Par exemple,

calculons, sur le losange Ω reliant les points $(0,0)$, $(2,1)$, $(4,0)$ et $(2,-1)$, l'intégrale

$$\iint_\Omega (x-2y)^7 e^{x+2y}\,dx\,dy.$$

Vu la forme du losange, il est naturel de poser $(u,v)=\varphi(x,y)=(x-2y,x+2y)$: en effet, dans ces coordonnées, le losange est transformé en carré $(u,v)\in U=[0,4]^2$, qui permettra un calcul plus simple par Fubini. La fonction dans l'intégrale est également simple à exprimer dans les coordonnées (u,v).

$$(x-2y)^7 e^{x+2y} = f(\varphi(x,y))J_\varphi(x,y),$$

autrement dit

$$f(u,v)=\frac{u^7e^v}{J_\varphi(\varphi^{-1}(u,v))}.$$

Cette démarche est donc valide si le jacobien de φ ne s'annule pas et si φ est une bijection C^1. On trouve que

$$J_\varphi(x,y)=\left|\det\begin{bmatrix}1 & -2\\ 1 & 2\end{bmatrix}\right|=4,$$

donc le jacobien de φ ne s'annule pas. φ est une application linéaire dont le déterminant est non nul, et donc est bijective. On définit donc f par

$$f(u,v)=\frac{1}{4}u^7e^v.$$

Enfin,

$$\begin{aligned}\iint_\Omega (x-2y)^7e^{x+2y}\,dx\,dy &= \iint_\Omega f(\varphi(x,y))J_\varphi(x,y)\,dx\,dy\\ &= \iint_U f(u,v)\,du\,dv\\ &= \int_0^4 \frac{1}{4}u^7\,du\int_0^4 e^v\,dv\\ &= \frac{1}{4}\frac{4^8}{8}(e^4-1)=2048(e^4-1).\end{aligned}$$

Exemple 2.2.8 On peut également calculer l'aire d'un secteur curviligne à l'aide d'un changement de variables en coordonnées polaires. Supposons que l'on veuille calculer l'aire du domaine

$$\Omega=\{(r\cos\theta,r\sin\theta)\in\mathbb{R}^2\ :\ \theta_1\le\theta\le\theta_2,\ 0<r\le R(\theta)\},$$

pour deux angles $\theta_1<\theta_2$ et une fonction continue $R:[\theta_1,\theta_2]\to\mathbb{R}_+$. En coordonnées polaires, on a $\Omega=\psi(U)$ avec

$$U=\{(r,\theta)\ :\ \theta_1\le\theta\le\theta_2,\ 0<r\le R(\theta)\}$$

et donc

$$\begin{aligned}\text{Aire}(\Omega) &= \iint_{\Omega} dx\,dy \\ &= \iint_{U} r\,dr\,d\theta \\ &= \int_{\theta_1}^{\theta_2}\int_{0}^{R(\theta)} r\,dr\,d\theta \\ &= \int_{\theta_1}^{\theta_2} \frac{R(\theta)^2}{2}\,d\theta.\end{aligned}$$

Par exemple, pour $\theta_1 = -\pi/4$, $\theta_2 = \pi/4$, et $R(\theta) = \cos 2\theta$ (représentée en Figure 2.8), on a

$$\text{Aire}(\Omega) = \frac{1}{2}\int_{-\pi/4}^{\pi/4} \cos^2(2\theta)\,d\theta = \frac{1}{2}\int_{-\pi/4}^{\pi/4} \frac{1+\cos 4\theta}{2}\,d\theta = \frac{\pi}{8}.$$

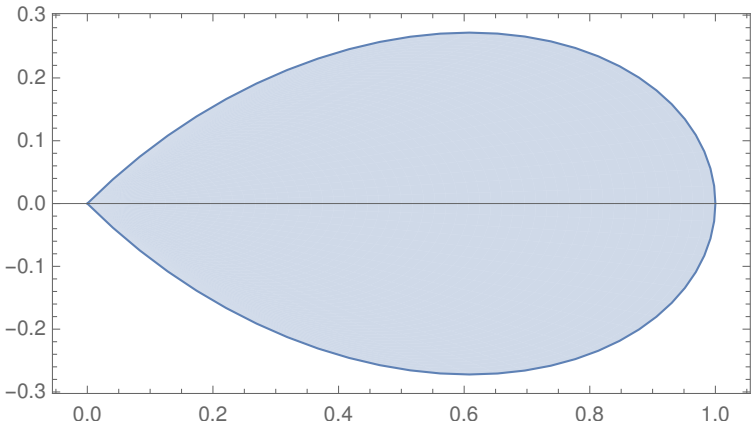

FIGURE 2.8 – Représentation de la région Ω

2.3 Intégrales triples

On passe maintenant aux intégrales des fonctions de trois variables, qui se définissent de manière identique au cas de deux variables.

2.3.1 Définition, propriétés

On part donc d'une fonction $f : \Omega \to \mathbb{R}$ continue, où Ω est cette fois-ci un domaine de $\mathbb{R}^3$, et on souhaite définir l'intégrale de f sur Ω. Là encore, l'intégrale mesure un "hypervolume" : celui compris entre l'hyperplan $(Oxyz)$ de $\mathbb{R}^4$ et le graphe de f (qui est une hypersurface de $\mathbb{R}^4$, d'équation $t = f(x, y, z)$). On procède de la même manière qu'en dimension inférieure : on découpe Ω en sous-domaines $\Omega_1, \ldots, \Omega_N$ de volumes au plus $\Delta\Omega$, où $\Delta\Omega > 0$ est très

petit. On choisit de plus des points $M_1 \in \Omega_1, M_2 \in \Omega_2, \ldots, M_N \in \Omega_N$ arbitrairement. On montre alors que la somme

$$\sum_{j=1}^{N} f(M_j)\,\text{Volume}(\Omega_j)$$

possède une limite lorsque $\Delta\Omega \to 0$ (et donc $N \to +\infty$). Cette limite est de plus indépendante du découpage de Ω choisi, et du choix des points (M_j). Elle est appelée intégrale de f sur Ω, et notée

$$\iiint_\Omega f(x,y,z)\,dx\,dy\,dz\,.$$

Là encore, on ne détaille pas sa définition mathématique précise, car il peut être compliqué de définir le volume des petits domaines (Ω_j). On se contente donc d'énoncer les propriétés de cette intégrale triple, et la manière de la calculer en pratique. Par exemple, l'intégrale triple vérifie des propriétés analogues à celle de la Proposition 2.2.1, ainsi que

$$\iiint_\Omega 1\,dx\,dy\,dz = \text{Volume}(\Omega).$$

On se focalise à présent sur les méthodes de calcul des intégrales triples.

2.3.2 Calcul des intégrales triples

Le premier outil de calcul des intégrales triples consiste à se ramener à des intégrales simples ou doubles. On a des versions $3d$ du théorème de Fubini : si Ω est un pavé droit (la version tridimensionnelle d'un rectangle),

$$\Omega = [a,b] \times [c,d] \times [e,f]$$

et que $g : \Omega \to \mathbb{R}$ est continue, alors

$$\iiint_\Omega g(x,y,z)\,dx\,dy\,dz = \int_a^b \left(\int_c^d \left(\int_e^f g(x,y,z)\,dz \right) dy \right) dx,$$

et on peut même choisir un ordre quelconque d'intégration.

Si maintenant Ω est de la forme

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 \ : \ (x,y) \in D,\ a(x,y) \le z \le b(x,y)\},$$

où $D \subset \mathbb{R}^2$ est un domaine de $\mathbb{R}^2$ et les fonctions $a, b : D \to \mathbb{R}$ sont continues (c'est-à-dire que Ω est compris entre les graphes de deux fonctions sur D), alors

$$\iiint_\Omega g(x,y,z)\,dx\,dy\,dz = \iint_D \left(\int_{a(x,y)}^{b(x,y)} g(x,y,z)\,dz \right) dx\,dy :$$

on se ramène alors au calcul d'intégrale $1d$ puis d'une intégrale $2d$. Ici le choix de fixer (x,y) et d'intégrer selon z est dû à la forme particulière de Ω.

Exemple 2.3.1 Calculons

$$\iiint_\Omega (x+y+z)^2\,dx\,dy\,dz,$$

où $\Omega = [0,1]\times[-1,1]\times[0,2]$ est un pavé droit. par Fubini, on a

$$\begin{aligned}
\iiint_\Omega (x+y+z)^2\,dx\,dy\,dz &= \int_0^1\int_{-1}^1\Big(\int_0^2 (x+y+z)^2\,dz\Big)\,dy\,dx\\
&= \int_0^1\int_{-1}^1 [(x+y+z)^3/3]_{z=0}^{z=2}\,dy\,dx\\
&= \frac{1}{3}\int_0^1\Big(\int_{-1}^1 ((x+y+2)^3-(x+y)^3)\,dy\Big)\,dx\\
&= \frac{1}{3}\int_0^1 [(x+y+2)^4/4-(x+y)^4/4]_{y=-1}^{y=1}\,dx\\
&= \frac{1}{12}\int_0^1 ((x+3)^4-2(x+1)^4+(x-1)^4)\,dx\\
&= \frac{1}{60}[(x+3)^5-2(x+1)^5+(x-1)^5]_{x=0}^{x=1}\\
&= \frac{1}{60}(4^5-2\times 2^5-3^5+2+1) = 12.
\end{aligned}$$

On remarque qu'il n'a pas été nécessaire de développer le polynôme $(x+y+z)^2$ pour faire le calcul (il est même recommandé de ne pas le développer, pour éviter des expressions trop longues).

Exemple 2.3.2 Calculons

$$\iiint_\Omega z\,dx\,dy\,dz$$

où Ω a la forme d'un biseau représentée en Figure 2.9 :

$$\Omega = \{(x,y,z)\in\mathbb{R}^3\ :\ \underbrace{0\le x\le 1,\ 0\le y\le 2}\,D,\ \underbrace{0}\,a(x)\le z\le \underbrace{1-x}\,b(x)\}.$$

Par Fubini on intègre d'abord sur la coordonnée z :

$$\begin{aligned}
\iiint_\Omega z\,dx\,dy\,dz &= \int_0^1\int_0^2\Big(\int_0^{1-x} z\,dz\Big)\,dy\,dx\\
&= \frac{1}{2}\int_0^1\int_0^2 (1-x)^2\,dy\,dx\\
&= \int_0^1 (1-x)^2\,dx = \frac{1}{3}.
\end{aligned}$$

On a aussi une notion de changement de variables en $3d$.

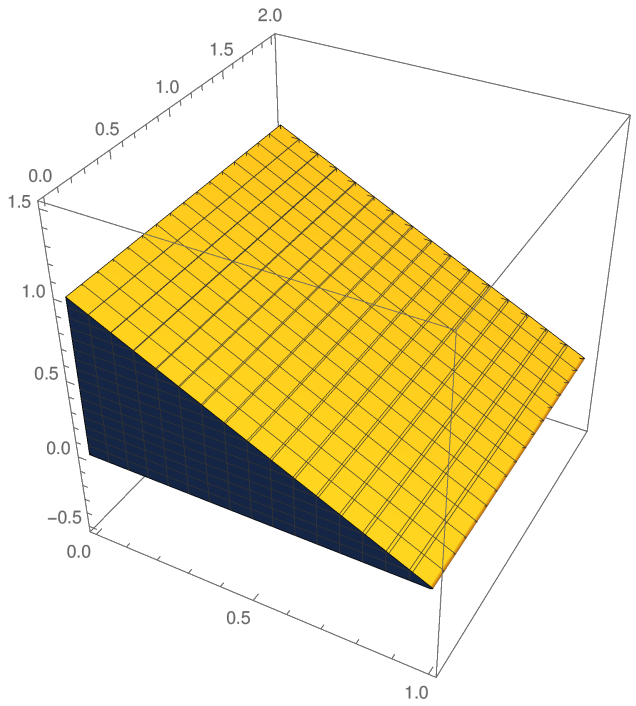

FIGURE 2.9 – Le "biseau" Ω

Proposition 2.3.1: Changement de variables dans les intégrales triples

Soient $\Omega_{x,y,z}$ et $U_{u,v,w}$ des domaines de $\mathbb{R}^3$, et $\psi : U \to \Omega$ une fonction de classe C^1, bijective. Soit $f : \Omega \to \mathbb{R}$ une fonction continue. Alors, on a

$$\iiint_{\Omega} f(x,y,z)\,dx\,dy\,dz = \iiint_{U} f(\psi(u,v,w))J_{\psi}(u,v,w)\,du\,dv\,dw,$$

où $J_{\psi}(u,v,w)$ désigne le *jacobien* de ψ au point $(u,v,w) \in U$, défini par

$$J_{\psi}(u,v,w) = |\det \mathrm{Jac}\,\psi(u,v,w)|,$$

la valeur absolue du déterminant de la matrice jacobienne de ψ au point (u,v,w).

Rappels sur les déterminants

Le déterminant de la matrice 2×2

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

est juste $ad - bc$, et pour le jacobien en dimension 2 on a pris $\left|\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}\right| = |ad - bc|$. La simplicité de cette formule cache les propriétés fondamentales des déterminants, que je vous conseille de regarder, et dont je ne vais rappeler que certaines. On regarde des matrices

carrées $n \times n$, mais on va prendre $n = 3$ pour simplifier. Et je n'explique pas non plus pourquoi toutes ces propriétés sont vraies.

1. Le développement qui permet de calculer $\det(M)$. Disons que

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

et écrivons le développement par rapport à la première ligne :

$$\det M = a \det \begin{bmatrix} e & f \\ h & i \end{bmatrix} - b \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + c \det \begin{bmatrix} d & e \\ g & h \end{bmatrix}.$$

Donc chaque élément de la matrice est affecté d'un signe (1 en haut à gauche, et qui change chaque fois qu'on se déplace de 1 vers la droite ou vers le bas), et on multiplie l'élément de la ligne par ce signe et par le déterminant de la matrice 2×2 qui reste quand on a ôté de M la ligne et la colonne qui contenaient l'élément.

On peut développer par rapport à n'importe quelle ligne, et aussi par rapport à n'importe quelle colonne. Je donne juste le développement par la seconde colonne ; j'espère que vous en déduirez comment on fait les autres :

$$\det M = -b \det \begin{bmatrix} d & f \\ g & i \end{bmatrix} + e \det \begin{bmatrix} a & c \\ g & i \end{bmatrix} - h \det \begin{bmatrix} a & c \\ d & f \end{bmatrix}.$$

2. Le déterminant de la matrice transposée de M est égal au déterminant de M. La matrice transposée, c'est

$$^tM = \begin{bmatrix} a & d & g \\ b & e & h \\ c & f & i \end{bmatrix}.$$

3. Le déterminant est une fonction linéaire de chaque ligne quand on fixe les deux autres. Et pareil pour les colonnes. Donc si on multiplie M par 2, le déterminant est multiplié par 8.
4. $\det(M) \neq 0$ si et seulement si M est inversible (donc si et seulement si les trois colonnes de M sont des vecteurs indépendants, et si et seulement si les trois lignes de M sont des vecteurs indépendants). C'est très pratique, et c'est en rapport avec la propriété suivante.
5. Le déterminant d'un produit de deux matrices $n \times n$ est le produit des déterminants : $\det(MN) = \det(M) \det(N)$.

Exemple 2.3.3 Un exemple fondamental de changement de variables $3d$ est donné par les coordonnées sphériques. Elles sont données par

$$(x, y, z) = \psi(r, \theta, \varphi) = (r \cos\varphi \sin\theta, r \sin\varphi \sin\theta, r \cos\theta),$$

et ψ est une bijection C^1 de $\mathbb{R}_+^*\times]0,\pi[\times[0,2\pi[$ dans $\mathbb{R}^3\setminus\{(0,0,z) : z\in\mathbb{R}\}$. Son jacobien vaut

$$\begin{aligned} J_\psi(r,\theta,\varphi) &= \left|\det\begin{bmatrix}\cos\varphi\sin\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta\\ \sin\varphi\sin\theta & r\sin\varphi\cos\theta & r\cos\varphi\sin\theta\\ \cos\theta & -r\sin\theta & 0\end{bmatrix}\right| \\ &= \cos\theta\times r\cos\theta\sin\theta + r\sin\theta\times r\sin^2\theta + 0 \\ &= r^2\sin\theta \end{aligned}$$

(on a développé par rapport à la 3e ligne). Si on veut par exemple calculer, sur $\Omega = B(0,R)$ la boule centrée en l'origine de rayon $R>0$, l'intégrale

$$I = \iiint_\Omega (x^2+y^2+z^2)\,dx\,dy\,dz,$$

on utilise les coordonnées sphériques $\psi : U\to\Omega\setminus\{(0,0,z) : z\in\mathbb{R}\}$, où

$$U =]0,R]\times]0,\pi[\times[0,2\pi[\quad \text{est un pavé droit,}$$

afin que ψ soit une bijection (le fait de retirer l'axe vertical $(Oz) = \{(0,0,z) : z\in\mathbb{R}\}$ est nécessaire pour avoir la bijectivité de ψ, et heureusement ne change pas la valeur de l'intégrale). En coordonnées sphériques, on a

$$r^2 = x^2+y^2+z^2,$$

de sorte que l'intégrale s'écrit

$$\begin{aligned} I &= \iiint_U r^2(r^2\sin\theta)\,dr d\theta d\varphi = \int_0^{2\pi}\int_0^\pi\int_0^R r^2\times r^2\sin\theta\,dr\,d\theta\,d\varphi \\ &= 2\pi\int_0^1 r^4\,dr\int_0^\pi \sin\theta\,d\theta = 2\pi\times\frac{1}{5}\times 2 = \frac{4\pi}{5}. \end{aligned}$$

Ici on a pu facilement appliquer Fubini, puisque U est un pavé droit et l'intégrand en (r,θ,φ) se factorise en un produit de fonctions.

Exemple 2.3.4 En trois dimensions, il existe un autre type de changement de variables utile : les coordonnées cylindriques. Elles sont définies par la relation

$$(x,y,z) = \psi(\rho,\varphi,z) = (\rho\cos\varphi,\rho\sin\varphi,z),$$

et ψ est une bijection C^1 de $\mathbb{R}_+^*\times[0,2\pi[\times\mathbb{R}$ dans $\mathbb{R}^3\setminus(Oz)$. Là encore, on doit retirer l'axe vertical (Oz) afin d'assurer la bijectivité. Le jacobien de ce changement de variables vaut

$$J_\psi(\rho,\varphi,z)\left|\det\begin{bmatrix}\cos\varphi & -\rho\sin\varphi & 0\\ \sin\varphi & \rho\cos\varphi & 0\\ 0 & 0 & 1\end{bmatrix}\right| = \rho.$$

Par exemple, si on veut calculer le volume du domaine Ω délimité par le cylindre d'équation $x^2 + y^2 = 1$ et les plans $z = 0$ et $z = 1 - x$, l'application ψ est une bijection C^1 de U dans $\Omega \setminus (Oz)$ (là encore, retirer l'origine ne modifie pas la valeur du volume), où

$$U = \{(\rho, \varphi, z) \ : \ 0 \leq \rho \leq 1, \ 0 \leq z \leq 1 - r\cos\varphi, \ 0 \leq \varphi < 2\pi\}.$$

Ainsi, on a

$$\begin{aligned} \text{Volume}(\Omega) &= \int_0^1 \int_0^{2\pi} \int_0^{1-\rho\cos\varphi} dz\, d\varphi\, \rho\, d\rho \\ &= \int_0^1 \int_0^{2\pi} (1 - \rho\cos\varphi)\, d\varphi\, \rho\, d\rho = 2\pi \int_0^1 \rho\, d\rho = \pi. \end{aligned}$$

La région Ω est représentée dans la Figure 2.10, il s'agit d'un cylindre biseauté.

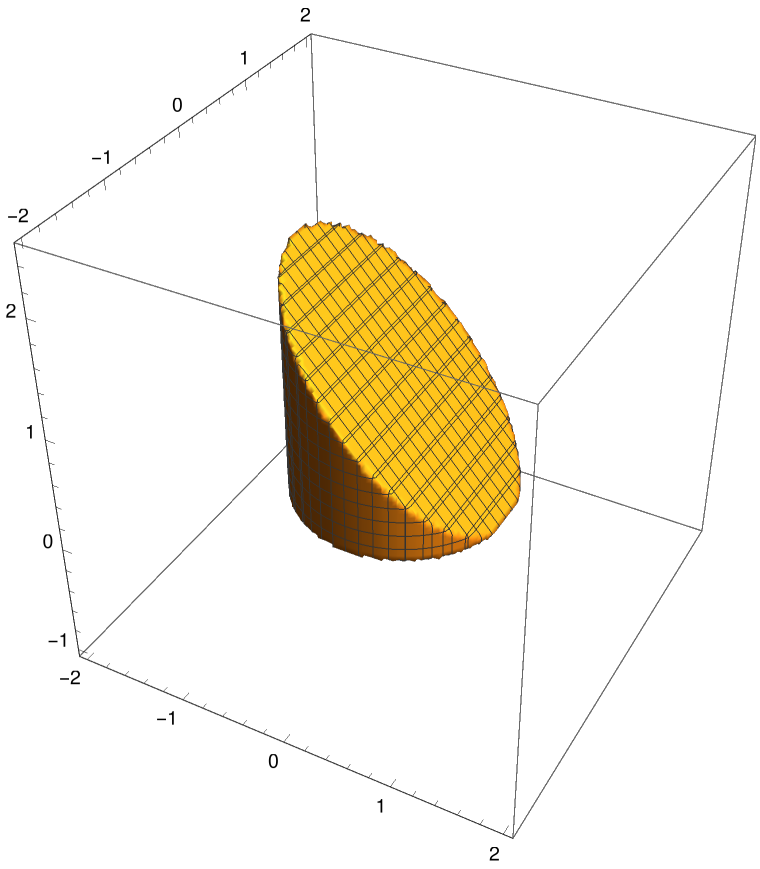

FIGURE 2.10 – Représentation du cylindre biseauté Ω.

On a vu que les intégrales doubles et triples se calculent de manière similaires aux intégrales simples, à l'aide du théorème de Fubini et de changements de variables. Il y a cependant un outil qui existe pour les fonctions d'une seule variable, et qui n'a pas d'analogue direct

en dimension supérieure : la formule

$$\int_a^b f'(t)\,dt = f(b) - f(a)$$

permet de calculer une intégrale en terme d'une primitive de son intégrand. Dans la suite du cours, on va rencontrer des analogues multidimensionnels (plus compliqués) de cette formule : la formule de Green-Riemann, la formule de Stokes, et la formule de Green-Ostrogradski.

Chapitre 3

Intégrales curvilignes

Le but de ce chapitre est de développer un concept d'intégration de fonctions (scalaires ou vectorielles) le long de courbes dans $\mathbb{R}^n$. En physique, ces notions peuvent être utiles pour calculer :

- la masse, le centre de masse, ou le moment d'inertie d'un fil inhomogène ;
- le travail d'une force agissant sur un corps ponctuel lors d'un déplacement ;
- l'intensité du courant traversant une surface en fonction du champ magnétique ambiant (loi d'Ampère) ;
- la force électromotrice d'un circuit électrique en fonction du champ électrique ambiant (loi de Faraday).

3.1 Courbes paramétrées

Avant de pouvoir intégrer une fonction le long d'une courbe, on définit ce qu'on entend par "courbe". Ce concept est abordé dans le cours de Math104 (voir le chapitre 3 du poly).

Définition 3.1.1: Courbe paramétrée de classe C^1

On appelle *courbe paramétrée de classe C^1 dans* $\mathbb{R}^n$ un couple $\gamma = (I, \varphi)$ où $I \subset \mathbb{R}$ est un intervalle de $\mathbb{R}$ et $\varphi : I \to \mathbb{R}^n$ est une application de classe C^1. On appelle *image* de la courbe paramétrée le sous-ensemble de $\mathbb{R}^n$ défini par

$$\Im(\gamma) := \Im(\varphi) = \{\varphi(t) \ : t \in I\}.$$

Les courbes paramétrées seront souvent notées γ dans la suite.

GD : J'insiste un peu que c'est une définition pour des courbes de classe C^1. C'est la régularité qu'on utilisera ci-dessous, parce qu'on veut calculer la longueur et intégrer des fonctions sur ces courbes. Autrement, le plus logique aurait été de définir les courbes paramétrées (tout court, ou continues) en demandant seulement à $\varphi : I \to \mathbb{R}^n$ d'être une fonction continue.

Rappelons aussi que "$\varphi : I \to \mathbb{R}^n$ est une de classe C^1" signifie que chacune de ses coordonnées est de classe C^1.

Exemple 3.1.2 — Si $I = [0, 2\pi]$ et $\varphi(t) = (\cos(t), \sin(t))$ pour tout $t \in I$, alors $\varphi : I \to \mathbb{R}^2$ est de classe C^1 et donc $\gamma = (I, \varphi)$ est une courbe paramétrée. Son image est le cercle centré en l'origine et de rayon 1 dans $\mathbb{R}^2$.

— Si $I = [0, 2\pi]$ et $\tilde{\varphi}(t) = (\cos(2t), \sin(2t))$ pour tout $t \in I$, alors $\tilde{\varphi}$ est C^1 et donc $\tilde{\gamma} = (I, \tilde{\varphi})$ est une courbe paramétrée. Son image est aussi le cercle centré en l'origine et de rayon 1 dans $\mathbb{R}^2$. Les courbes paramétrées γ et $\tilde{\gamma}$ ont la même image, mais sont pourtant des courbes paramétrées différentes. On verra dans la suite pourquoi il est important de les distinguer.

— Si $I = [0, 1]$ et $\varphi(t) = (t, 2t) = t(1, 2)$ pour tout $t \in I$, alors $\varphi : I \to \mathbb{R}^2$ est C^1 et donc (I, φ) est une courbe paramétrée de $\mathbb{R}^2$. Son image est le segment reliant les points $(0, 0)$ et $(1, 2)$ dans $\mathbb{R}^2$.

— Si $I = [0, 1]$ et $\varphi(t) = (1+t, 1-t, 3t)$ pour tout $t \in I$, alors $\varphi : I \to \mathbb{R}^3$ est C^1 et donc (I, φ) est une courbe paramétrée de $\mathbb{R}^3$. Son image est le segment reliant les points $(1, 1, 0)$ et $(2, 0, 3)$ dans $\mathbb{R}^3$.

— On parlera un peu plus loin du graphe d'une fonction C^1 $f : I \to \mathbb{R}$, qui est un bon exemple de courbe tracée dans $\mathbb{R}^2$.

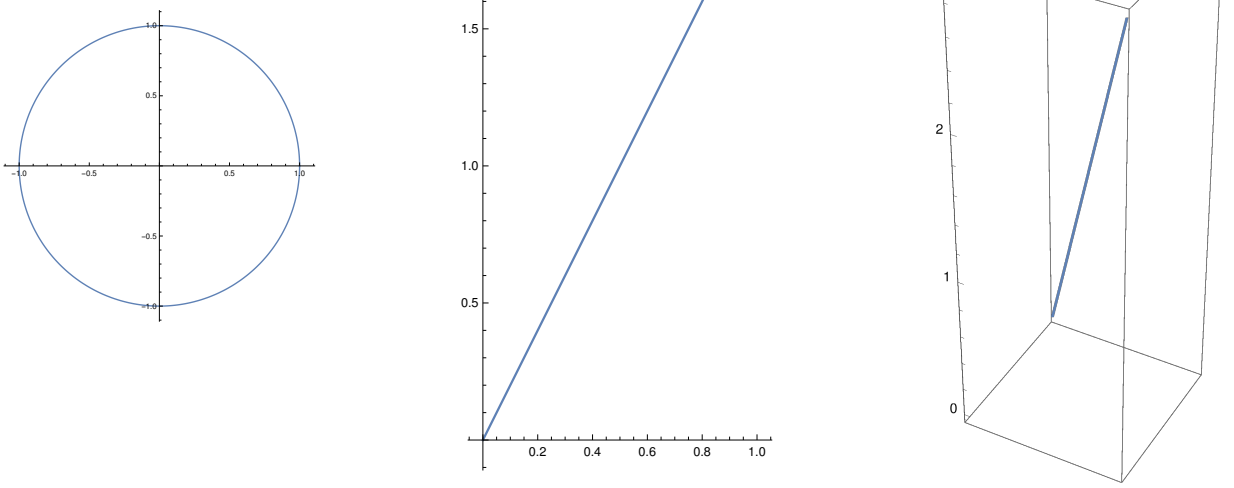

FIGURE 3.1 – Images des courbes paramétrées de l'exemple

Remarque 3.1.1 L'idée des courbes paramétrées est donc de formaliser le fait qu'une courbe est un objet de "dimension 1" dans $\mathbb{R}^n$. Ici, cela signifie que cela correspond à une "déformation" de l'intervalle réel I par l'application φ, qui part de l'intervalle I qui est "plat" dans $\mathbb{R}$ pour le déformer dans $\mathbb{R}^n$.

Remarque 3.1.2 On peut imaginer des courbes paramétrées plus compliquées, comme $([0, 2\pi], t \mapsto ((1+\cos(t^2/(2\pi))/2)\cos(t), (1+\cos(t^2/(2\pi))/2)\sin(t))$ représentée en Figure 3.2, ou $([0, 4\pi], t \mapsto (\cos(t), \sin(t), t))$ qui représente une spirale en $3d$, représentée en Figure 3.3.

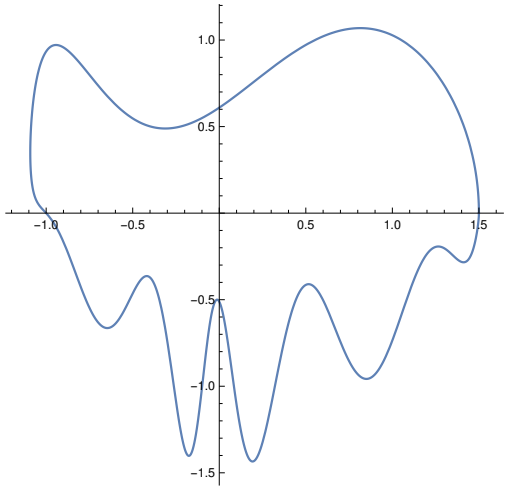

FIGURE 3.2 – Une courbe compliquée en $2d$

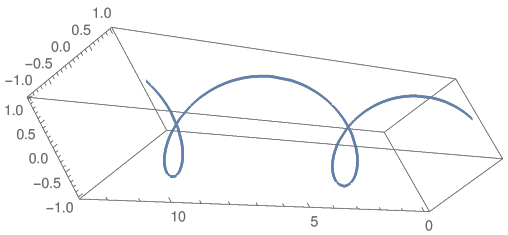

FIGURE 3.3 – Une courbe compliquée en $3d$

Remarque 3.1.3 Dans notre présentation, on part d'une courbe paramétrée (qui est un couple (I, φ)) et on lui associe une "courbe géométrique" : son image (qui est un sous-ensemble de $\mathbb{R}^n$). Il est peut-être plus naturel de renverser le point de vue, en partant d'une courbe géométrique (par exemple un cercle dans $\mathbb{R}^2$). Comme on l'a vu dans les exemples, pour une courbe géométrique donnée, il peut exister plusieurs courbes paramétrées (I, φ) dont elle est l'image. On dit qu'un tel couple (I, φ) est un *paramétrage* de la courbe géométrique. Mathématiquement, il est plus simple définir une courbe paramétrée (qui est juste une fonction) plutôt qu'une courbe géométrique (qui est un sous-ensemble de $\mathbb{R}^n$ satisfaisant un certain nombre de propriétés compliquées).

On a vu dans les exemples que pour une courbe géométrique donnée (par exemple, un cercle), il existe plusieurs paramétrages possibles. Pour distinguer ces paramétrages, on utilise la notion suivante.

Définition 3.1.3: Courbes paramétrées (de classe C^1) équivalentes

Deux courbes paramétrées (I, φ) et (J, ψ) sont dites *équivalentes* s'il existe une bijection $\theta : I \to J$ de classe C^1, strictement croissante et dont la dérivée ne s'annule pas sur I, telle que

$$\varphi = \psi \circ \theta \,.$$

GD : j'ai ajouté la condition supplémentaire que $\theta'(x) > 0$ (puisque θ est croissante) pour tout $x \in I$. C'est parce que je veux que si (I, φ) est équivalente à (J, ψ), alors (J, ψ) est

équivalente à (I, φ). Et pour vérifier ceci, j'ai besoin que la fonction réciproque $\theta^{-1} : J \to I$ soit de classe C^1. Pour ceci, on a un théorème d'inversion bien pratique que vous avez vu en L1, mais qui demande bien que $\theta'(x) > 0$.

Je rappelle que la fonction $x \mapsto x^3$, de $[-1, 1]$ dans $[-1, 1]$, est de classe C^1, bijective, croissante, mais sa dérivée s'annule en 0 et son inverse $y \mapsto y^{1/3}$ n'est pas dérivable en 0.

Remarque 3.1.4 De manière évidente, deux courbes paramétrées équivalentes possèdent la même image.

Exemple 3.1.4 Les courbes $([0, 2\pi], t \mapsto (\cos(t), \sin(t)))$ et $([0, \pi], t \mapsto (\cos(2t), \sin(2t))$ sont équivalentes, en choisissant $\theta : [0, \pi] \to [0, 2\pi]$ telle que $\theta(t) = 2t$ pour tout $t \in [0, \pi]$, qui est bien une bijection de classe C^1 strictement croissante.

On va voir que des courbes paramétrées équivalentes possèdent beaucoup de propriétés communes.

Remarque 3.1.5 Dans notre définition de courbe paramétrée, on a supposé que la fonction φ est de classe C^1. On peut relâcher cette hypothèse, en demandant juste qu'elle soit continue et C^1 *par morceaux* : cela signifie qu'elle est continue sur I, et qu'il existe une partition de I en un nombre fini de sous-intervalles I_j telle que φ est C^1 sur chacun des I_j. Mais φ n'est pas forcément C^1 sur tout I, elle pourrait par exemple ne pas être dérivable au bord des I_j. Un exemple de fonction C^1 par morceaux est la fonction $x \mapsto |x|$ sur $[-1, 1]$: elle n'est pas dérivable en 0 donc elle n'est pas C^1 sur $[-1, 1]$, par contre elle est bien C^1 sur $[-1, 0[$ et sur $[0, 1]$: elle est donc C^1 par morceaux sur $[-1, 1]$. Il est naturel d'étendre la notion de courbes paramétrées au cadre C^1 par morceaux, qui est adapté aux courbes qui possèdent des *coins* : par exemple, le bord d'un carré est une courbe paramétrée C^1 par morceaux.

GD : j'ai quand même demandé que φ soit continue. C'est pas vraiment obligé a priori, mais on ne sera pas intéressé par des courbes qui sautent, et il faudrait faire attention à certains détails pour appppliquer les théorèmes de la suite. Pour l'équivalence aussi, on peut alors se contenter de demander que θ soit continue, strictement croissante, et seulement de classe C^1 par morceaux, ainsi que son inverse.

On décrit à présent quelques courbes paramétrées basiques.

1. Si A et B sont deux points dans $\mathbb{R}^n$, on peut paramétrer le segment $[AB]$ en posant $I = [0, 1]$ et $\varphi : I \to \mathbb{R}^n$ telle que

$$\varphi(t) = A + t\overrightarrow{AB}$$

pour tout $t \in [0, 1]$. [NB : voici l'intrusion d'un vecteur. Si cela vous dérange un peu, pensez juste en termes de coordonnées, et que $\overrightarrow{AB} = B - A$.] La fonction φ ainsi définie est clairement de classe C^1 (elle est affine), donc (I, φ) est une courbe paramétrée d'image le segment $[AB]$. Par exemple, si $A = (1, 1)$ et $B = (2, 3)$ dans $\mathbb{R}^2$, alors $\overrightarrow{AB} = (1, 2)$ et

$$\varphi(t) = (1, 1) + t(1, 2) = (1 + t, 1 + 2t)$$

pour tout $t \in [0,1]$.

2. Si l'on veut paramétrer le graphe $\{y = f(x)\}$ d'une fonction $f : I \to \mathbb{R}$ de classe C^1, on peut poser
$$\varphi(t) = (t, f(t)), \quad \forall t \in I\,,$$
et alors (I, φ) est une courbe paramétrée de classe C^1, d'image le graphe de f.

3. Pour paramétrer un cercle de centre $(x_0, y_0) \in \mathbb{R}^2$ et de rayon $R > 0$ dans $\mathbb{R}^2$, on peut introduire $I = [0, 2\pi]$ et définir
$$\varphi(t) = (x_0 + R\cos(t), y_0 + R\sin(t))$$
pour tout $t \in I$. Alors, φ est clairement C^1 et (I, φ) est donc une courbe paramétrée C^1 dont l'image est le cercle de centre (x_0, y_0) et de rayon $R > 0$.

On peut également paramétrer des courbes à l'aide des coordonnées polaires.

Définition 3.1.5: Courbe paramétrée en coordonnées polaires

Soit $I \subset \mathbb{R}$ un intervalle et $R : I \to \mathbb{R}$ une fonction C^1. La courbe d'équation notée formellement $\{r = R(\theta),\ \theta \in I\}$ (ou plus simplement par l'équation $r = R(\theta)$) est la courbe paramétrée (I, φ) de $\mathbb{R}^2$ où
$$\varphi(\theta) = (R(\theta)\cos\theta, R(\theta)\sin\theta), \quad \forall \theta \in I.$$

Exemple 3.1.6 — La courbe d'équation $r = 1$ pour $0 \leq \theta \leq 2\pi$ est le cercle centré en l'origine de rayon 1.

— La courbe d'équation $r = \cos\theta$ pour $-\pi/2 \leq \theta \leq \pi/2$ est également un cercle, ce que l'on justifie brièvement. En effet, $r = \cos\theta$ implique que $r^2 = r\cos\theta$ et comme $x = r\cos\theta$ et $y = r\sin\theta$, on en déduit que
$$x^2 + y^2 = r^2 = r\cos\theta = x,$$
ce que l'on peut réécrire
$$\left(x - \frac{1}{2}\right)^2 + y^2 = \frac{1}{4},$$
autrement dit il s'agit du cercle de centre $(1/2, 0)$ de rayon $1/2$.

GD : on a décidé ici d'écrire R en fonction de θ, mais on aurait pu faire dans l'aitre sens, se donner une fonction Θ définie sur un intervalle I de $\mathbb{R}$ (ou pour simplifier de $\mathbb{R}_+$) et d'utiliser le paramétrage $\varphi : I \to \mathbb{R}^2$ donné par $\varphi(r) = (r\cos(\Theta(r)), r\sin(\Theta(r))$.

Autres exemples classiques de courbes : les spirales d'Archimède données par $r = a\theta + b$ (avec des paramètres $a, b \in \mathbb{R}$, $a \neq 0$ si on veut autre chose qu'un cercle), et les spirales logarithmiques, données par $r = ae^{b\theta}$.

3.2 Longueur d'une courbe, intégration d'une fonction par rapport à la longueur

On commence par définir la longueur d'une courbe paramétrée, puis on va intégrer des fonctions par rapport à la longueur d'une courbe.

Comme on pense souvent à des trajectoires de particules, on appellera la dérivée de $\gamma : I \to \mathbb{R}^n$ le vecteur vitesse. Ainsi, si $n = 3$ et $\gamma(t) = (\gamma_1(t), \gamma_2(t), \gamma_3(t))$, le vecteur vitesse au point t est $v(t) = (\gamma_1'(t), \gamma_2'(t), \gamma_3'(t))$. Et ce que j'appellerais plutôt "vitesse instantanée" serait le nombre $||v'(t)|| = \sqrt{\gamma_1'(t)^2 + \gamma_2'(t)^2 + \gamma_3'(t)^2}$.

3.2.1 Longueur d'une courbe

Définition 3.2.1: Longueur d'une courbe paramétrée

Soit $\gamma = (I, \varphi)$ une courbe paramétrée (de classe C^1). On appelle *longueur de* γ, notée $\ell(\gamma)$, la quantité

$$\ell(\gamma) = \int_I \|\varphi'(t)\| \, dt.$$

Cette définition de la longueur peut se justifier de la manière suivante : si l'on découpe l'intervalle I en petits intervalles $[t_j, t_j + \Delta t]$ avec $\Delta t > 0$ petit, l'ensemble

$$\{\varphi(t) \ : \ t \in [t_j, t_j + \Delta t]\}$$

peut raisonnablement être approché par le segment $[\varphi(t_j)\varphi(t_j + \Delta t)]$ qui est de longueur

$$\|\varphi(t_j + \Delta t) - \varphi(t_j)\| \sim \Delta t \|\varphi'(t_j)\|.$$

En sommant les longueurs de ces petits segments, on obtient

$$\sum_j \Delta t \|\varphi'(t_j)\| \, ,$$

qui converge bien dans la limite $\Delta t \to 0$ vers $\int_I \|\varphi'(t)\| \, dt$ (penser aux sommes de Riemann).

Exemple 3.2.2 Calculons les longueurs des courbes usuelles que l'on vient d'introduire.

— Dans le cas d'un segment $[AB]$, on a

$$\varphi(t) = A + t\overrightarrow{AB}, \quad t \in [0, 1],$$

donc $\varphi'(t) = \overrightarrow{AB}$ pour tout t. Ainsi,

$$\ell(\gamma) = \int_0^1 \|\varphi'(t)\| \, dt = \int_0^1 \|\overrightarrow{AB}\| \, dt = \|\overrightarrow{AB}\|,$$

qui est bien la longueur du segment $[AB]$.

— Dans le cas d'un graphe $\{(x,y)\,;\,y=f(x)\}$ avec $f:I\to\mathbb{R}$ de classe C^1, on a

$$\varphi(t)=(t,f(t)),\quad t\in[0,1],$$

donc $\varphi'(t)=(1,f'(t))$. Ainsi,

$$\ell(\gamma)=\int_0^1\|(1,f'(t))\|\,dt=\int_0^1\sqrt{1+(f'(t))^2}\,dt.$$

— Pour un cercle de centre $(x_0,y_0)\in\mathbb{R}^2$ et de rayon $R>0$, on avait choisi

$$\varphi(t)=(x_0+R\cos(t),y_0+R\sin(t)),\quad t\in[0,2\pi],$$

et donc $\varphi'(t)=(-R\sin(t),R\cos(t))$ pour tout $t\in[0,2\pi]$. Ainsi,

$$\ell(\gamma)=\int_0^{2\pi}\|(-R\sin(t),R\cos(t))\|\,dt=\int_0^{2\pi}\sqrt{R^2\sin^2(t)+R^2\cos^2(t)}\,dt=2\pi R.$$

Exemple 3.2.3 Si l'on considère la courbe de la Figure 3.3, on a $I=[0,4\pi]$ et pour tout $t\in I$,

$$\varphi(t)=(\cos(t),\sin(t),t)$$

donc

$$\varphi'(t)=(-\sin(t),\cos(t),1)$$

et ainsi

$$\|\varphi'(t)\|=\sqrt{(-\sin(t))^2+(\cos(t))^2+1}=\sqrt{2}.$$

Il est donc facile de calculer sa longueur,

$$\ell(I,\varphi)=\int_0^{4\pi}\|\varphi'(t)\|\,dt=4\pi\sqrt{2}.$$

Attention La longueur d'une courbe paramétrée ne coïncide pas toujours avec la longueur naturelle de son image : par exemple si $I=[0,4\pi]$ et $\varphi(t)=(\cos(t),\sin(t))$ pour tout $t\in I$, on a pour tout $t\in I$,

$$\|\varphi'(t)\|=\|(-\sin(t),\cos(t))\|=\sqrt{\sin^2(t)+\cos^2(t)}=1.$$

Ainsi,

$$\ell(I,\varphi)=\int_0^{4\pi}1\,dt=4\pi.$$

Or l'image de (I,φ) est le cercle centré en l'origine et de rayon 1 dans $\mathbb{R}^2$, qui n'est manifestement pas de longueur 4π ! Il faut donc distinguer la longueur d'une courbe paramétrée, qui dépend fortement du paramétrage choisi, et la longueur de l'image de cette courbe. Pour que les deux coïncident, il faut imposer d'autres conditions sur le paramétrage. Dans notre

cas, on voit le problème : comme on a choisi $I = [0, 4\pi]$ et non pas $I = [0, 2\pi]$, on a parcouru deux fois le cercle et donc on trouve deux fois la longueur attendue.

Et encore, si on fait un nombre entier connu de tours, on comprend aisément comment passer d'une longueur à l'autre, mais on peut aussi imaginer que $\gamma(t)$ reste sur le cercle tout le temps, mais y fait un parcours erratique et compliqué, d'une longueur totale parcourue n'a alors plus rien à voir avec 2π.

Pour être sûr que les deux notions de longueur coïncident, il faut imposer de plus que la fonction φ soit injective, donc bijective dans son image.

Deux courbes équivalentes ont la même longueur.

Proposition 3.2.1

Soient (I, φ) et (J, ψ) deux courbes paramétrées (de classe C^1) équivalentes. Alors

$$\ell(I, \varphi) = \ell(J, \psi).$$

Démonstration. On va voir qu'il s'agit d'une simple application du théorème de changement de variables dans les intégrales simples. Comme (I, φ) et (J, ψ) sont équivalentes, il existe $\theta : I \to J$ une bijection de classe C^1 strictement croissante telle que $\varphi = \psi \circ \theta$. Ainsi, pour tout $t \in I$, la règle de chaîne implique que

$$\forall t \in I, \qquad \varphi'(t) = \psi'(\theta(t))\, \theta'(t)$$

et donc en changeant de variable $s = \theta(t)$ on trouve

$$\ell(I, \varphi) = \int_I \|\varphi'(t)\|\, dt = \int_I \|\psi'(\theta(t))\| \theta'(t)\, dt = \int_J \|\psi'(s)\|\, ds = \ell(J, \psi).$$

□

Proposition 3.2.2: Longueur d'une courbe donnée en coordonnées polaires

La courbe d'équation $r = R(\theta)$ pour $\theta_1 \leq \theta \leq \theta_2$ a pour longueur

$$\int_{\theta_1}^{\theta_2} \sqrt{R(\theta)^2 + R'(\theta)^2}\, d\theta.$$

Démonstration. Par définition, la courbe paramétrée associée est donnée par $([\theta_1, \theta_2], \varphi)$, où

$$\varphi(\theta) = (R(\theta)\cos\theta, R(\theta)\sin\theta), \quad \theta_1 \leq \theta \leq \theta_2.$$

Pour tout $\theta \in [\theta_1, \theta_2]$, on a donc

$$\varphi'(\theta) = (R'(\theta)\cos\theta - R(\theta)\sin\theta, R'(\theta)\sin\theta + R(\theta)\cos\theta),$$

d'où

$$\|\varphi'(\theta)\| = \sqrt{(R'(\theta)\cos\theta - R(\theta)\sin\theta)^2 + (R'(\theta)\sin\theta + R(\theta)\cos\theta)^2} = \sqrt{R(\theta)^2 + R'(\theta)^2}.$$

Cela donne le résultat, car

$$\ell([\theta_1, \theta_2], \varphi) = \int_{\theta_1}^{\theta_2} \|\varphi'(\theta)\|\, d\theta.$$

□

3.2.2 Intégration d'une fonction par rapport à la longueur d'une courbe

Définition 3.2.4: Intégration par rapport à la longueur d'une courbe

Soit $\gamma = (I, \varphi)$ une courbe paramétrée de $\mathbb{R}^n$ et $f : \mathbb{R}^n \to \mathbb{R}$ une fonction continue. L'*intégrale de f par rapport à la longueur de γ*, notée $\int_\gamma f\, d\ell$, est le nombre

$$\int_\gamma f\, d\ell = \int_I f(\varphi(t))\|\varphi'(t)\|\, dt.$$

Cette définition peut s'interpréter physiquement de la façon suivante : si l'image de γ représente un fil matériel inhomogène de densité linéique de masse donnée par la fonction f (c'est-à-dire que la masse d'un élément infinitésimal de fil de longueur $d\ell$ au point M du fil est $f(M)\, d\ell$), alors la masse totale du fil vaut $\int_\gamma f\, d\ell$.

Remarque 3.2.1

1. De la même manière qu'en Proposition 3.2.1, l'intégrale de f sur γ ne change pas lorsqu'on considère deux courbes paramétrées équivalentes.
2. On appelle *$d\ell$ élément de longueur sur γ*, et on note souvent formellement
$$d\ell = \|\varphi'(t)\|dt.$$
3. Pour une courbe donnée en coordonnées polaires par l'équation $r = R(\theta)$, l'élément de longueur vaut
$$d\ell = \sqrt{R(\theta)^2 + R'(\theta)^2}\, d\theta.$$
4. Lorsque $f \equiv 1$ est la fonction constante égale 1, on retrouve la longueur de la courbe :
$$\int_\gamma 1\, d\ell = \ell(\gamma).$$

Exemple 3.2.5 — Calculons

$$\int_\gamma (x + y)\, d\ell,$$

où γ est le graphe de $y = 2x - 3$ pour x allant de -1 à 1. On a vu (voir les courbes paramétrées basiques avant la Définition 3.1.5) qu'on peut paramétrer ce graphe par (I, φ), où $I = [-1, 1]$ et

$$\varphi(t) = (t, 2t - 3), \quad t \in I.$$

Ici, on doit donc calculer

$$\int_I f(\varphi(t))\|\varphi'(t)\|\, dt,$$

avec $f(x, y) = x + y$ pour tout $(x, y) \in \mathbb{R}^2$. Pour tout $t \in I$, on a

$$f(\varphi(t)) = t + 2t - 3 = 3t - 3, \quad \|\varphi'(t)\| = \|(1, 2)\| = \sqrt{5},$$

et donc

$$\int_\gamma (x + y)\, d\ell = \int_{-1}^{1} (3t - 3)\sqrt{5}\, dt = \sqrt{5}[3t^2/2 - 3t]_{t=-1}^{t=1} = -6\sqrt{5}.$$

— Calculons

$$\int_\gamma x^2\, d\ell,$$

où γ est le cercle de $\mathbb{R}^2$ centré en $(1, 0)$ et de rayon 2. On a vu qu'un tel cercle est paramétré par (I, φ) où $I = [0, 2\pi]$ et

$$\varphi(t) = (1 + 2\cos(t), 2\sin(t)), \quad t \in I.$$

On doit donc calculer

$$\int_I f(\varphi(t))\|\varphi'(t)\|\, dt,$$

avec $f(x, y) = x^2$ pour tout $(x, y) \in \mathbb{R}^2$. Pour tout $t \in I$, on a

$$f(\varphi(t)) = (1 + 2\cos(t))^2, \quad \|\varphi'(t)\| = \|(-2\sin(t), 2\cos(t))\| = 2,$$

donc

$$\int_\gamma x^2\, d\ell = \int_0^{2\pi} (1 + 2\cos(t))^2 2\, dt = 12\pi.$$

Remarque 3.2.2 Dans les deux exemples précédents, on a juste donné γ par la courbe géométrique qu'il représente (une portion de graphe ou un cercle ici), sans donner le paramétrage que l'on choisissait. Cela est rigoureusement insuffisant comme on l'a vu, car plusieurs paramétrages non-équivalents du cercle existent par exemple. Néanmoins, il s'agit ici de courbes "usuelles" qui possèdent un paramétrage naturel : c'est celui-ci que l'on choisit toujours.

Application : calcul de l'aire d'une surface de révolution

On présente une application de l'intégrale d'une fonction par rapport à la longueur : le calcul de l'aire d'une surface de révolution dans $\mathbb{R}^3$. Pour construire une telle surface, on considère une courbe paramétrée γ de $\mathbb{R}^2$ dont l'image est incluse dans le demi-plan supérieur $\{(x,y) \in \mathbb{R}^2 \ : \ y \geq 0\}$, comme celle dessinée en Figure 3.4. On peut voir cette courbe comme une courbe tracée dans $\mathbb{R}^3$, en plongeant ce demi-plan supérieur dans $\mathbb{R}^3$, autrement dit en identifiant

$$\{(x,y) \in \mathbb{R}^2 \ : \ y \geq 0\} \simeq \{(x,y,z) \in \mathbb{R}^3 \ : \ y \geq 0, \ z = 0\}.$$

On fait à présent tourner l'image de la courbe γ autour de l'axe (Ox) en lui faisant faire un tour complet : l'ensemble balayé par la courbe pendant ce procédé définit une *surface de révolution*, comme représentée en Figure 3.5. On peut déterminer l'aire de cette surface à l'aide de la courbe γ :

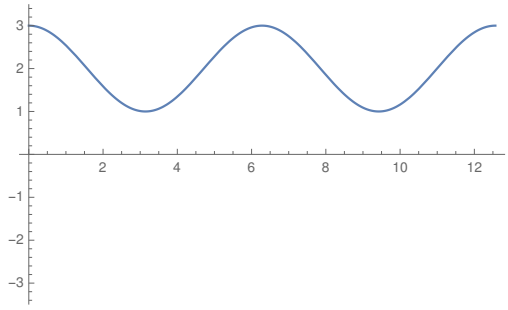

FIGURE 3.4 – Courbe à faire tourner autour de (Ox)

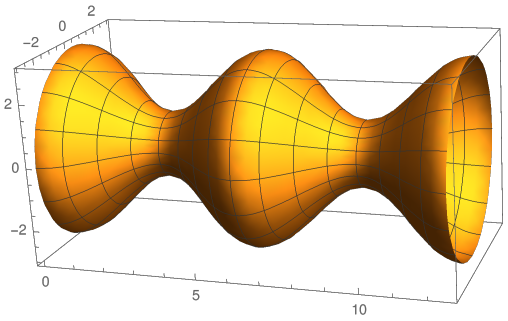

FIGURE 3.5 – Surface obtenue après rotation de la courbe

Proposition 3.2.3: Aire d'une surface de révolution

Soit γ une courbe paramétrée de $\mathbb{R}^2$ contenue dans le demi-plan supérieur, et S la surface de révolution obtenue en faisant tourner l'image de γ autour de l'axe (Ox) comme décrit ci-dessus. Alors l'aire de S est donnée par

$$\text{Aire}(S) = 2\pi \int_\gamma y \, d\ell.$$

Cette formule pourra se démontrer rigoureusement en définissant l'aire d'une surface, comme on le fera dans le chapitre suivant. Néanmoins, on peut en donner une explication heuristique : si M est un point de (l'image de) la courbe γ de coordonnées cartésiennes $(x, y = \varphi(x))$, on peut considérer un petit tronçon de γ de longueur $\Delta\ell$ autour de M. Ce tronçon peut être approché par un segment de longueur $\Delta\ell$ centré en M, donc de hauteur approximative y. Lorsque l'on fait tourner ce segment autour de l'axe (Ox), on obtient une surface de révolution qui est le bord d'un tronçon de cône de rayon approximatif y et de « hauteur »

$\Delta\ell$. Son aire vaut donc approximativement $2\pi y\Delta\ell$. En découpant la surface de révolution en un grand nombre de tels petits tronçons de cônes de rayons y_j, son aire totale est donc approchée par

$$\sum 2\pi y_j \Delta\ell\,,$$

qui converge bien vers $\int_\gamma y\,d\ell$ lorsque $\Delta\ell \to 0$.

Exemple 3.2.6 — Considérons la surface

$$S = \{(x,y,z) \in \mathbb{R}^3 \ : \ x^2+y^2+z^2 = R^2\},$$

qui est juste la sphère centrée en l'origine de rayon $R > 0$. On peut voir S comme une surface de révolution, partant de la courbe $\gamma = (I,\varphi)$ donnée par $I = [0,\pi]$ et

$$\varphi(t) = (R\cos(t), R\sin(t)), \quad t \in I.$$

On a déjà vu que pour tout $t \in I$, $\|\varphi'(t)\| = R$ et donc

$$\text{Aire}(S) = 2\pi \int_\gamma y\,d\ell = 2\pi \int_0^\pi R\sin(t)\,R\,dt = 4\pi R^2,$$

qui est bien la formule d'aire d'une sphère de rayon R.

— On peut calculer également l'aire d'un tore de révolution construit de la manière suivante : on considère deux paramétres $0 < r < R$ et $\gamma = (I,\varphi)$ d'image le cercle de centre $(0,R)$ et de rayon r. On prend donc $I = [0,2\pi]$ et

$$\varphi(t) = (r\cos(t), R + r\sin(t)), \quad t \in I.$$

Le tore de révolution S est construit comme la surface de révolution basée sur la courbe γ, comme représenté en Figure 3.6. Son aire vaut alors

$$\text{Aire}(S) = 2\pi \int_\ell y\,d\ell = 2\pi \int_I \varphi_2(t)\|\varphi'(t)\|\,dt.$$

Pour tout $t \in I$, on a

$$\varphi_2(t) = R + r\sin(t), \quad \varphi'(t) = (-r\sin(t), r\cos(t)),$$

et donc $\|\varphi'(t)\| = r$. Ainsi,

$$\text{Aire}(S) = 2\pi \int_0^{2\pi} (R + r\sin(t))r\,dt = 4\pi^2 Rr.$$

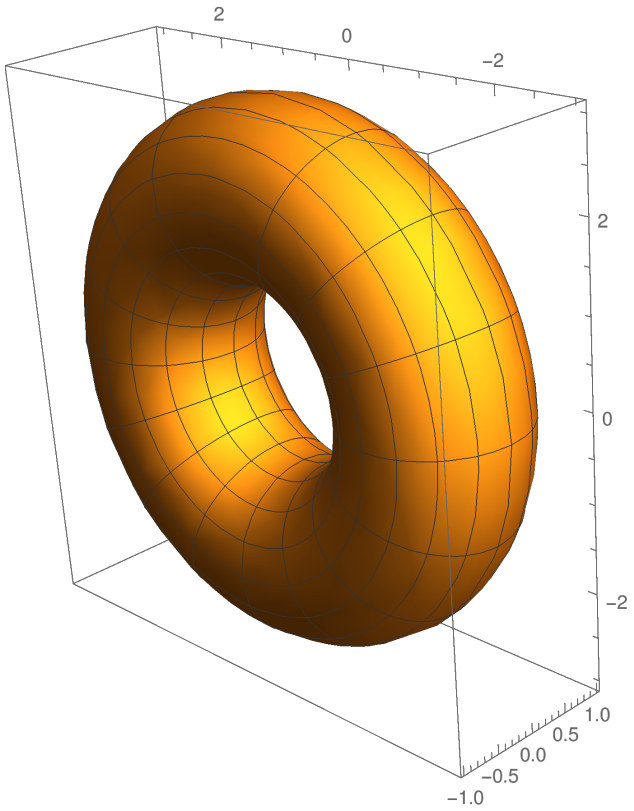

FIGURE 3.6 – Le tore de révolution avec $r = 1$ et $R = 2$.

3.3 Circulation d'un champ de vecteurs le long d'une courbe orientée

GD : J'ai fait un peu différemment en cours, mais je n'ose pas changer tout le poly ; je laisserai mon successeur décider. En attendant, j'explique ici les différences de présentation (rassurez-vous, qui ne changent pas trop le fond). J'ai souvent gardé la notation γ pour parler des courbes paramétrées, et comme dans la définition 3.1.3 de l'équivalence de deux paramétrages, on n'a autorisé que des changements de variables strictement croissants, les courbes paramétrées dont on parle depuis le début viennent toutes avec une orientation. En particulier si $\gamma = (I, \varphi)$ avec $I = [a, b]$, la courbe commence au point $\varphi(a)$ et se termine au point $\varphi(b)$. Si on suppose de plus que φ est injective, alors on peut vraiment parler d'un sens de parcours le long de l'image $\Im(\gamma)$, donc d'une orientation de γ.

Dans cette section on fera implicitement l'hypothèse que l'application $\varphi : I \to \mathbb{R}^n$ est injective.

Maintenant, pour toute courbe $\gamma = (I, \varphi)$, avec $I = [a, b]$, on peut lui associer une courbe parcourue dans le sens opposé (on dira, d'orientation opposée) et que j'ai appelée γ^*. Un paramétrage possible de γ^* est $\varphi^* : I \to \mathbb{R}^n$ définie par $\varphi^*(t) = \varphi(a + b - t)$ pour $t \in [a, b]$, où l'on s'est débrouillé pour utiliser le même intervalle I que pour φ, mais on aurait pu aussi

prendre $\varphi^* : [-b, -a] \to \mathbb{R}^n$ définie par $\varphi^*(t) = \varphi(-t)$. Cette nouvelle courbe γ^* n'est pas (en général) équivalente à γ : elle part de $\varphi(b)$ et se termine en $\varphi(a)$. Donc par exemple si γ est le paramétrage de l'intervalle $[AB]$ défini plus haut, φ^* est un paramétrage de $[BA]$.

Une fois cette définition faite, on peut vérifier que les deux manières de définir γ^* ci-dessus donnent deux paramétrages équivalents, et que si $\gamma = (I, \varphi)$ et $\widetilde{\gamma} = (J, \psi)$ sont deux paramétrages équivalents, les courbes γ^* et $\widetilde{\gamma}^*$ obtenues par changement d'orientation sont équivalentes entre elles.

Pourquoi parler d'orientation maintenant ? Pour la définition de l'intégrale $\int_\gamma f \, d\ell$ d'une fonction sur γ, on peut vérifier que changer l'orientation de γ ne change pas l'intégrale. Autrement dit, $\int_\gamma^* f \, d\ell = \int_\gamma f \, d\ell$, et donc on n'a pas besoin de faire trop attention. La vérification est la même que pour la proposition 3.3.1. Mais ici on va voir que pour la circulation $Circ(F, \gamma)$ d'un champ de vecteur $F : \mathbb{R}^n \to \mathbb{R}^n$, changer l'orientation de γ change le signe de la circulation. Autrement dit, $Circ(F, \gamma^*) = -Circ(F, \gamma)$. Donc quand on parle de la circulation de F sur un cercle, par exemple, il faudra bien dire comment le cercle est orienté, c.-à-d., dans quel sens il est parcouru. Bien entendu, il n'est pas question ici de changer la définition de la circulation, à savoir

$$Circ(F, \gamma) = \int_I F(\varphi(t)) \cdot \varphi'(t) \, dt$$

où l'on a intégré le produit scalaire entre les deux vecteurs $F(\varphi(t))$ et $\varphi'(t)$ (le vecteur-vitesse). La notation abrégée $Circ(F, \gamma) = \int_\gamma \vec{F} \cdot \vec{dx}$ est bien utile (et y mettre des flèches set un bon moyen de se souvenir qu'on a des vecteurs) ; donc ici on pense que $\vec{dx} = \overrightarrow{\varphi'(t)dt}$ est un petit déplacement *vectoriel* le long de la courbe. On remarque que l'élément infinitésimal de longueur $d\ell = \|\vec{dx}\|$.

J'ai aussi dit les chose dans un ordre un peu différent : d'abord la définition de la circulation, puis le fait qu'elle ne dépend pas du paramétrage de γ (tant qu'il reste équivalent), puis définition du changement d'orientation et vérification qu'alors on change le signe de la circulation. Revenons au cours (ancien) du poly.

Orientation d'une courbe paramétrée

Nous remarquons que dans le cas où φ est injective [1], la définition d'une courbe paramétrée $\gamma = (I, \varphi)$ implique que l'image $\Im(\gamma)$ est parcourue dans un certain sens lorsque t parcourt $I = [a, b]$: on part du point $\varphi(a)$ pour aller vers le point $\varphi(b)$. Comme φ est injective, on ne revient jamais sur ses pas, mais on reste toujours dans le même sens. Ce sens de parcours fournit une *orientation* à $\Im(\gamma)$, elle indique que la courbe γ est intrinsèquement orientée.

Définissons la fonction

$$\varphi^*(t) = \varphi(a + b - t), \quad t \in [a, b].$$

1. Ci-dessous on supposera que φ est injective, sauf possiblement sur ses points extrémaux : on autorisera parfois l'égalité $\varphi(a) = \varphi(b)$.

La courbe (I, φ^*) est encore une courbe paramétrée, qui possède la même image que (I, φ). Cependant, ces deux courbes ne sont pas équivalentes en général. On a bien $\varphi^* = \varphi \circ \theta$, où $\theta : [a, b] \to [a, b]$ est donnée par $\theta(t) = a + b - t$, mais ici, θ est une bijection C^1 strictement *décroissante*. L'application φ^* part de $\varphi(b)$ et parcourt l'image de la courbe en sens inverse de φ, pour arriver au point final $\varphi(a)$.

On notera cette courbe en sens inverse

$$\gamma^* = (I, \varphi^*)$$

Pour indiquer cette orientation de γ (ou l'orientation inverse de γ^*), on représente graphiquement le sens de parcours le long de $\Im(\gamma)$ à l'aide d'une flèche, comme en Figure 3.7.

FIGURE 3.7 – Si γ est la courbe de gauche, alors γ^* est la courbe de droite.

Remarque 3.3.1 Il est facile de montrer que la longueur ne dépend pas du sens de parcours :

$$\ell(\gamma) = \ell(\gamma^*)\,.$$

Plus généralement, pour tout fonction scalaire f,

$$\int_\gamma f\, d\ell = \int_{\gamma^*} f\, d\ell.$$

Par contre, le type d'intégrales que nous présentons ci-dessous va, lui, dépendre du sens de parcours choisi.

Exemple 3.3.1 Si γ est le segment $[AB]$ parcouru de A vers B, c'est-à-dire paramétré par $([0, 1], \varphi)$ avec

$$\varphi(t) = A + t\overrightarrow{AB}, \quad t \in [0, 1],$$

alors γ^* est par définition paramétré par $([0, 1], \varphi^*)$ où

$$\varphi^*(t) = \varphi(1 - t) = A + (1 - t)\overrightarrow{AB} = A + \overrightarrow{AB} - t\overrightarrow{AB} = B - t\overrightarrow{AB} = B + t\overrightarrow{BA},$$

ce qui correspond bien au segment $[BA]$ parcouru de B vers A.

Circulation d'un champ de vecteur le long d'une courbe paramétrée

Maintenant qu'on a défini la notion d'orientation d'une courbe paramétrée, on peut définir la circulation d'un champ de vecteurs le long d'une telle courbe.

Définition 3.3.2: Circulation d'un champ de vecteurs le long d'une courbe orientée

Soit $\gamma = (I, \varphi)$ une courbe paramétrée de $\mathbb{R}^n$, et $F : \mathbb{R}^n \to \mathbb{R}^n$ un champ de vecteurs C^1. On appelle *circulation de F le long de γ*, notée $\mathrm{Circ}(F, \gamma)$, l'intégrale

$$\mathrm{Circ}(F, \gamma) = \int_I F(\varphi(t)) \cdot \varphi'(t)\, dt.$$

Notation On rencontre parfois les autres notations

$$\mathrm{Circ}(F, \gamma) = \int_\gamma \vec{F} \cdot \overrightarrow{dx} = \int_\gamma F_1\, dx_1 + F_2\, dx_2 + \cdots + F_n\, dx_n.$$

La circulation dépend de l'orientation de la courbe :

Proposition 3.3.1

$$\mathrm{Circ}(F, \gamma^*) = -\,\mathrm{Circ}(F, \gamma).$$

Démonstration. Si $\gamma = (I, \varphi)$, alors $\gamma^* = (I, \varphi^*)$ avec

$$\varphi^*(t) = \varphi(a + b - t), \quad t \in [a, b] = I.$$

On déduit que pour tout $t \in I$,

$$\varphi^{*\prime}(t) = -\varphi'(a + b - t)$$

et ainsi

$$\mathrm{Circ}(F, \gamma^*) = \int_a^b F(\varphi^*(t)) \cdot \varphi^{*\prime}(t)\, dt = -\int_a^b F(\varphi(a + b - t)) \cdot \varphi'(a + b - t)\, dt.$$

En opérant le changement de variable $s = a + b - t$ dans cette intégrale, on trouve

$$\begin{aligned} -\int_a^b F(\varphi(a + b - t)) \cdot \varphi'(a + b - t)\, dt &= \int_b^a F(\varphi(s)) \cdot \varphi'(s)\, ds \\ &= -\int_a^b F(\varphi(s)) \cdot \varphi'(s)\, ds \\ &= -\,\mathrm{Circ}(F, \gamma), \end{aligned}$$

ce qui démontre le résultat. □

On voit donc que la circulation change de signe lorsque l'on change l'orientation.

Remarque 3.3.2 La circulation d'un champ de vecteurs possède l'application suivante en physique. Si $t \in [a, b] \mapsto \varphi(t)$ représente la trajectoire d'une particule ponctuelle dans $\mathbb{R}^n$ (le paramètre t jouant alors le rôle du temps) et F est une force agissant sur cette particule lors de son déplacement, la circulation de F le long de $\overset{\frown}{\gamma}$ correspond au *travail* de la force F le long de γ, c'est-à-dire à l'énergie fournie par la force F pendant le déplacement.

Notation Si la courbe γ est *fermée*, c'est-à-dire si son point de départ est le même que son point d'arrivée ($\varphi(a) = \varphi(b)$ où $\gamma = ([a,b], \varphi)$), on note

$$\int_\gamma \vec{F} \cdot \overrightarrow{dx} = \oint_\gamma \vec{F} \cdot \overrightarrow{dx}.$$

Même dans cette situation où l'image $\Im(\gamma)$ est une boucle, la circulation de F dépend bien de l'orientation de la courbe.

Proposition 3.3.2

Si $\gamma_1 = (I_1, \varphi_1)$ et $\gamma_2 = (I_2, \varphi_2)$ sont deux courbes paramétrées équivalentes, on a pour tout champ de vecteurs F :

$$\mathrm{Circ}(F, \gamma_1) = \mathrm{Circ}(F, \gamma_2)\,.$$

Démonstration. Par définition, il existe $\theta : I_1 \to I_2$ une bijection croissante C^1 telle que $\varphi_1 = \varphi_2 \circ \theta$. Pour tout $t \in I_1$, on a donc

$$\varphi_1'(t) = \varphi_2'(\theta(t))\theta'(t),$$

et ainsi

$$\mathrm{Circ}(F, \gamma_1) = \int_{I_1} F(\varphi_1(t)) \cdot \varphi_1'(t)\, dt = \int_{I_1} F(\varphi_2(\theta(t))) \cdot \varphi_2'(\theta(t))\theta'(t)\, dt.$$

En changeant de variables $s = \theta(t)$, on obtient donc

$$\int_{I_1} F(\varphi_2(\theta(t))) \cdot \varphi_2'(\theta(t))\theta'(t)\, dt = \int_{I_2} F(\varphi_2(s)) \cdot \varphi_2'(s)\, ds = \mathrm{Circ}(F, \gamma_2),$$

ce qui démontre le résultat (ici on a utilisé la positivité de $\theta'(s)$ pour dire que $\theta'(t)\, dt = ds$). □

Exemple 3.3.3 Calculons

$$\int_\gamma \vec{F} \cdot \overrightarrow{dx}$$

avec $F : \mathbb{R}^2 \to \mathbb{R}^2$ définie par $F(x,y) = (x^2, xy)$ pour tout $(x,y) \in \mathbb{R}^2$, et où γ est le segment $[AB]$ parcouru de $A(0,1)$ vers $B(1,0)$. Ici encore, on se donne le chemin γ via son image et une orientation ; il faut donc définir un paramétrage ayant la bonne orientation. On utilisera le paramétrage standard des segments : $\overset{\frown}{\gamma} = ([0,1], \varphi)$ avec

$$\varphi(t) = A + t\overrightarrow{AB} = (0,1) + t(1,-1) = (t, 1-t), \quad t \in [0,1].$$

Comme $\varphi'(t) = (1,-1)$ et que $F(\varphi(t)) = F(t, 1-t) = (t^2, t(1-t))$ pour tout $t \in [0,1]$, on en déduit que

$$\int_\gamma \vec{F} \cdot \overrightarrow{dx} = \int_0^1 F(\varphi(t)) \cdot \varphi'(t)\, dt = \int_0^1 (t^2, t(1-t)) \cdot (1,-1)\, dt = \int_0^1 (t^2 - t(1-t))\, dt,$$

donc après un calcul élémentaire de l'intégrale à un paramètre :

$$\int_\gamma \vec{F} \cdot \overrightarrow{dx} = \int_0^1 (2t^2 - t)\, dt = \frac{1}{6}.$$

Circulation d'un champ de gradients

Le résultat suivant est un analogue de la formule

$$\int_a^b f'(t)\,dt = f(b) - f(a)$$

pour des fonctions de plusieurs variables.

Proposition 3.3.3

Si f est un champ scalaire C^1 sur $\mathbb{R}^n$ et si γ est une courbe paramétrée de $\mathbb{R}^n$ de point de départ A et de point d'arrivée B (c'est-à-dire que $A = \varphi(a)$ et $B = \varphi(b)$ où $\gamma = ([a,b], \varphi)$), alors

$$\int_\gamma \overrightarrow{\nabla f} \cdot \overrightarrow{dx} = f(B) - f(A).$$

Démonstration. Par définition, on a

$$\int_\gamma \overrightarrow{\nabla f} \cdot \overrightarrow{dx} = \int_a^b \nabla f(\varphi(t)) \cdot \varphi'(t)\,dt.$$

Or, on peut réécrire pour tout $t \in [a,b]$,

$$\nabla f(\varphi(t)) \cdot \varphi'(t) = \frac{\partial f}{\partial x_1}(\varphi(t))\varphi_1'(t) + \frac{\partial f}{\partial x_2}(\varphi(t))\varphi_2'(t) + \cdots + \frac{\partial f}{\partial x_n}(\varphi(t))\varphi_n'(t),$$

et l'on reconnait ici la dérivée d'une fonction composée (autrement dit, la règle de chaîne) :

$$\nabla f(\varphi(t)) \cdot \varphi'(t) = g'(t)$$

avec $g = f \circ \varphi$. Ainsi,

$$\int_a^b \nabla f(\varphi(t)) \cdot \varphi'(t)\,dt = \int_a^b g'(t)\,dt = g(b) - g(a) = f(\varphi(b)) - f(\varphi(a)) = f(B) - f(A).$$

□

GD : Je me permets d'insister sur le fait que c'est une formule remarquable. Quel que soit le chemin que vous prenez pour aller de $A = \varphi(a)$ à $B = \varphi(b)$, si $F = \nabla f$, la circulation de F le long du chemin sera toujours $f(B) - f(A)$.

Le plus souvent vous verrez ceci sous la forme "si la force F dérive d'un potentiel V, son travail le long du chemin de dépend pas du chemin choisi, juste des points de départ et d'arrivée". C'est d'autant mieux que beaucoup de forces dérivent d'un potentiel.

Corollaire 3.3.4

Si f est un champ scalaire C^1 sur $\mathbb{R}^n$ et si γ est une courbe paramétrée de $\mathbb{R}^n$ qui est de plus **fermée**, on a

$$\oint_\gamma \overrightarrow{\nabla f} \cdot \overrightarrow{dx} = 0.$$

Démonstration. Si $\overset{\frown}{\gamma}$ est une courbe fermée, cela signifie que $A = B$. En utilisant la proposition précédente, on en déduit que

$$\oint_\gamma \overrightarrow{\nabla f} \cdot \overrightarrow{dx} = f(B) - f(A) = f(A) - f(A) = 0.$$

□

Ce résultat donne une condition nécessaire pour être un champ de gradient : si un champ de vecteurs F est un champ de gradient, alors sa circulation le long de toute courbe fermée est nulle.

GD : En fait, pour deviner ce que devrait être la fonction f connaissant F, il s'agit simplement de partir d'un point fixe x_0 du domaine (et de la valeur $f(x_0)$ qui est une constante), puis la proposition dit que $f(x) - f(x_0)$ sera forcément la circulation de F sur n'importe quel chemin qui va de x_0 à x. On peut en choisir un au hasard, mais encore faut-il montrer que cette circulation ne dépend pas du chemin choisi. Et c'est pareil, en parcourant le premier chemin dans le sens direct, puis en repartant à l'envers le long du second chemin, que de vérifier que la circulation le long de la boucle obtenue, qui est la différence des deux circulations (puisqu'on a changé l'orientation du second chemin) est nulle. La condition sur le rotationnel qu'on a donnée au chapitre 1 permet de s'assurer que si on modifie un peu le chemin, la circulation ne change pas. Mais l'exemple qui suit montre que cela ne suffit pas pour assurer que la circulation sur n'importe quelle boucle est nulle.

Exemple 3.3.4 On définit le champ de vecteurs

$$F(x,y) = \left(-\frac{y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)$$

pour tout $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$. On le représente en Figure 3.8.

GD : En fait on a parlé de ce champs de vecteur à la toute fin du chapitre 1 : c'est celui dont le rotationnel était nul et qui pourtant n'était pas le gradient d'une fonction, parce qu'alors la fonction aurait dû être la fonction $\theta(x,y) + C$, où $\theta(x,y)$ l'angle du vecteur (x,y) avec l'axe des x. Et comme il ne dérive pas d'un potentiel, on n'est pas trop surpris que ca circulation le long d'un cercle puisse être non nulle. Quand on fait un tour le long du cercle (dans le sens trigonométrique), l'angle $\theta(x,y)$ augmente de 2π, ce qui est cohérent avec le calcul ci-dessous.

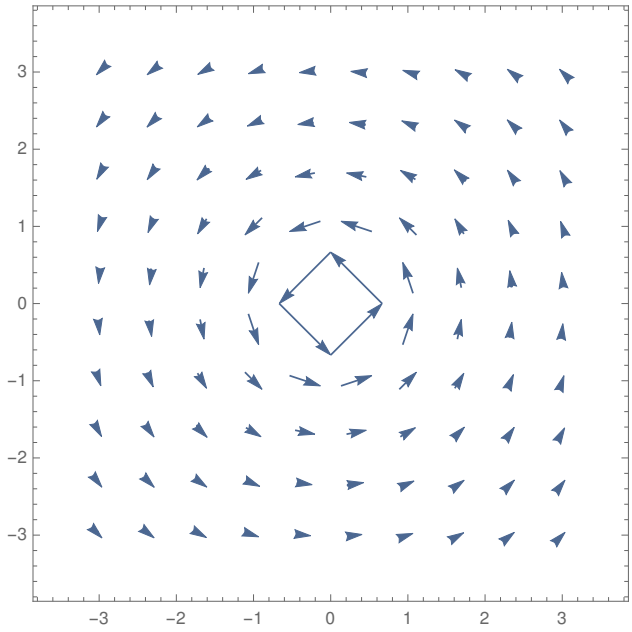

FIGURE 3.8 – Le champ de vecteurs F

Calculons sa circulation le long du cercle centré en $(0,0)$ et de rayon 1. Celui-ci est paramétré par $\widehat{\gamma} = ([0,2\pi], \varphi)$, où

$$\varphi(t) = (\cos(t), \sin(t)), \quad t \in [0, 2\pi].$$

Ainsi,

$$\oint_{\widehat{\gamma}} \vec{F} \cdot \vec{dx} = \int_0^{2\pi} F(\varphi(t)) \cdot \varphi'(t)\, dt = \int_0^{2\pi} \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} \cdot \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} dt$$
$$= \int_0^{2\pi} (\sin^2(t) + \cos^2(t))\, dt = \int_0^{2\pi} 1\, dt = 2\pi.$$

La circulation de F le long de $\widehat{\gamma}$, qui est une courbe fermée, étant non-nulle, on en déduit par le Corollaire 3.3.4 que F n'est pas un champ de gradient (car si F l'était, sa circulation le long de toute courbe fermée serait nulle). C'est particulièrement intéressant, car si on calcule le rotationnel de F, on trouve

$$\operatorname{rot} F(x,y) = \frac{\partial}{\partial x}\frac{x}{x^2+y^2} - \frac{\partial}{\partial y}\frac{-y}{x^2+y^2} = \frac{1}{x^2+y^2} - \frac{2x^2}{(x^2+y^2)^2} + \frac{1}{x^2+y^2} - \frac{2y^2}{(x^2+y^2)^2} = 0.$$

Le rotationnel de F est donc nul, et pourtant F n'est pas un champ de gradient. Cela paraît contredire le Théorème 1.5.3 du Chapitre 1, mais il n'en est rien : en effet, le domaine de définition de F est $\mathbb{R}^2 \setminus \{(0,0)\}$ qui n'est pas un domaine étoilé. Ainsi, le Théorème 1.5.3 ne s'applique pas dans ce cas. Le champ de vecteurs F est donc un exemple de champ de vecteurs de rotationnel nul sur un domaine non-étoilé qui n'est pas un champ de gradient.

Nous résumons dans le tableau suivant les deux notions d'intégrales de fonctions de plusieurs variables le long de courbes que nous avons vues jusqu'à présent.

	Notation	Type de fonction que l'on intègre	Dépend de l'orientation	Formule à partir d'un paramétrage (I, φ)
Intégrale par rapport à la longueur	$\int_\gamma f \, d\ell$	$f : \mathbb{R}^n \to \mathbb{R}$ champ scalaire	Non	$\int_I f(\varphi(t)) \|\varphi'(t)\| \, dt$
Circulation d'un champ de vecteurs	$\int_{\widehat{\gamma}} \vec{F} \cdot \overrightarrow{dx}$	$F : \mathbb{R}^n \to \mathbb{R}^n$ champ vectoriel	Oui	$\int_I F(\varphi(t)) \cdot \varphi'(t) \, dt$

3.4 La formule de Green-Riemann

GD. Je l'ai aussi vu appeler Gauss-Riemann. Ce sera notre première formule d'intégration par parties pour des intégrales multiples (justification et explication plus tard) ! Et c'est très utile (et parfois un peu mystérieux). Je ne peux m'empêcher de décrire un de mes dessins humoristique (humour de matheux) préférés. On y voit un étudiant en thèse désespéré, sur le point de se jeter par la fenêtre parce que visiblement il n'arrive pas à démontrer son théorème, et le directeur de thèse qui lui dit : "peut-être que vous devriez essayer une intégration par parties ?".

Considérons un domaine $\Omega \subset \mathbb{R}^2$, comme représenté en bleu dans la Figure 3.9.

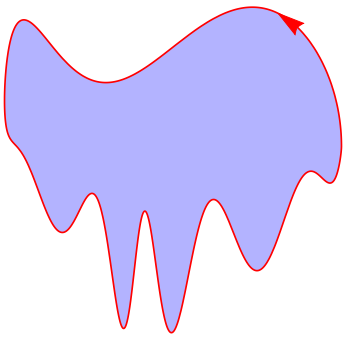

FIGURE 3.9 – Le domaine Ω en bleu et son bord orienté $\widehat{\gamma}$ en rouge.

Le bord de Ω est un chemin (représenté en rouge dans la figure précédente) associé à une courbe paramétrée $\widehat{\gamma}$, que l'on oriente de telle façon à ce que Ω soit à gauche lorsque l'on parcourt $\widehat{\gamma}$. Dans cette partie, on va présenter un résultat permettant de relier l'intégrale

double d'une fonction sur Ω à l'intégrale d'une autre fonction sur $\overset{\frown}{\gamma}$. Il peut arriver que le bord de Ω soit constitué de plusieurs courbes disjointes, comme représenté en Figure 3.10. Dans ce cas, on oriente toutes les courbes de telles manière à ce que Ω reste à gauche lorsqu'on les parcourt.

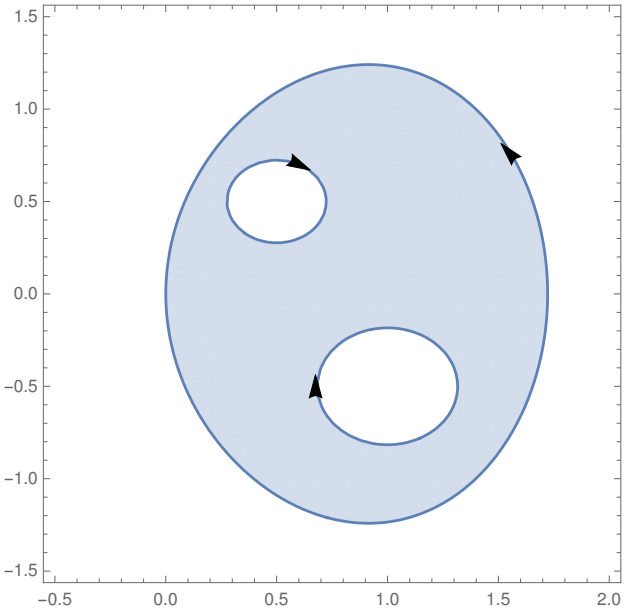

FIGURE 3.10 – Un exemple où le bord de Ω est constitué de plusieurs courbes

Le résultat principal de cette section est le suivant.

Théorème 3.4.1 (Formule de Green-Riemann) *Soit $\Omega \subset \mathbb{R}^2$ un domaine de $\mathbb{R}^2$ dont le bord est formé de n courbes paramétrées $\overset{\frown}{\gamma}_1, \ldots, \overset{\frown}{\gamma}_n$ fermées et orientées de telle sorte que Ω reste à leur gauche lorsqu'on les parcourt. Si $F : \overline{\Omega} \to \mathbb{R}^2$ est un champ de vecteurs de classe C^1, alors*

$$\iint_{\Omega} \operatorname{rot} F(x,y)\, dx\, dy = \sum_{j=1}^{n} \oint_{\overset{\frown}{\gamma}_j} \vec{F} \cdot \vec{dx}. \tag{3.1}$$

Remarque 3.4.1 Le théorème précédent relie deux notions d'intégrales différentes : le terme de gauche de (3.1) représente l'intégrale double de la fonction scalaire de deux variables $\operatorname{rot} F$ vue au Chapitre 2, alors que le terme de droite représente la circulation du champ de vecteurs F le long de courbes, qu'on vient de voir dans ce chapitre.

GD : A défaut de démonstration (désolé, mais une vraie démonstration prendrait longtemps), voici une longue liste de remarques pour essayer de faire que ceci vous paraisse naturel. Ce qui suit n'est pas au programme, mais peut quand même vous aider à comprendre à quel jeu on joue.

D'abord, les choses sont plus simple pour un domaine sans trou, qui est juste le domaine entouré par une seule courbe γ_1 (qu'on a donc paramétrée dans le sens trigonométrique pour que Ω reste à gauche de la courbe). On n'a pas dit, mais ici on suppose en fait que cette

courbe est une boucle simple, c.-à-d. qui ne se recoupe pas comme un ∞, par exemple. Et aussi que cette courbe n'est pas trop compliquée (de classe C^1 par morceaux suffit largement ; sans hypothèse, vous n'imaginez pas à quel point le bord d'un ouvert Ω peut être compliqué.

Maintenant si on sait démontrer le résultat pour un domaine comme ceci, on peut essayer d'en déduire le résultat pour un domaine avec plusieurs trous, comme dans l'énoncé général. Je donne juste le programme ; il y aurait sans quelques vérifications un peu désagréables à faire. Commençons par supposer que Ω est comme dans l'énoncé général, et aussi que le champ de vecteurs F est en fait défini sur $\mathbb{R}^2$ tout entier. Notons nos courbes $\gamma_1, \dots, \gamma_n$. Il se trouve que chaque courbe γ_j sépare le plan en deux parties, l'une qui est à l'intérieur de la courbe, et l'autre à l'extérieur. Pour chaque γ_j, appelons Ω_j le domaine qui se trouve à l'intérieur. C'est un théorème que la situation comme cela, mais vous n'en aurez pas besoin, parce que dans chaque exemple concret que vous aurez, ce sera vrai. Ensuite, il est vrai aussi que l'un des domaines Ω_i, et on va dire que c'est Ω_1, contient toutes les autres courbes γ_j et tout les autres Ω_j. Dans le dessin (et toutes les situations où vous vous truverez), γ_1 est la courbe orientée dans le sens trigo, et toutes les autres sont orientées dans le sens inverse (et Ω, qui est à l'extérieur de ces Ω_j, est bien à gauche de ces γ_j (orientées donc dans le sens inverse). Et maintenant, on peut dire que, modulo les courbes γ_j, mais sur lesquelles la contribution intégrales de $\operatorname{rot} F(x,y)$ sont nulles, on a que Ω_1 est l'union disjointe de Ω et des Ω_j, $j \neq 1$. De sorte que (par la version "intégrales multiples" de Chasles)

$$\iint_{\Omega_1} \operatorname{rot} F(x,y)\,dx\,dy = \iint_{\Omega} \operatorname{rot} F(x,y)\,dx\,dy + \sum_{j\neq 1} \iint_{\Omega_j} \operatorname{rot} F(x,y)\,dx\,dy.$$

Autrement dit,

$$\iint_{\Omega} \operatorname{rot} F(x,y)\,dx\,dy = \iint_{\Omega_1} \operatorname{rot} F(x,y)\,dx\,dy - \sum_{j\neq 1} \iint_{\Omega_j} \operatorname{rot} F(x,y)\,dx\,dy. \tag{3.2}$$

Si on suppose le théorème vrai pour les domaines sans trous, on trouve que

$$\iint_{\Omega_1} \operatorname{rot} F(x,y)\,dx\,dy = \oint_{\overset{\frown}{\gamma_1}} \vec{F}\cdot\vec{dx}.$$

et, pour $j > 1$,

$$\iint_{\Omega_1} \operatorname{rot} F(x,y)\,dx\,dy = -\oint_{\overset{\frown}{\gamma_j}} \vec{F}\cdot\vec{dx},$$

où le signe $-$ vient du fait qu'on a justement inversé l'orientation de γ_j par rapport à celle qui donnait une intégrale sur Ω_j. Il ne nous reste plus qu'à reporter dans (3.2) et on trouve (3.1).

Ensuite, comment passer au cas où F est défini et C^1 seulement sur l'adhérence de $\overline{\Omega}$ (si vous n'aimez pas l'adhérence, demandez, sur Ω et un peu autour), à partir du cas où F est défini sur $\mathbb{R}^2$ tout entier. Une option est de prendre F, et d'en trouver une extension de classe C^1 (c'est-à-dire de trouver une fonction $\widetilde{F}$ de classe C^1 sur $\mathbb{R}^2$ telle que $\widetilde{F}(x) = F(x)$

pour $x \in \overline{\Omega}$. C'est possible avec nos hypothèses (mais il faut un peu de travail pour le faire), et ensuite, on peut appliquer le théorème à $\widetilde{F}$ et s'apercevoir qu'il donne exactement le théorème pour F.

Voilà. Ceci, c'était pour pouvoir se concentrer sur le cas plus simple des domaines sans trous. Maintenant parlons d'un cas encore plus simple, celui où

$$\Omega = I \times J = [a, b] \times [c, d] \tag{3.3}$$

est un rectangle. Je dis que dans ce cas, le théorème n'est rien de plus que la combinaison du théorème de Fubini (un peut forcé, si on veut calculer des intégrales doubles), et la formule unidimensionelle

$$\int_a^b f'(t)\, dt = f(b) - f(a) \tag{3.4}$$

J'a parlé plus haut d'intégration par parties (au lieu de cette formule de base), mais d'une part la différence est maigre (la formule d'IPP se démontre directement en appliquant la formule de base (3.4) à un produit fg, puis en calculant $(fg)'$, et d'autre part les utilisations standard de Green-Riemann et les autres formules qu'on verra dans la suite sont plus directement apparentées à des intégrations par parties.

Démontrons carrément (3.1) dans le cas particulier de (3.3). En fait on doit le faire pour $F = (f_1, f_2)$, mais si on sait le faire avec $F = (f_1, 0)$ et avec $F = (0, f_2)$, vous pouvez vérifier que, comme chacun des deux membres de (3.1) est une fonction linéaure de F, on en déduira aussitôt la formule pour F. De plus, les deux vérifications sont semblables, et donc je ne vais vérifier que le cas de $(f_1, 0)$ et vous laisser faire le cas de $F = (0, f_2)$. Quand $F = (f_1, 0)$, on voit que $\operatorname{rot} F(x, y) = -\frac{\partial f_1}{\partial x}(x, y)$, donc

$$\iint_{\Omega_1} \operatorname{rot} F(x, y)\, dx\, dy = -\iint_{I\times J} \frac{\partial f_1}{\partial x}(x, y)\, dx\, dy = -\int_{x=a}^{b} \Big\{ \int_{y=c}^{d} \frac{\partial f_1}{\partial x}(x, y) dy \Big\} dx \tag{3.5}$$

Maintenant, pour calculer l'intégrale intérieure, on applique la formule de base (3.4) à la fonction $y \mapsto f_1(x, y)$ (où x est fixé), dont la dérivée est justement $\frac{\partial f_1}{\partial x}(x, y)$. On trouve donc

$$\iint_{\Omega_1} \operatorname{rot} F(x, y)\, dx\, dy = -\int_{x=a}^{b} \big\{ f_1(x, d) - f_1(x, c) \big\} dx \tag{3.6}$$

Il nous reste à calculer la circulation de $F = (f_1, 0)$ sur le bord de Ω pour voir si c'est pareil. Notons $A = (a, c)$, $B = (b, c)$, $C = (b, d)$ et $D = (b, c)$ les quatre coins du rectangles. Le bord en question est composé de 4 morceaux mis bout-à-bout. Le premier est le segment horizontal $[AB]$. Le second est le segment vertical $[BC]$, puis $[CD]$ puis $[DA]$. A chaque fois on se donne un paramétrage raisonable du segment, on calcule la circulation de F sur ce segment et à la fin on somme les quatre morceaux. Commençons par $[A, B]$, et appelons γ_1 le chemin correspondant. Le plus simple est de le paramétrer avec l'intervalle I, en posant $\varphi_1(t) = (t, c)$ pour $t \in I$. Alors $\varphi_1'(t) = (1, 0)$, et

$$circ(F; \gamma_1) = \int_I F_1(\varphi_1(t)) \cdot \varphi_1'(t) dt = \int_I F_1(t, c) \cdot \varphi_1'(t) dt = \int_I f_1(t, c) dt = \int_a^b f_1(x, c) dx$$

puisque $F_1(x,y) = (f_1(x,y),0)$ et $\varphi_1'(t) = (1,0)$. Passons à γ_3, qui va de C à D. Le plus simple est de commencer par paramètrer le chemin inverse γ_3 qui va de D à C, parce qu'il est facile à paramétrer par I aussi, en prenant $\varphi_3^*(t) = (t,d)$ pour $t \in I$; comme on sait que c'est la mauvaise orientation on changera juste de signe. On trouve comme ci-dessus que

$$circ(F;\gamma_3) = -circ(F;\gamma_3^*) - \int_I F_1(t,d)\cdot(\varphi_2^*)'(t)dt = -\int_I f_1(t,d)dt = -\int_a^b f_1(x,d)dx.$$

A cause de (3.6), on voit qu'il ne nous reste plus qu'à vérifier que les deeux circulations sur les segments verticaux sont nulles. Mais sur les segments verticaux, la première coordonnée de φ' est nulle (le vecteur vitesse est vertical), et la seconde coordonnée de F est nulle, donc $F(\varphi(x))\cdot\varphi'(x) = 0$ et la circulation, qui est l'intégrale de ce produit, est nulle.

Pour un champ de vecteur $F = (0,f_2)$, on aurait procédé pareil, mais en appliquant Fubini intégrant d'abord en x, et en notant que pour la circulation, c'est les deux segments horizontaux qui ne contribuent pas.

Donc on a une démonstation de (3.1) dans le cas particulier où Ω est un rectangle. Mais on peut en déduire aussi cette formule quand Ω est un rectangle R_1, plus un autre rectangle R_2 collé le long d'un bord. Appelons L la partie commune aux deux bords, qui est un intervalle. On écrit la formule (3.1) pour R_1, la formule pour R_2, on additionne, et on trouve exactement la formule pour $R_1 \cup R_2$, sauf qu'on obtient deux termes supplémentaires, un pour R_1 et l'autre pour R_2, qui viennent de L qui n'est plus dans le bord de $R_1 \cup R_2$. Ca tombe bien, comme R_1 et R_2 sont chacun d'un coté de L, L est en fait orienté de manière opposée en tant que bout de bord de R_1 ou R_2, les contributions de L sont exactement opposées (on calculait la circulation du même F), et donc on a bien démonté la formule pour $R_1 \cup R_2$. Et rien n'empêche de recommencer. On obtient la formule pour tous les domaines Ω qui sont des unions finies de rectangles avec des bords parallèles aux axes. Cela fait en fait pas mal de domaines ; ensuite on peut finir la démonstration par un argument d'approximation par des domaines comme plus haut, ou en remarquant qu'on peut aussi modifier la démonstration dans le cas d'un rectangle pour faire le domaine compris entre deux graphes de fonctions de classe C^1. C'est plutôt comme cela qu'il faudrait faire en vrai (l'approximation est pénible).

J'aime bien l'idée de démonstration par unions, même si elle n'est pas pratique à la fin, parce que si on veut que la formule (3.1) soit vraie pour tous les domaines, c'est en fair pratiquement obligatoire qu'elle soit stable par recollement de deux domaines (comme avec deux rectangles collés), sinon on aurait un problème.

Fin des remarques sur la démonstration. Vous pouvez reprendre le cours normal du poly. Quand même, disons que la formule (3.1) est encore un analogue de la formule unidimensionnelle (3.4) ci-dessus. En effet, les termes de gauche représentent l'intégrale sur un domaine (Ω ou $[a,b]$) d'une fonction faisant intervenir des dérivées (rot F ou f'), alors que le terme de droite fait intervenir les valeurs de la fonction au "bord" du domaine d'intégration ($\widehat{\gamma}$ est le bord de Ω et les points a et b sont en quelque sorte le bord du segment $[a,b]$). Lorsqu'on calcule $f(b) - f(a)$, on considère les valeurs de f au bord du segment $[a,b]$ de la même

manière que lorsqu'on calcule $\oint_{\overset{\frown}{\gamma}} \vec{F} \cdot \overrightarrow{dx}$, on considère les valeurs de F au bord de Ω qui est $\overset{\frown}{\gamma}$.

On peut se servir de la formule de Green-Riemann "dans les deux sens" : à la fois pour calculer des intégrales doubles (terme de gauche dans (3.1)), ou pour calculer des circulations de champs de vecteurs (terme de droite dans (3.1)), comme on l'illustre sur différents exemples.

Exemple 3.4.1 La formule de Green-Riemann peut servir pour calculer l'aire du domaine bordé par une courbe paramétrée. En effet, si $\overset{\frown}{\gamma}$ est une courbe paramétrée orientée fermée, et si Ω est le domaine qu'elle borde sur sa gauche, on a vu que

$$\text{Aire}(\Omega) = \iint_\Omega 1 \, dx \, dy.$$

Pour pouvoir calculer cette intégrale double à l'aide de la formule de Green-Riemann, il faut donc pouvoir écrire $1 = \operatorname{rot} F$ pour un certain champ de vecteurs F. On peut par exemple prendre les champs de vecteurs suivants :

$$F(x,y) = (0,x), \quad G(x,y) = (-y,0), \quad H(x,y) = \frac{1}{2}(-y,x),$$

qui vérifient tous

$$\operatorname{rot} F(x,y) = \operatorname{rot} G(x,y) = \operatorname{rot} H(x,y) = 1$$

pour tout $(x,y) \in \mathbb{R}^2$. Ainsi, par la formule de Green-Riemann,

$$\text{Aire}(\Omega) = \iint_\Omega 1 \, dx \, dy = \begin{cases} \displaystyle\iint_\Omega \operatorname{rot} F(x,y) \, dx \, dy = \oint_{\overset{\frown}{\gamma}} \vec{F} \cdot \overrightarrow{dx} = \oint_{\overset{\frown}{\gamma}} x \, dy \\ \displaystyle\iint_\Omega \operatorname{rot} G(x,y) \, dx \, dy = \oint_{\overset{\frown}{\gamma}} \vec{G} \cdot \overrightarrow{dx} = -\oint_{\overset{\frown}{\gamma}} y \, dx \\ \displaystyle\iint_\Omega \operatorname{rot} H(x,y) \, dx \, dy = \oint_{\overset{\frown}{\gamma}} \vec{H} \cdot \overrightarrow{dx} = \frac{1}{2} \oint_{\overset{\frown}{\gamma}} (x \, dy - y \, dx). \end{cases}$$

Par exemple, calculons l'aire de l'ellipse

$$\Omega = \left\{ (x,y) \in \mathbb{R}^2 \ : \ \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 \leq 1 \right\}$$

pour $a, b > 0$ à l'aide de ces formules. Le bord de cette ellipse est la courbe d'équation

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1$$

que l'on peut paramétrer par

$$\varphi(t) = (a\cos(t), b\sin(t)), \quad t \in [0, 2\pi].$$

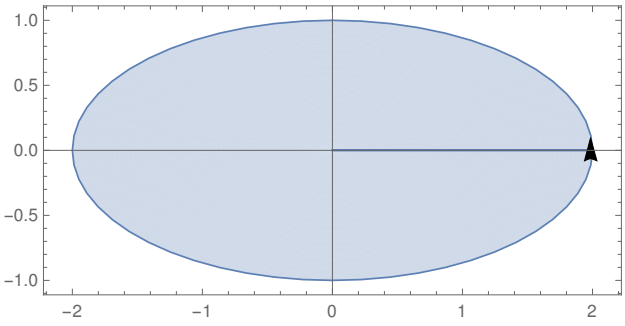

FIGURE 3.11 – L'ellipse Ω pour $a = 2$ et $b = 1$ et son bord orienté

Cette courbe a bien la bonne orientation : en $t = 0$, on a $\varphi(0) = (a, 0)$ et pour $t > 0$ petit, $b\sin(t) > 0$. On se déplace donc "vers le haut" en partant de $(a, 0)$ et donc Ω se situe bien à gauche de la courbe $\overset{\frown}{\gamma}$, comme représenté en Figure 3.11. Par les formules que l'on vient de voir, on a

$$\text{Aire}(\Omega) = \oint_{\overset{\frown}{\gamma}} \vec{H} \cdot \vec{dx} = \int_0^{2\pi} H(\varphi(t)) \cdot \varphi'(t)\, dt.$$

Pour tout $t \in [0, 2\pi]$, on a

$$H(\varphi(t)) = \frac{1}{2}(-b\sin(t), a\cos(t)), \quad \varphi'(t) = (-a\sin(t), b\cos(t)),$$

donc

$$H(\varphi(t)) \cdot \varphi'(t) = \frac{1}{2}(-b\sin(t)(-a\sin(t)) + a\cos(t)b\cos(t)) = \frac{1}{2}ab.$$

Ainsi,

$$\text{Aire}(\Omega) = \int_0^{2\pi} \frac{1}{2}ab\, dt = \pi ab.$$

Pour $a = b = R$, on retrouve bien l'aire d'un disque de rayon R.

Exemple 3.4.2 On peut également se servir de la formule de Green-Riemann pour calculer une circulation de champ de vecteurs. Par exemple, si

$$F(x, y) = (y(1 - x^2 + y^2), x(1 + x^2 - y^2)),$$

calculons la circulation de F le long de $\overset{\frown}{\gamma}$, le cercle d'équation $x^2 + y^2 - 2x = 0$ parcouru dans le sens trigonométrique. La courbe d'équation $x^2 + y^2 - 2x = 0$ est bien un cercle, puisqu'en complétant le carré on trouve

$$x^2 + y^2 - 2x = (x - 1)^2 + y^2 - 1,$$

donc $\overset{\frown}{\gamma}$ représente le cercle de centre $(1, 0)$ et de rayon 1. Ce cercle borde le disque Ω de même centre et de même rayon,

$$\Omega = \{(x, y) \in \mathbb{R}^2 \ : \ (x - 1)^2 + y^2 \leq 1\},$$

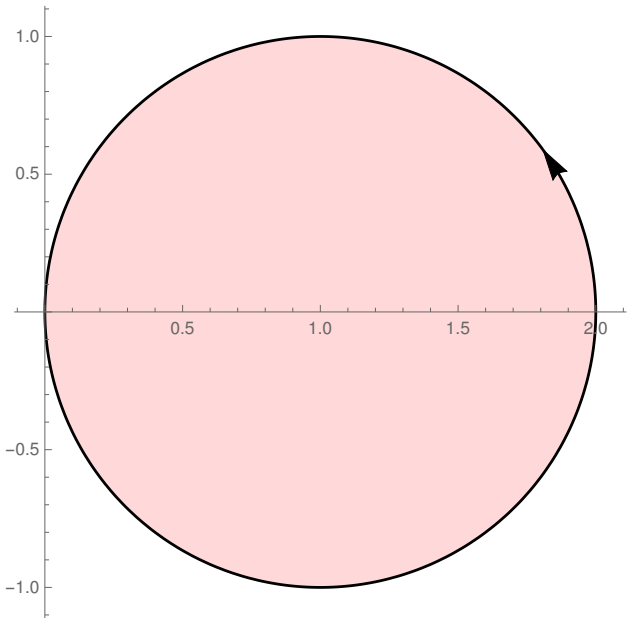

FIGURE 3.12 – Le disque Ω en rose et son bord orienté $\widehat{\gamma}$ en noir

et l'orientation est bien celle donnée par le théorème de Green-Riemann comme représenté en Figure 3.12. Par la formule de Green-Riemann, on a donc

$$\oint_{\widehat{\gamma}} \vec{F} \cdot \overrightarrow{dx} = \iint_{\Omega} \operatorname{rot} F(x,y)\, dx\, dy.$$

Pour tout $(x,y) \in \mathbb{R}^2$, on a

$$\frac{\partial F_2}{\partial x}(x,y) = 1 + x^2 - y^2 + 2x^2 = 1 + 3x^2 - y^2,$$

$$\frac{\partial F_1}{\partial y}(x,y) = 1 - x^2 + y^2 + 2y^2 = 1 - x^2 + 3y^2,$$

et donc

$$\operatorname{rot} F(x,y) = \frac{\partial F_2}{\partial x}(x,y) - \frac{\partial F_1}{\partial y}(x,y) = 4(x^2 - y^2).$$

Pour calculer l'intégrale

$$4 \iint_{\Omega} (x^2 - y^2)\, dx\, dy,$$

on introduit les coordonnées polaires

$$x = 1 + r\cos\theta, \quad y = r\sin\theta,$$

pour $0 \leq r \leq 1$ et $0 \leq \theta \leq 2\pi$. Comme pour les coordonnées polaires usuelles, le jacobien du

changement de variables vaut r. Ainsi, en passant en coordonnées polaires,

$$\begin{aligned}
4\iint_\Omega (x^2-y^2)\,dx\,dy &= 4\int_0^{2\pi}\int_0^1 ((1+r\cos\theta)^2 - r^2\sin^2\theta)\,r\,dr\,d\theta \\
&= 4\int_0^{2\pi}\int_0^1 (1+2r\cos\theta + r^2(\cos^2\theta-\sin^2\theta)\,r\,dr\,d\theta \\
&= 4\int_0^{2\pi} (\frac{1}{2}+\frac{2}{3}\cos\theta + \frac{1}{4}(\cos^2\theta-\sin^2\theta))\,d\theta \\
&= 4\pi + \int_0^{2\pi}\cos(2\theta)\,d\theta \\
&= 4\pi.
\end{aligned}$$

Ainsi, on trouve que

$$\oint_{\widehat{\gamma}} \vec{F}\cdot\overrightarrow{dx} = 4\pi.$$

Remarque 3.4.2 Si jamais l'orientation de $\widehat{\gamma}$ n'est pas la bonne par rapport à la formule de Green-Riemann (c'est-à-dire si Ω est à droite de $\widehat{\gamma}$ et pas à gauche), il suffit de renverser l'orientation. En effet, dans ce cas Ω est à gauche de $\overleftarrow{\gamma}$ (qui est $\widehat{\gamma}$ parcouru dans le sens inverse) et donc on peut appliquer la formule de Green-Riemann à $\overleftarrow{\gamma}$: on a donc

$$\oint_{\overleftarrow{\gamma}} \vec{F}\cdot\overrightarrow{dx} = \iint_\Omega \operatorname{rot} F(x,y)\,dx\,dy.$$

Comme de plus on a vu que

$$\oint_{\overleftarrow{\gamma}} \vec{F}\cdot\overrightarrow{dx} = -\oint_{\widehat{\gamma}} \vec{F}\cdot\overrightarrow{dx},$$

on en déduit que dans ce cas on a

$$\oint_{\widehat{\gamma}} \vec{F}\cdot\overrightarrow{dx} = -\iint_\Omega \operatorname{rot} F(x,y)\,dx\,dy.$$

La formule de Green-Riemann doit donc s'écrire avec un signe moins si Ω est à droite de $\widehat{\gamma}$.

Chapitre 4

Intégrales surfaciques

Dans ce chapitre, nous allons voir la notion d'intégrale d'une fonction sur une surface (objet à deux dimension) dans $\mathbb{R}^3$: par exemple une sphère, un plan, un cylindre... Une notion importante sera celle de flux d'un champ de vecteur à travers une surface, utilisée en électromagnétisme, mécanique des fluides, etc. Dans le chapitre précédent, on a vu la notion d'intégrale le long de courbes. On verra ici une notion similaire en remplaçant les courbes par des surfaces (donc en passant d'un objet de dimension 1 à un objet de dimension 2). On commence donc par définir une surface paramétrée, de la même manière qu'on avait commencé par définir une courbe paramétrée au chapitre précédent.

4.1 Surfaces paramétrées

Au chapitre précédent, on avait préféré le point de vue "analytique" en définissant des courbes à l'aide de fonctions les paramétrant (les paramétrages), plutôt que par leur aspect plus visuel (leur image géométrique). On procède de même ici.

Définition 4.1.1: Surface paramétrée

On appelle *surface paramétrée de* $\mathbb{R}^3$ un couple (Ω, X) où $\Omega \subset \mathbb{R}^2$ est un domaine de $\mathbb{R}^2$ et $X : \Omega \to \mathbb{R}^3$ est une application de classe C^1. Si (Ω, X) est une surface paramétrée de $\mathbb{R}^3$, son *image géométrique* est l'ensemble

$$\mathcal{S} = \{X(u, v) \ : \ (u, v) \in \Omega\}.$$

On dit que (Ω, X) est un *paramétrage* de $\mathcal{S}$.

Comme pour les courbes paramétrées, on se contentera de connaître un certain nombre de surfaces paramétrées "usuelles" de $\mathbb{R}^3$.

GD : La relation entre l'image géométrique et la surface paramétré est la même que pour les courbes. On fera tous nos calculs à partir du paramétrage, mais on pense souvent très fort à l'image géométrique. Ainsi, on donnera une formule pour calculer l'intégrale d'une

fonction sur une sphère à partir d'un paramétrage de la sphère, et tout ceci reposera sur un certain nombre de faits qu'on ne vérifiera pas dans le détail. D'abord, le paramétrage qu'on donnera (coordonnées sphériques, voir ci-dessous), est bien une bijection, sauf pour une partie de la sphère (comme un unique méridien) qui est de surface nulle, et où il y a un recollement. Donc (sauf sur le bout de recollement qui ne compte pas), on s'assure qu'on recouvre bien toute la sphère, et aussi qu'on na passe pas trois fois au même endroit (ce qui pourait donner une intégrale trois fois trop grande). Ensuite, il y a aussi le même théorème que pou les courbes, qui dit que si on a deux paramétrages équivalents (ici, de la sphère), le calcul donne le même résultat pour les deux (voir plus loin)

Définition 4.1.2: Surfaces paramétrées usuelles de $\mathbb{R}^3$

1. On appelle *morceau de plan de* $\mathbb{R}^3$ une surface paramétrée (Ω, X) avec Ω de la forme $\Omega = [0,1]^2$ et
$$X(u,v) = A + ue_1 + ve_2, \quad (u,v) \in \Omega,$$
où A est un point de $\mathbb{R}^3$ et $\{e_1, e_2\}$ est une famille libre de $\mathbb{R}^3$. L'image géométrique de (Ω, X) est un morceau du plan affine $A + \mathrm{Vect}\{e_1, e_2\}$ de $\mathbb{R}^3$.
2. On appelle *graphe de* f, où $f : \Omega \to \mathbb{R}$ est C^1 et $\Omega \subset \mathbb{R}^2$, la surface paramétrée (Ω, X), où
$$X(u,v) = (u, v, f(u,v)), \quad (u,v) \in \Omega.$$
3. On utilisera volontiers le paramétrage suivant de la sphère de $\mathbb{R}^3$ centrée en $A \in \mathbb{R}^3$ et de rayon $R > 0$: on prend $\Omega = [0,\pi] \times [0, 2\pi[$ et
$$X(\theta,\varphi) = (x_0 + R\cos\varphi\sin\theta, y_0 + R\sin\varphi\sin\theta, z_0 + R\cos\theta), \ (\theta,\varphi). \in \Omega,$$
(On a donc noté $x_0 + R\cos\varphi\sin\theta$ la première coordonnée de $X(\theta,\varphi)$, etc. C'est évidemment ce qui sort des coordonnées sphériques décrites ci-dessus, en fixant R).
4. On utilise aussi le paramétrage du *cylindre vertical de* $\mathbb{R}^3$ *centré en* $A(x_0, y_0, 0) \in \mathbb{R}^3$ *de rayon* $R > 0$, qui est le couple (Ω, X) avec $\Omega = [0, 2\pi[\times \mathbb{R}$ et
$$X(\theta, z) = (x_0 + R\cos\theta, y_0 + R\sin\theta, z), \quad (\theta, z) \in \Omega.$$

Exemple 4.1.3 On remarque la similarité entre le paramétrage des plans et celui des segments vu au chapitre précédent ($\varphi(t) = A + t\overrightarrow{AB}$, $t \in [0,1]$). Remarquons aussi qu'on a décrit un plan à l'aide d'un point A et d'une base $\{e_1, e_2\}$ de l'espace vectoriel associé. Si on sait que le plan passe par trois points (non alignés), A, A_1, et A_2, on peut prendre $e_1 = A_1 - A$ et $e_2 = A_2 - A$ (faire la différence coordonnée par coordonnée).

Mais on peut aussi le faire à l'aide d'une équation cartésienne. Il faut alors en trouver une base pour pouvoir déterminer le paramétrage : par exemple, si l'on veut paramétrer le morceau $\mathcal{P}$ de plan d'équation $x - y + 2z = 3$ pour $0 \leq y \leq 1$ et $-1 \leq z \leq 1$, on rappelle

que

$$\begin{aligned}\mathcal{P} &= \{(x,y,z) \in \mathbb{R}^3 \ : \ x-y+2z=3,\ 0\le y\le 1,\ -1\le z\le 1\},\\ &= \{(3+y-2z,y,z) \ : \ 0\le y\le 1,\ -1\le z\le 1\},\\ &= \{(3,0,0)+y(1,1,0)+z(-2,0,1) \ : \ 0\le y\le 1,\ -1\le z\le 1\},\end{aligned}$$

et donc on peut choisir $\Omega = [0,1]\times[-1,1]$ et

$$X(u,v) = (3,0,0)+u(1,1,0)+v(-2,0,1), \quad (u,v)\in[0,1]\times[-1,1].$$

Dans notre définition, on avait dit qu'on prenait $\Omega = [0,1]^2$ pour des portions de plan : c'est possible ici en reparamétrant la variable v et en choisissant donc

$$\begin{aligned}\tilde{X}(u,w) &= (3,0,0)+u(1,1,0)+(2w-1)(-2,0,1)\\ &= (5,0,-1)+u(1,1,0)+w(-4,0,2), \quad (u,w)\in[0,1]^2\end{aligned}$$

où l'on a posé $v = 2w-1$ (et donc v parcourt $[-1,1]$ si et seulement si w parcourt $[0,1]$). Un paramétrage avec $\Omega = [0,1]^2$ est facile à se représenter géométriquement : $\mathcal{P}$ est ici un parallélogramme basé en $(5,0,-1)$ et de côtés $(1,1,0)$ et $(-4,0,2)$.

GD : pour un plan d'équation $2x+3y+z=4$, par exemple, il se trouve que le vecteur $V=(2,3,1)$ est orthogonal au plan. C'est assez facile à voir, parce que l'ensemble des points du plan est l'ensemble des vecteurs $X=(x,y,z)$ tels que le produit scalaire de X avec V vaut 4. Ajouter un vecteur du plan vectoriel parallèle ne doit pas changer le produit scalaire. Donc si on cherche une base du plan tangent vectoriel, il s'agit de prendre deux vecteurs indépendants tous les deux orthogonaux à V. Bon, on pouvait aussi choisir 3 points comme ci-dessus.

Exemple 4.1.4 Comme pour les courbes paramétrées, on peut parfois restreindre le domaine Ω pour considérer des sous-surfaces. Par exemple, si l'on veut paramétrer le graphe de $f : (u,v)\in\mathbb{R}^2 \mapsto u^2+v^2$ correspondant à $u^2+v^2\le 1$, on posera $\Omega = \{(u,v)\in\mathbb{R}^2 \ : \ u^2+v^2\le 1\}$ et $X(u,v) = (u,v,u^2+v^2)$ pour $(u,v)\in\Omega$. De même, si on veut paramétrer une demi-sphère centrée en $(0,0,0)$ et de rayon 1 correspondant à $z\ge 0$, on prendra

$$X(\theta,\varphi) = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta),$$

avec $(\theta,\varphi)\in\Omega = [0,\pi/2]\times[0,2\pi[$. Si l'on veut considérer le cylindre centré en $(0,0,0)$ de rayon 1 uniquement pour z entre -1 et 2, on prendra

$$X(\theta,z) = (\cos\theta,\sin\theta,z)$$

pour $(\theta,z)\in[0,2\pi[\times[-1,2]$.

Définition 4.1.5: Plan tangent d'une surface paramétrée

Soit $S = (\Omega, X)$ une surface paramétrée de $\mathbb{R}^3$ et $(u_0, v_0) \in \Omega$. On appelle *plan tangent vectoriel à S en $X(u_0, v_0)$* le plan vectoriel

$$\text{Vect}\left\{\frac{\partial X}{\partial u}(u_0, v_0), \frac{\partial X}{\partial v}(u_0, v_0)\right\}.$$

Remarque 4.1.1 On a une notion similaire pour les courbes paramétrées : la droite tangente à une courbe (I, φ) au point $\varphi(t_0)$ est $\text{Vect}\{\varphi'(t_0)\}$: la droite vectorielle engendrée par $\varphi'(t_0)$. Elle est bien géométriquement tangente (à une translation près) à l'image géométrique de (I, φ). De même, le plan tangent est géométriquement tangent à l'image géométrique d'une surface paramétrée.

GD : Pour que ceci marche bien, on suppose que les deux vecteurs $\frac{\partial X}{\partial u}(u_0, v_0)$ et $\frac{\partial X}{\partial v}(u_0, v_0)$ sont indépendants ; sinon ils n'engendrent pas un plan, et on s'attend à une situation géométrique plus compliquée. Ceci reste vrai pour un autre paramétrage équivalent (voir ci-dessous), et le plan tangent calculé ainsi reste le même (les vecteurs changent, mais l'espace engendré reste le même). Et de plus on montrerait assez aisément que les points de l'image géométrique sont assez proche du plan tangent $X(u_0, v_0) + P$, où P est le plan tangent vectoriel. Quand nos deux vecteurs sont indépendants, on appelle $X(u_0, v_0)$ un point régulier de S (ou encore mieux, (u_0, v_0) est un point régulier de Ω).

On a vu quelque chose de semblable avec les courbes, où l'on supposait que $\varphi'(u_0) \neq 0$ avant de parler de la tangente à γ en $\varphi(u_0)$. C'est un peu pareil ici, sauf qu'on cherche un espace vectoriel de dimension 2.

GD : je vous laisse vérifier que quand on paramètre le graphe d'une fonction $f : \Omega \to \mathbb{R}$, le plan tangent qu'on vient de définir est aussi celui qu'on aurait pu définir pour un graphe de manière naturelle. C'est aussi le graphe de la fonction affine qui approxime f au point $(u_0, v_0, f(u_0, v_0))$ (comme dans la formule de Taylor à l'ordre 1).

Exemple 4.1.6 1. Si (Ω, X) paramètre une portion de plan, c'est-à-dire si (Ω, X) est de la forme $\Omega = [0, 1]^2$ et $X(u, v) = A + ue_1 + ve_2$, alors

$$\frac{\partial X}{\partial u}(u, v) = e_1, \quad \frac{\partial X}{\partial v}(u, v) = e_2,$$

et donc le plan tangent en $X(u_0, v_0)$ est

$$\text{Vect}\left\{\frac{\partial X}{\partial u}(u_0, v_0), \frac{\partial X}{\partial v}(u_0, v_0)\right\} = \text{Vect}\{e_1, e_2\}.$$

On voit ici que le plan tangent ne dépend pas du point de base $X(u_0, v_0)$. C'est géométriquement clair : le plan tangent à un plan est lui-même (à une translation près) !

2. Si (Ω, X) paramètre un graphe de $f : \Omega \to \mathbb{R}$, c'est-à-dire si $X(u,v) = (u, v, f(u,v))$ avec $(u,v) \in \Omega$, alors

$$\frac{\partial X}{\partial u}(u,v) = (1, 0, \frac{\partial f}{\partial u}(u,v)), \quad \frac{\partial X}{\partial v}(u,v) = (0, 1, \frac{\partial f}{\partial v}(u,v)),$$

et donc le plan tangent au point $X(u_0, v_0)$ est

$$\text{Vect}\{(1, 0, \frac{\partial f}{\partial u}(u_0,v_0)), (0, 1, \frac{\partial f}{\partial v}(u_0,v_0))\}.$$

On peut montrer que celui-ci est d'équation cartésienne

$$z = \frac{\partial f}{\partial u}(u_0,v_0)x + \frac{\partial f}{\partial v}(u_0,v_0)y,$$

ce qui correspond à la formule obtenue avec la formule de Taylor à l'ordre 1 dans le chapitre 1 (à une translation près).

3. La sphère de centre (x_0, y_0, z_0) et de rayon R est paramétrée par

$$X(\theta, \varphi) = (x_0 + R\cos\varphi\sin\theta, y_0 + R\sin\varphi\sin\theta, z_0 + R\cos\theta),$$

avec $(\theta, \varphi) \in [0, \pi] \times [0, 2\pi[$. Ce paramétrage vérifie

$$\begin{cases} \dfrac{\partial X}{\partial \theta}(\theta,\varphi) = (R\cos\varphi\cos\theta, R\sin\varphi\cos\theta, -R\sin\theta), \\ \\ \dfrac{\partial X}{\partial \varphi}(\theta,\varphi) = (-R\sin\varphi\sin\theta, R\cos\varphi\sin\theta, 0) \end{cases},$$

et donc le plan tangent à cette sphère en $X(\theta_0, \varphi_0)$ est

$$\text{Vect}\{(R\cos\varphi_0\cos\theta_0, R\sin\varphi_0\cos\theta_0, -R\sin\theta_0), (-R\sin\varphi_0\cos\theta_0, R\cos\varphi_0\sin\theta_0, 0)\}.$$

Ce plan s'interprète géométriquement de la façon suivante : en remarquant que

$$\frac{\partial X}{\partial \theta}(\theta,\varphi) \cdot (X(\theta,\varphi) - (x_0, y_0, z_0)) = 0, \quad \frac{\partial X}{\partial \varphi}(\theta,\varphi) \cdot (X(\theta,\varphi) - (x_0, y_0, z_0)) = 0,$$

on en déduit que le plan tangent en $X(\theta_0, \varphi_0)$ n'est autre que le plan orthogonal au vecteur $X(\theta,\varphi) - (x_0, y_0, z_0)$.

4. Le cylindre centré en $(x_0, y_0, 0)$ de rayon R est paramétré par

$$X(\theta, z) = (x_0 + R\cos\theta, y_0 + R\sin\theta, z),$$

pour $(\theta, z) \in [0, 2\pi[\times \mathbb{R}$. On a donc

$$\frac{\partial X}{\partial \theta}(\theta, z) = (-R\sin\theta, R\cos\theta, 0), \quad \frac{\partial X}{\partial z}(\theta, z) = (0, 0, 1).$$

Ainsi, le plan tangent au cylindre en $X(\theta_0, z_0)$ est

$$\mathrm{Vect}\{(-R\sin\theta, R\cos\theta, 0), (0,0,1)\}.$$

Le premier vecteur correspond au vecteur tangent au cercle de rayon R dans le plan $(0xy)$, le second vecteur montre que le plan tangent contient toujours l'axe $(0z)$, ce qui est clair géométriquement.

Remarque 4.1.2 Pour que le plan tangent soit bien un plan (c'est-à-dire un espace vectoriel de dimension 2), on voit qu'il faut que la famille $\{\frac{\partial X}{\partial u}(u_0, v_0), \frac{\partial X}{\partial v}(u_0, v_0)\}$ soit libre. Lorsque c'est le cas, on dit que le point $X(u_0, v_0)$ est un point *régulier* de S.

4.2 Aire d'une surface, intégration par rapport à l'aire d'une surface

De la même manière que pour les courbes paramétrées, on peut calculer l'aire d'une surface paramétrée à l'aide de son paramétrage. La formule pour l'élément d'intégration est légèrement plus compliquée.

Définition 4.2.1

Soit $S = (\Omega, X)$ une surface paramétrée de $\mathbb{R}^3$. L'*aire de* S est définie par

$$\mathrm{Aire}(S) = \iint_\Omega \left\| \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \right\| du\, dv,$$

où l'on rappelle que $\wedge$ désigne le produit vectoriel, dont la définition est rappelée dans le chapitre 2 du cours d'algèbre linéaire.

Remarque 4.2.1 Cette formule est de nature similaire à celle de la longueur d'une courbe

$$\ell(\gamma) = \int_I \|\varphi'(t)\|\, dt.$$

Exemple 4.2.2 Cette formule se comprend bien lorsque S est un morceau de plan : en effet dans ce cas on a vu que $\frac{\partial X}{\partial u}(u,v) = e_1$ et $\frac{\partial X}{\partial v}(u,v) = e_2$ pour tout $(u,v) \in [0,1]^2$ et donc

$$\mathrm{Aire}(S) = \int_0^1 \int_0^1 \|e_1 \wedge e_2\|\, du\, dv = \|e_1 \wedge e_2\|.$$

Cette dernière quantité représente bien l'aire du parallélogramme de côtés e_1 et e_2 comme on le justifie à présent. Par définition du produit vectoriel, on a

$$\det(e_1, e_2, e_1 \wedge e_2) = \|e_1 \wedge e_2\|^2.$$

Comme $\det(e_1, e_2, e_1 \wedge e_2)$ représente le volume du parallélépipède de côtés e_1, e_2, $e_1 \wedge e_2$ et que $e_1 \wedge e_2$ est orthogonal à e_1 et à e_2, on en déduit que

$$\text{Volume(parallélépipède de côtés } e_1,\ e_2,\ e_1 \wedge e_2) = \\ \text{Aire(parallélogramme de côtés } e_1,\ e_2) \times \text{longueur de } e_1 \wedge e_2$$

et donc

$$\text{Volume(parallélépipède de côtés } e_1,\ e_2,\ e_1 \wedge e_2) = \\ = \text{Aire(parallélogramme de côtés } e_1,\ e_2) \times \|e_1 \wedge e_2\|.$$

Ainsi,

$$\begin{aligned}\text{Aire(parallélogramme de côtés } e_1,\ e_2) &= \frac{\text{Volume(parallélépipède de côtés } e_1,\ e_2,\ e_1 \wedge e_2)}{\|e_1 \wedge e_2\|} \\ &= \frac{\det(e_1, e_2, e_1 \wedge e_2)}{\|e_1 \wedge e_2\|} \\ &= \|e_1 \wedge e_2\|.\end{aligned}$$

La formule dans le cas général se comprend alors en approchant S par des "plans infinitésimaux" d'aires $\|\frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v)\|$.

GD : Ici on utilise le produit vectoriel parce que c'est pratique, et aussi parce qu'il donne une orientation de la surface qui va se révéler utile. Pour des surfaces de dimension 2 dans $\mathbb{R}^n$, $n > 2$, on peut encore prendre $e_1 = \frac{\partial X}{\partial u}(u,v)$ et $e_2 = \frac{\partial X}{\partial v}(u,v)$, calculer l'aire du palallèlogramme tracé dans $\mathbb{R}^n$ (en fait dans le plan engendré par ces deux vecteurs), et calculer avec ça. Autrement dit on peut remplacer $\|\frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v)\|$ par $\|\frac{\partial X}{\partial u}(u,v)\|\,\|\frac{\partial X}{\partial v}(u,v)\|\,|\cos(\alpha)|$, où α est l'angle des deux vecteurs. Et pour des surfaces de dimension 3 ou plus on utiliserait l'aire d'un parallèlépipède de dimension 3. Par contre, pour les histoires qui suivent avec l'orientation, l'utilisation du produit vectoriel est bien pratique.

GD : Il est important, pour la cohérence de définitions, de s'assurer que lorsqu'on paramètre S de deux manières équivalentes, alors on trouve la même surface dans la définition ci-dessus. Commençons par la définition d'équivalence. Soit (Ω, X) un paramétrage de classe C^1. On dit que $(\widetilde{\Omega}, \widetilde{X})$ est un paramétrage équivalent quand il existe une bijection $\theta : \Omega \to \widetilde{\Omega}$, de classe C^1, dont l'inverse $\theta^{-1} : \widetilde{\Omega} \to \Omega$ est également de classe C^1, et aussi qui préserve l'orientation, telle que $X = \widetilde{X} \circ \theta$ sur Ω. Comparez à la définition 3.1.3.

Plus haut on a demandé que θ soit croissante, et maintenant on demande qu'elle préserve l'orientation. Cela se vérifie aisément en pratique : les bijections C^1 (dont la réciproque est aussi C^1) qui préservent l'orientation sont celles dont la matrice Jacobienne (la matrice carrrée $((\frac{\partial \theta_i}{\partial u_j}))$), qui est inversible par définition (l'inverse est C^1), a un déterminant positif. Si le déterminant est négatif, θ renverse l'orientation. Ainsi, le changement de variable $(x,y,z) \mapsto (y,x,z)$ renverse l'orientation, alors que $(x,y,z) \mapsto (y,z,x)$ la préserve. Noter

que comme le déterminant du jacobien ne s'annule pas et est une fonction continue sur Ω, son signe ne change pas (si on prend Ω connexe), donc la notion a un sens.

Le fait que l'aire $\text{Aire}(S) = \iint_\Omega \left\| \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \right\| du\, dv$, ne change pas quand on remplace (Ω, X) par un paramétrage équivalent se démontre sans trop de mal (un peu quand même, donc on passe) en faisant le changement de variable dans l'intégrale double.

En fait, l'aire ne change toujours pas quand on compose avec un changement de variable qui renverse l'orientation, mais par contre dans la suite, quand on calculera des flux, le signe changera (donc il faudra faire attention à l'orientation), comme pour les circulations de champs de vecteurs.

Exemple 4.2.3 On calcule l'aire dans le cas des surfaces usuelles définies plus haut. On a déjà calculé l'aire d'un plan. On calcule à présent les aires des autres surfaces.

1. Si S est un graphe d'une fonction $f : \Omega \to \mathbb{R}$, on a vu que

$$\frac{\partial X}{\partial u}(u,v) = (1, 0, \frac{\partial f}{\partial u}), \quad \frac{\partial X}{\partial v}(u,v) = (0, 1, \frac{\partial f}{\partial v}),$$

et donc

$$\frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) = (-\frac{\partial f}{\partial u}(u,v), -\frac{\partial f}{\partial v}(u,v), 1),$$

et donc

$$\left\| \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \right\| = \sqrt{1 + \left(\frac{\partial f}{\partial u}(u,v)\right)^2 + \left(\frac{\partial f}{\partial u}(u,v)\right)^2}.$$

Ainsi,

$$\text{Aire}(S) = \iint_\Omega \sqrt{1 + \left(\frac{\partial f}{\partial u}(u,v)\right)^2 + \left(\frac{\partial f}{\partial v}(u,v)\right)^2}\, du\, dv.$$

On peut comparer cette formule à celle qui donne la longueur d'une courbe donnée par un graphe,

$$\ell(\gamma) = \int_I \sqrt{1 + f'(t)^2}\, dt.$$

2. Dans le cas d'une sphère de rayon R, on a vu que

$$\begin{cases} \dfrac{\partial X}{\partial \theta}(\theta,\varphi) = (R\cos\varphi\cos\theta, R\sin\varphi\cos\theta, -R\sin\theta), \\ \dfrac{\partial X}{\partial \varphi}(\theta,\varphi) = (-R\sin\varphi\sin\theta, R\cos\varphi\sin\theta, 0) \end{cases},$$

et donc

$$\frac{\partial X}{\partial \theta}(\theta,\varphi) \wedge \frac{\partial X}{\partial \varphi}(\theta,\varphi) = (R^2\cos\varphi\sin^2\theta, R^2\sin\varphi\sin^2\theta, R^2\cos\theta\sin\theta),$$

ainsi que

$$\left\| \frac{\partial X}{\partial \theta}(\theta, \varphi) \wedge \frac{\partial X}{\partial \varphi}(\theta, \varphi) \right\| = R^2 \sin\theta.$$

On trouve donc que

$$\text{Aire}(S) = \int_0^{2\pi} \int_0^{\pi} R^2 \sin\theta \, d\theta \, d\varphi = 4\pi R^2,$$

qui est bien l'aire d'une sphère de rayon R.

3. Soit S un cylindre de rayon R compris entre z_0 et $z_0 + h$. On a vu que

$$\frac{\partial X}{\partial \theta}(\theta, z) = (-R\sin\theta, R\cos\theta, 0), \quad \frac{\partial X}{\partial z}(\theta, z) = (0, 0, 1),$$

et donc

$$\frac{\partial X}{\partial \theta}(\theta, z) \wedge \frac{\partial X}{\partial z}(\theta, z) = (R\cos\theta, R\sin\theta, 0),$$

ainsi que

$$\left\| \frac{\partial X}{\partial \theta}(\theta, z) \wedge \frac{\partial X}{\partial z}(\theta, z) \right\| = R.$$

On trouve donc que

$$\text{Aire}(S) = \int_{z_0}^{z_0+h} \int_0^{2\pi} R \, d\theta \, dz = 2\pi R h,$$

ce qui est bien l'aire d'un cylindre de rayon R et de hauteur h.

Exemple 4.2.4 Calculons l'aire de la portion du graphe de $z = xy$ comprise à l'intérieur du cylindre d'équation $x^2 + y^2 = 1$. Comme S est un graphe de la fonction $f : (u, v) \mapsto uv$, on peut le paramétrer par $S = (\Omega, X)$ avec

$$X(u, v) = (u, v, f(u, v)) = (u, v, uv),$$

pour tout $(u, v) \in \Omega$, où Ω est choisi pour que S reste à l'intérieur du cylindre $x^2 + y^2 = 1$. L'intérieur de ce cylindre a pour équation $x^2 + y^2 \leq 1$, et donc on choisir Ω afin que pour tout $(u, v) \in \Omega$, on ait

$$X(u, v) \in \{(x, y, z) \in \mathbb{R}^3 \ : \ x^2 + y^2 \leq 1\},$$

c'est-à-dire

$$\Omega = \{(u, v) \in \mathbb{R}^2 \ : \ u^2 + v^2 \leq 1\},$$

et donc Ω est le disque centré en $(0, 0)$ de rayon 1 dans $\mathbb{R}^2$. Comme S est un graphe, on a vu que

$$\left\| \frac{\partial X}{\partial u}(u, v) \wedge \frac{\partial X}{\partial v}(u, v) \right\| = \sqrt{1 + \left(\frac{\partial f}{\partial u}(u, v) \right)^2 + \left(\frac{\partial f}{\partial u}(u, v) \right)^2}.$$

Ici, comme $f(u,v) = uv$ on a

$$\frac{\partial f}{\partial u}(u,v) = v, \quad \frac{\partial f}{\partial v}(u,v) = u,$$

et donc

$$\left\| \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \right\| = \sqrt{1+v^2+u^2}.$$

Ainsi, on en déduit que

$$\text{Aire}(S) = \iint_\Omega \sqrt{1+u^2+v^2}\, du\, dv.$$

Il est naturel de calculer cette intégrale en coordonnées polaires, en posant $u = r\cos\theta$ et $v = r\sin\theta$. On a

$$(u,v) \in \Omega \Longleftrightarrow 0 \leq r \leq 1,\ 0 \leq \theta \leq 2\pi,$$

et comme $du\, dv = r\, dr\, d\theta$ et que $u^2+v^2 = r^2$ on en déduit que

$$\begin{aligned} \text{Aire}(S) &= \int_0^{2\pi}\int_0^1 \sqrt{1+r^2} r\, dr\, d\theta \\ &= 2\pi[\frac{1}{3}(1+r^2)^{3/2}]_{r=0}^{r=1} \\ &= \frac{2\pi}{3}(2\sqrt{2}-1). \end{aligned}$$

On peut également intégrer une fonction scalaire par rapport à l'aire d'une surface.

Définition 4.2.5: Intégration d'une fonction scalaire par rapport à l'aire d'une surface

Soit $S = (\Omega, X)$ une surface paramétrée de $\mathbb{R}^3$ et $f : \mathbb{R}^3 \to \mathbb{R}$ une fonction continue. L'*intégrale de f par rapport à l'aire de S*, notée $\iint_S f\, dS$, est le nombre

$$\iint_S f\, dS = \iint_\Omega f(X(u,v)) \left\| \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \right\| du\, dv.$$

C'est pourquoi on utilise la notation pour l'*élément d'aire sur S*,

$$dS = \left\| \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \right\| du\, dv.$$

Remarque 4.2.2 Cette formule est à comparer avec l'intégrale d'une fonction scalaire par rapport à la longueur d'une courbe $\gamma = (I, \varphi)$,

$$\int_\gamma f\, d\ell = \int_I f(\varphi(t)) \|\varphi'(t)\|\, dt.$$

Remarque 4.2.3 On a

$$\text{Aire}(S) = \iint_S 1\, dS.$$

GD : Et les remarques générales qui ont été faites sur l'aire valent encore ici : si on prend un autre paramétrage équivalent $(\widetilde{\Omega}, \widetilde{X})$ de S, on trouve le même nombre $\iint_S f\,dS$. C'est ce qui autorise la notation (autrement, on serait obligé de parler d'intégrale de f pour un paramétrage donné). Et même, si on compose avec un changement de paramétrage qui renverse l'orientation, on ne change toujours pas $\iint_S f\,dS$. Ce sera différent pour les flux.

Exemple 4.2.6 Calculons $\iint_S x^2\,dS$ où S est la sphère centrée en $(0,0,0)$ de rayon 1. On a vu que dans ce cas on a $S = (\Omega, X)$ avec $\Omega = [0,\pi] \times [0, 2\pi[$ et

$$X(\theta,\varphi) = (\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta).$$

De plus, on a déjà calculé l'élément d'aire sur la sphère de rayon 1,

$$dS = \sin\theta\,d\theta\,d\varphi.$$

Ici, la fonction $f : \mathbb{R}^3 \to \mathbb{R}$ à intégrer est

$$f(x,y,z) = x^2, \quad (x,y,z) \in \mathbb{R}^3.$$

On a donc pour tout $(\theta,\varphi) \in \Omega$ on a

$$f(X(\theta,\varphi)) = f(\cos\varphi\sin\theta, \sin\varphi\sin\theta, \cos\theta) = \cos^2\varphi\sin^2\theta.$$

L'intégrale de f par rapport à l'aire de la sphère vaut donc

$$\begin{aligned}
\iint_S f\,dS &= \int_0^{2\pi}\int_0^{\pi} \cos^2\varphi\sin^2\theta \times \sin\theta\,d\theta\,d\varphi \\
&= \int_0^{2\pi} \cos^2\varphi\,d\varphi \int_0^{\pi} \sin^3\theta\,d\theta \\
&= \pi\int_0^{\pi} (1-\cos^2\theta)\sin\theta\,d\theta \\
&= \pi[-\cos\theta + \frac{1}{3}\cos^3\theta]_{\theta=0}^{\theta=\pi} \\
&= \pi(2-\frac{2}{3}) = \frac{4\pi}{3}.
\end{aligned}$$

Exemple 4.2.7 Calculons $\iint_S x\,dS$ où S est le morceau du plan $x+y+z=0$ pour $0 \le x \le a$ et $0 \le y \le b$ avec $a, b \ge 0$ des constantes fixées. Comme le plan est ici donné par son équation cartésienne, on en détermine une base comme suit :

$$\begin{aligned}
\{(x,y,z) \in \mathbb{R}^3 \ :\ & x+y+z=0,\ 0 \le x \le a,\ 0 \le y \le b\} \\
&= \{(x,y,-x-y),\ 0 \le x \le a,\ 0 \le y \le b\} \\
&= \{x(1,0,-1) + y(0,1,-1),\ 0 \le x \le a,\ 0 \le y \le b\}.
\end{aligned}$$

On en déduit qu'on peut prendre

$$X(u,v) = ue_1 + ve_2, \quad (u,v) \in \Omega = [0,a] \times [0,b],$$

avec $e_1 = (1, 0, -1)$ et $e_2 = (0, 1, -1)$. On a vu que dans ce cas,

$$dS = \left\| \frac{\partial X}{\partial u}(u, v) \wedge \frac{\partial X}{\partial v}(u, v) \right\| du\, dv = \|e_1 \wedge e_2\|\, du\, dv = \|(1, 1, 1)\|\, du\, dv = \sqrt{3}\, du\, dv.$$

De plus, la fonction à intégrer est

$$f(x, y, z) = x, \quad (x, y, z) \in \mathbb{R}^3,$$

et donc

$$f(X(u, v)) = f(u(1, 0, -1) + v(0, 1, -1)) = f(u, v, -u - v) = u.$$

Ainsi

$$\iint_S x\, dS = \int_0^b \int_0^a u\sqrt{3}\, du\, dv = \frac{\sqrt{3}}{2} a^2 b.$$

On remarque qu'ici on n'a pas paramétré le morceau de plan par $\Omega = [0, 1]^2$ mais plutôt par $\Omega = [0, a] \times [0, b]$: on aurait pu le faire ici en posant $x = a\tilde{x}$ et $y = b\tilde{y}$, et en remarquant que

$$\begin{aligned} \{x(1, 0, -1) + y(0, 1, -1),\ 0 \le x \le a,\ 0 \le y \le b\} \\ &= \{a\tilde{x}(1, 0, -1) + b\tilde{y}(0, 1, -1),\ 0 \le a\tilde{x} \le a,\ 0 \le b\tilde{y} \le b\} \\ &= \{\tilde{x}(a, 0, -a) + \tilde{y}(0, b, -b),\ 0 \le \tilde{x} \le 1,\ 0 \le \tilde{y} \le 1\}. \end{aligned}$$

On aurait pu alors prendre

$$X(u, v) = ue_1 + ve_2 = (au, bv, -au - bv), \quad (u, v) \in [0, 1]^2,$$

avec $e_1 = (a, 0, -a)$ et $e_2 = (b, 0, -b)$. Dans ce cas, on aurait eu

$$\frac{\partial X}{\partial u}(u, v) \wedge \frac{\partial X}{\partial v}(u, v) = e_1 \wedge e_2 = (ab, ab, ab),$$

et donc $dS = \sqrt{3}ab\, du\, dv$. On aurait aussi eu

$$f(X(u, v)) = f(au, bv, -au - bv) = au,$$

et donc

$$\iint_S f\, dS = \int_0^1 \int_0^1 au\sqrt{3}ab\, du\, dv = \frac{\sqrt{3}}{2} a^2 b.$$

On trouve bien le même résultat.

4.3 Surfaces orientées

On va maintenant définir une notion d'intégrale de champ de vecteurs sur une surface, similaire à la notion de circulation de champ de vecteurs le long d'une courbe. On a vu au chapitre précédent que cette notion dépendait de l'orientation de la courbe. Sur une surface, on va aussi définir une notion d'orientation. Vous le sentez venir : on aura besoin de ceci pour parler de flux d'un champ de vecteur à travers la surface, voir ci-dessous.

Définition 4.3.1: Vecteur normal à une surface en un point

Soit $S = (\Omega, X)$ une surface paramétrée et $M = X(u_0, v_0)$ un point de S avec $(u_0, v_0) \in \Omega$. On appelle *vecteur normal à S au point M* un vecteur $n \in \mathbb{R}^3$ satisfaisant les deux conditions :

1. n est orthogonal au plan tangent à S au point M ;
2. n est de norme 1.

GD : Vous avez deviné que si on change de paramétrage (en remplaçant par un paramétrage équivalent), on ne change pas la notion de vecteur normal à S en M. En chaque point de S (et je devais dire en chaque $(u, v) \in \Omega$), il n'y a que deux chois possibles de normale n : la première condition dit que n doit être dans un espace vectoriel de dimension 1, et la seconde ne laisse le choix qu'entre n et $-n$. On va utiliser le produit vectoriel pour faire un choix systématique de $n(X(u, v))$, qu'on appellera orientation de S.

Définition 4.3.2: Orientation d'une surface paramétrée

Soit $S = (\Omega, X)$ une surface paramétrée dont tous les points sont réguliers (donc $\frac{\partial X}{\partial u}(u, v) \wedge \frac{\partial X}{\partial v}(u, v) \neq 0$). Alors, l'application

$$n_X : \begin{cases} \Omega & \longrightarrow & \mathbb{R}^3 \\ (u, v) & \longmapsto & \dfrac{\frac{\partial X}{\partial u}(u, v) \wedge \frac{\partial X}{\partial v}(u, v)}{\left\| \frac{\partial X}{\partial u}(u, v) \wedge \frac{\partial X}{\partial v}(u, v) \right\|} \end{cases}$$

est définie et continue et vérifie que pour tout $(u, v) \in \Omega$, le vecteur $n_X(X(u, v))$ est un vecteur normal à S en $X(u, v)$. On dira que n_X définit l'orientation de S (associée au paramétrage (Ω, X)).

GD : C'est un peu bizarre, mais c'est dû au fait qu'on n'a vraiment défini que des surfaces paramétrées. Normalement, il y a une définition de surface (disons de classe C^1) de dimension 2 dans $\mathbb{R}^3$, qui est plus intrinsèque (c.à.d., pas directement donnée par un paramétrage). Grosso modo, pour chaque point M de S, on demande qu'il y ait un petit voisinage de M dans lequel S est le graphe (dans certaines coordonnées de $\mathbb{R}^n$, qui ont le droit de changer quand on change M) d'une fonction de classe C^1. Avec cette définition, on peut encore parler de plan tangent en M, puis de vecteur normal $n = n(M)$ au point M. Et une orientation de S, c'est juste le choix q'une fonction continue $n : S \to \mathbb{R}^3$ (en fait à valeurs dans la sphère unité), telle que pour tout $M \in S$, $n(M)$ est un vecteur normal à S au point M. Ici on a pris des surfaces paramétrées, on a choisi une fonction n associée, et on a dit que c'était celle-là qui définit l'orientation du support associée au paramétrage.

Prenons l'exemple de la sphère $S = \left\{ M \in \mathbb{R}^3 \,;\, ||M|| = 1 \right\}$, pour une fois vue comme un ensemble pas forcément paramétré. Pour tout $M \in S$, on a seulement deux chois possibles de $n(M)$: $n(M) = M$ ou $n(M) = -M$. En fait, comme n doit être une fonction continue de M

on peut vérifier que l'on doit prendre, soit $n(M) = M$ pour tout $M \in S$, soit $n(M) = -M$ pour tout $M \in S$. Voir plus bas pour un peu plus de détails. Cela donne deux orientations possibles à S, une avec la normale qui pointe vers l'extérieur de la boule, une avec la normale qui pointe vers l'intérieur.

Pour une surface S de classe C^1 plus générale, mais connexe (comprendre, en un seul morceau), il arrive la même chose : à cause de la continuité de n, on a seulement deux choix possibles de n, parce que quand on a choisi $n(M)$ pour un M, on n'a plus le choix du signe pour les voisins proche, et par propagation de l'information, on n'a plus le choix sur S tout entier. En fait, j'aurais du dire qu'il y a au plus deux orientations possibles, parce qu'il se pourrait a priori qu'on ne puisse pas trouver d'orientation du tout sur S. C'est ce qui se produit sur un ruban de Möbius (prendre une bandelette comme un long rectangle, et recoller les deux bords opposés éloignés, mais en faisant tourner le ruban d'un demi-tour). On peut vérifier que si on part avec $n(M)$ en un point, et qu'on suit ce que doit être n pour les voisins, et ainsi de suite, on se retrouve après un tour de l'autre coté de la bande, ce qui veut dire qu'on aurait aussi dû pendre $n(M) = -n(M)$, une contradiction. Mais par ailleurs, le ruban de Möbius, si je le prends fermé (si j'inclus les bord) n'est pas vraiment une surface de classe C^1 comme j'ai dit plus haut, parce que près du bord in ne ressemble pas à un graphe complet, juste à un demi-graphe. Et en fait un joli théorème dit que toute surface lisse (C^1) compacte (pour éviter Möbius) de dimension 2 dans $\mathbb{R}^3$ est effectivement orientable. Mais tout ceci était un peu une digression, et on a réglé le problème en ne regardant que des surfaces paramétrées.

Vérifications liées à la définition. Comme le plan tangent à S en $X(u,v)$ est par définition engendré par les vecteurs $\frac{\partial X}{\partial u}(u,v)$ et $\frac{\partial X}{\partial v}(u,v)$, et comme $\frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v)$ est orthogonal à ces deux vecteurs, on en déduit que le vecteur $\frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v)$ est orthogonal au plan tangent à S en $X(u,v)$. De plus, comme tous les points de S sont supposés réguliers, la famille $\{\frac{\partial X}{\partial u}(u,v), \frac{\partial X}{\partial v}(u,v)\}$ est libre en tout point et donc $\frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v) \neq 0$: l'application n_X est donc bien définie et continue. Enfin, le vecteur $n_X(X(u,v))$ est manifestement de norme 1 et orthogonal au plan tangent à S en tout point.

Maintenant, vérifions (presque) qu'il n'y a que si Ω est en un seul morceau (on dit connexe), il n'y a que deux choix possibles de fonctions continues $n : \Omega \mapsto \mathbb{R}^n$; telles qu'en tout point $(u,v) \in \Omega$, le vecteur $n(u,v)$ est un vecteur normal à S au point $X(u,v)$.

Supposons que n est un autre champ de vecteurs satisfaisant les même propriétés. Comme l'orthogonal au plan tangent de S au point $X(u,v)$ est de dimension 1, il est engendré par $n_X(X(u,v))$ et donc pour tout $(u,v) \in \Omega$, il existe $\lambda(u,v)$ un scalaire tel que

$$n(X(u,v)) = \lambda(u,v) n_X(X(u,v)).$$

En prenant la norme de cette relation, on trouve que

$$1 = \|n(X(u,v))\| = |\lambda(u,v)| \|n_X(X(u,v))\| = |\lambda(u,v)|,$$

et ainsi $\lambda(u,v) = \pm 1$ pour tout $(u,v) \in \Omega$. Enfin, comme on peut écrire

$$\lambda(u,v) = n_X(X(u,v)) \cdot n(X(u,v)),$$

on en déduit également que

$$(u,v) \in \Omega \mapsto \lambda(u,v) \in \{-1,1\}$$

est continue. Comme l'ensemble Ω est "en un seul morceau", on a deux possibilités : soit

$$\forall (u,v) \in \Omega,\ \lambda(u,v) = 1,$$

soit

$$\forall (u,v) \in \Omega,\ \lambda(u,v) = -1,$$

(penser au théorème des valeurs intermédiaires).

Exemple 4.3.3 Si S est une sphère centrée en l'origine et de rayon 1, on a vu qu'en tout point on a

$$\frac{\partial X}{\partial \theta}(\theta,\varphi) \wedge \frac{\partial X}{\partial \varphi}(\theta,\varphi) = (\cos\varphi \sin^2\theta, \sin\varphi \sin^2\theta, \cos\theta \sin\theta),$$

ainsi que

$$\left\| \frac{\partial X}{\partial \theta}(\theta,\varphi) \wedge \frac{\partial X}{\partial \varphi}(\theta,\varphi) \right\| = \sin\theta.$$

On en déduit que

$$n_X(\theta,\varphi) = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta) = X(\theta,\varphi).$$

Si l'on oriente S par n_X, on l'oriente donc par *normale sortante* et si on l'oriente par $-n_X$, on l'oriente par *normale entrante*

4.4 Flux d'un champ de vecteurs à travers une surface orientée

Une fois définie la notion de surface orientée, on définit une notion d'intégration de champ de vecteurs sur une surface orientée.

Définition 4.4.1

Soit $S = (\Omega, X)$ une surface paramétrée de $\mathbb{R}^3$ orientée par n_X, et soit $F : \mathbb{R}^3 \to \mathbb{R}^3$ un champ de vecteurs continu. Le *flux de F à travers S*, noté $\iint_S F \cdot \overrightarrow{dS}$, est

$$\iint_S F \cdot \overrightarrow{dS} = \iint_\Omega F(X(u,v)) \cdot \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v)\, du\, dv.$$

Si S est orientée par $-n_X$, alors

$$\iint_S F \cdot \overrightarrow{dS} = -\iint_\Omega F(X(u,v)) \cdot \frac{\partial X}{\partial u}(u,v) \wedge \frac{\partial X}{\partial v}(u,v)\, du\, dv.$$

Remarque 4.4.1 Si F désigne le champ de vitesse d'un fluide incompressible s'écoulant à travers une surface S, le flux de F à travers S s'interprète physiquement comme la quantité totale de fluide s'écoulant à travers S par unité de temps.

Remarque 4.4.2 On peut comparer la formule avec celle de la circulation d'un champ de vecteurs F le long d'une courbe orientée $\overset{\frown}{\gamma}$:

$$\int_{\overset{\frown}{\gamma}} F \cdot dx = \int_I F(\varphi(t)) \cdot \varphi'(t)\, dt,$$

qui elle aussi change de signe lorsqu'on change l'orientation de la courbe.

Remarque 4.4.3 Dans tous les cas, si S est orientée par un choix de normale n, on a

$$\iint_S F \cdot \overrightarrow{dS} = \iint_S F \cdot n\, dS,$$

où dS est l'élément d'aire sur S.

GD : Comme d'habitude, la notation suggère que les définitions ci-dessus s'étendent au cas de surfaces lisses orientées, sans qu'elles soient nécessairement paramétrées. Et dans notre cas de surfaces paramétrées, le flux ne change pas quand on remplace un paramétrage de S par un paramétrage équivalent. Par contre, le flux change de signe quand on change l'orientation de S, soit en remplaçant n par $-n$ sur tout S, soit en prenant un paramétrage avec une orientation différente (provenant d'un cha,gement de variable avec déterminant jacobien strictement négatif).

GD : je n'ai pas retrafiqué les notations ci-dessus ; il est assez difficile de répartir les flèches de vecteurs de manière totalement juste et transparente. On a gardé l'idée générale que F, avec une lettre capitale, est plutôt un vecteur, mais on aurait aussi pu lui donner une flèche. De même, il eut été logique de mettre une flèche à n mais la tradition est souvent de ne pas le faire.

Exemple 4.4.2 Calculons le flux de $F(x, y, z) = (x, y, z)$ à travers la sphère centrée en $(0, 0, 0)$ de rayon 1, orientée par normale sortante. Dans ce cas, on a vu que $\Omega = [0, \pi] \times [0, 2\pi[$ et que

$$X(\theta, \varphi) = (\cos\varphi \sin\theta, \sin\varphi \sin\theta, \cos\theta),\ (\theta, \varphi) \in \Omega.$$

On a également vu que

$$\frac{\partial X}{\partial \theta}(\theta, \varphi) \wedge \frac{\partial X}{\partial \varphi}(\theta, \varphi) = (\cos\varphi \sin^2\theta, \sin\varphi \sin^2\theta, \cos\theta \sin\theta),$$

et que ce paramétrage était compatible avec l'orientation choisie (puisque $n_X = X$ est bien sortante). Ainsi,

$$\begin{aligned}\iint_S F \cdot \overrightarrow{dS} &= \int_0^{2\pi} \int_0^{\pi} \begin{bmatrix} \cos\varphi \sin\theta \\ \sin\varphi \sin\theta \\ \cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\varphi \sin^2\theta \\ \sin\varphi \sin^2\theta \\ \cos\theta \sin\theta \end{bmatrix} d\theta\, d\varphi \\ &= \int_0^{2\pi} \int_0^{\pi} \sin\theta\, d\theta\, d\varphi = 4\pi.\end{aligned}$$

Un autre manière de justifier ce calcul consiste à remarquer que comme $n_X = X$ et que X paramètre la sphère on a $\|X\|^2 = 1$ donc

$$F(X(\theta,\varphi)) \cdot n_X(\theta,\varphi) = X(\theta,\varphi) \cdot X(\theta,\varphi) = \|X(\theta,\varphi)\|^2 = 1,$$

d'où

$$\iint_S F \cdot \overrightarrow{dS} = \iint_S F \cdot n \, dS = \iint_S 1 \, dS = \text{Aire}(S) = 4\pi.$$

Exemple 4.4.3 Calculons le flux de $F(x,y,z) = (x,z,x)$ à travers le cylindre $x^2 + y^2 = 1$ compris entre $z = 0$ et $z = 1$, orienté par normale sortante. On a vu que pour un cylindre, on pouvait prendre

$$X(\theta,z) = (\cos\theta, \sin\theta, z), \ (\theta,z) \in [0,2\pi[\times[0,1],$$

et que

$$\frac{\partial X}{\partial \theta}(\theta,z) \wedge \frac{\partial X}{\partial z}(\theta,z) = (\cos\theta, \sin\theta, 0).$$

Pour justifier que le paramétrage est compatible avec l'orientation, on calcule la normale en un point particulier : par exemple $(\theta,z) = (0,0)$, ce qui donne

$$\frac{\partial X}{\partial \theta}(0,0) \wedge \frac{\partial X}{\partial z}(0,0) = (1,0,0),$$

ce qui est clairement normal sortant au cylindre (faire un dessin). Si n_X est sortant en un point, elle l'est en tout point : le paramétrage est donc compatible avec l'orientation. On a de plus

$$F(X(\theta,z)) = (\cos\theta, z, \cos\theta),$$

et donc

$$\begin{aligned}\iint_S F \cdot \overrightarrow{dS} &= \int_0^1 \int_0^{2\pi} \begin{bmatrix}\cos\theta \\ z \\ \cos\theta\end{bmatrix} \cdot \begin{bmatrix}\cos\theta \\ \sin\theta \\ 0\end{bmatrix} d\theta \, dz \\ &= \int_0^1 \int_0^{2\pi} (\cos^2\theta + z\sin\theta) \, d\theta \, dz = \pi\end{aligned}$$

4.5 Formules de Stokes et de Green-Ostrogradski

[Dans la version précédente des notes, c'était Gauss-Ostrogradski ; je crois que Green a une majorité.] Enfin, on donne deux formules reliant le flux d'un champ de vecteurs à travers une surface orientée à deux autres types d'intégrales déjà vues : la circulation d'un champ de vecteurs le long d'une courbe (formule de Stokes) et l'intégrale triple d'une fonction scalaire sur un domaine de $\mathbb{R}^3$ (formule de Gauss-Ostrogradski). On commence par la formule de Stokes.

Le contexte de cette formule est le suivant : on considère une surface paramétrée S de $\mathbb{R}^3$ orientée par une normale n. On suppose que cette surface possède un bord, de la même manière qu'un disque dans $\mathbb{R}^2$ possède un bord qui est un cercle. On peut par exemple penser à une demi-sphère dans $\mathbb{R}^3$: elle possède un bord qui est aussi un cercle. Le bord de S est une courbe paramétrée $\widehat{\gamma} = (I, \varphi)$ orientée de manière compatible avec l'orientation de S. L'orientation de γ est choisie de la façon suivante. Soit $M = \varphi(t_0)$ un point de $\widehat{\gamma}$ (qui est aussi un point de S puisque $\widehat{\gamma}$ est le bord de S). En particulier, le vecteur tangent $t(M) = \varphi'(t_0)$ à $\widehat{\gamma}$ en M appartient à l'espace tangent à S en M. On peut alors choisir un autre vecteur tangent à S en M, qu'on appelle $v(M)$, de telle manière à ce que $v(M)$ pointe vers la surface S. Les orientations de $\widehat{\gamma}$ et de S sont dites compatibles si la famille $\{t(M), v(M), n(M)\}$ forme une base directe de $\mathbb{R}^3$ (c'est-à-dire si elle satisfait la règle de la main droite : ils doivent être orientée de telle manière à ce que $t(M)$ soit dans la direction de l'index, $v(M)$ dans la direction du majeur, et $n(M)$ dans la direction du pouce), ou autrement dit si $n(M) = t(M) \wedge v(M)$.

Théorème 4.5.1 (Formule de Stokes) *Soit S une surface paramétrée orientée de $\mathbb{R}^3$ bordée par une courbe paramétrée $\widehat{\gamma}$ orientée de manière compatible avec l'orientation de S comme expliqué ci-dessus. Alors, pour tout champ de vecteurs $F : \mathbb{R}^3 \to \mathbb{R}^3$ de classe C^1, on a*

$$\iint_S \operatorname{rot} F \cdot \overrightarrow{dS} = \oint_{\widehat{\gamma}} F \cdot \overrightarrow{dx}.$$

GD : je me suis permis d'ajouter une flèche au rotationnel, pour insister sur le fait que comme on est dans $\mathbb{R}^3$, il s'agit bien d'un champ de vecteur. Mais je le laisse sans flèche dans la suite, maintenant qu'on a compris. A la fin, les deux membres de l'identité sont des scalaires.

Exemple 4.5.1 Prenons $F(x, y, z) = (x, -y, 2)$ et calculons le flux de F à travers la demi-sphère S d'équations $x^2+y^2+z^2 = 1$, $z \geq 0$, orientée par la normale sortante. Pour appliquer la formule de Stokes, on veut écrire

$$\iint_S F \cdot \overrightarrow{dS} = \iint_S \operatorname{rot} G \cdot \overrightarrow{dS} = \oint_{\widehat{\gamma}} G \cdot \overrightarrow{dx},$$

et donc F doit être le rotationnel d'une certaine fonction G. On peut prendre $G(x, y, z) = (-y, x, xy)$, et constater que

$$\operatorname{rot} G(x, y, z) = \begin{bmatrix} \frac{\partial}{\partial x} \\ \frac{\partial}{\partial y} \\ \frac{\partial}{\partial z} \end{bmatrix} \wedge \begin{bmatrix} -y \\ x \\ xy \end{bmatrix} = \begin{bmatrix} x \\ -y \\ 2 \end{bmatrix} = F(x, y, z).$$

La courbe $\widehat{\gamma}$ est le bord de la demi-sphère S : c'est donc le cercle de centre $(0, 0, 0)$ de rayon 1 dans le plan $(0xy)$, qu'on peut paramétrer par $I = [0, 2\pi[$ et

$$\varphi(t) = (\cos(t), \sin(t), 0), \ t \in I.$$

Vérifions que les orientations de S et de $\widehat{\gamma}$ sont compatibles. En un point $M = \varphi(t)$ de $\widehat{\gamma}$, la tangente $t(M)$ à $\widehat{\gamma}$ est

$$t(M) = \varphi'(t) = (-\sin(t), \cos(t), 0),$$

et la normale $n(M)$ à S est

$$n(M) = M = \varphi(t) = (\cos(t), \sin(t), 0).$$

Comme la relation $n(M) = t(M) \wedge v(M)$ est équivalente à $v(M) = n(M) \wedge t(M)$, une manière de vérifier que $n(M) = t(M) \wedge v(M)$ est de calculer $n(M) \wedge t(M)$ est de vérifier que le vecteur obtenu pointe bien vers S. Ici, on a

$$n(M) \wedge t(M) = (0, 0, 1)$$

qui pointe bien vers S puisque S est située dans le demi-espace $z \geq 0$. Ainsi, les orientations de S et de $\widehat{\gamma}$ sont compatibles, et par la formule de Stokes on a

$$\begin{aligned}
\iint_S F \cdot \overrightarrow{dS} &= \iint_S \operatorname{rot} G \cdot \overrightarrow{dS} \\
&= \oint_{\widehat{\gamma}} G \cdot \overrightarrow{dx} = \int_0^{2\pi} G(\varphi(t)) \cdot \varphi'(t)\, dt \\
&= \int_0^{2\pi} \begin{bmatrix} -\sin(t) \\ \cos(t) \\ \cos(t)\sin(t) \end{bmatrix} \cdot \begin{bmatrix} -\sin(t) \\ \cos(t) \\ 0 \end{bmatrix} dt \\
&= \int_0^{2\pi} (\sin^2(t) + \cos^2(t))\, dt = 2\pi.
\end{aligned}$$

Remarque 4.5.1 L'énoncé de la formule de Stokes fait fortement penser à celui de la formule de Green-Riemann, qui dit que

$$\iint_\Omega \operatorname{rot} F\, dx\, dy = \oint_{\widehat{\gamma}} F \cdot \overrightarrow{dx},$$

la différence étant que le flux de $\operatorname{rot} F$ à travers S est remplacé par l'intégrale de $\operatorname{rot} F$ sur le domaine Ω. Pour Stokes, on intègre sur S qui est une surface dans $\mathbb{R}^3$ et pour Green-Riemann, on intègre sur un domaine de $\mathbb{R}^2$ (donc une surface "plate"). On va voir qu'on peut quand même retrouver la formule de Green-Riemann à partir de celle de Stokes, à l'aide du raisonnement suivant : considérons un domaine Ω de $\mathbb{R}^2$, qu'on va identifier à une surface de $\mathbb{R}^3$ par le paramétrage

$$X(x, y) = (x, y, 0), \quad (x, y) \in \Omega.$$

Il est facile de voir que

$$n_X(x, y) = \frac{\partial X}{\partial x}(x, y) \wedge \frac{\partial X}{\partial y}(x, y) = (1, 0, 0) \wedge (0, 1, 0) = (0, 0, 1),$$

et on choisit d'orienter S par n_X. Si $F : \mathbb{R}^2 \to \mathbb{R}^2$ est de classe C^1, on l'étend à un champ de vecteurs sur $\mathbb{R}^3$ de la façon suivante, en définissant

$$\tilde{F}(x, y, z) = (F_1(x, y), F_2(x, y), 0), \quad (x, y, z) \in \mathbb{R}^3.$$

On a alors

$$\operatorname{rot} \tilde{F}(x, y, z) = (0, 0, \operatorname{rot} F(x, y)), \quad (x, y, z) \in \mathbb{R}^3.$$

De plus, si le bord de Ω est $\overset{\frown}{\gamma} = (I, \varphi)$ est orienté de telle manière à ce que Ω reste à gauche lorsqu'on parcours $\overset{\frown}{\gamma}$, alors le bord de S est $\overset{\frown}{\gamma_S} = (I, \varphi_S)$ avec

$$\varphi_S(t) = (\varphi_1(t), \varphi_2(t), 0) \in \mathbb{R}^3, \quad t \in I.$$

En un point $M = \varphi_S(t)$ du bord de S, on a donc

$$n_X(M) \wedge t(M) = (0, 0, 1) \wedge (\varphi_1'(t), \varphi_2'(t), 0) = (-\varphi_2'(t), \varphi_1'(t), 0),$$

qu'on identifie comme la rotation d'angle $\pi/2$ du vecteur $\varphi_S'(t)$ dans le plan (Oxy). Comme Ω reste à gauche lorsqu'on parcourt $\overset{\frown}{\gamma}$, on déduit que le vecteur $n_X(M) \wedge t(M)$ pointe vers la surface S (faire un dessin), et donc que les orientations de S et de $\overset{\frown}{\gamma_S}$ sont compatibles. En appliquant la formule de Stokes au champ de vecteurs $\tilde{F}$ sur la surface S, on obtient

$$\iint_S \operatorname{rot} \tilde{F} \cdot \overrightarrow{dS} = \oint_{\overset{\frown}{\gamma_S}} \tilde{F} \cdot \overrightarrow{dx}.$$

Identifions à présent les deux termes de cette égalité. A gauche, on a

$$\begin{aligned} \iint_S \operatorname{rot} \tilde{F} \cdot \overrightarrow{dS} &= \iint_\Omega \operatorname{rot} \tilde{F}(X(x, y)) \cdot \frac{\partial X}{\partial x}(x, y) \wedge \frac{\partial X}{\partial y}(x, y)\, dx\, dy \\ &= \iint_\Omega \begin{bmatrix} 0 \\ 0 \\ \operatorname{rot} F(x, y) \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} dt = \iint_\Omega \operatorname{rot} F(x, y)\, dx\, dy. \end{aligned}$$

A droite, on a

$$\begin{aligned} \oint_{\overset{\frown}{\gamma_S}} F \cdot \overrightarrow{dx} &= \int_I \tilde{F}(\varphi_S(t)) \cdot \varphi_S(t)\, dt \\ &= \int_I \begin{bmatrix} F_1(\varphi_1(t), \varphi_2(t)) \\ F_2(\varphi_1(t), \varphi_2(t)) \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \varphi_1'(t) \\ \varphi_2'(t) \\ 0 \end{bmatrix} dt \\ &= \int_I (F_1(\varphi_1(t), \varphi_2(t))\varphi_1'(t) + F_2(\varphi_1(t), \varphi_2(t))\varphi_2'(t))\, dt \\ &= \int_I F(\varphi(t)) \cdot \varphi'(t)\, dt = \oint_{\overset{\frown}{\gamma}} F \cdot \overrightarrow{dx}. \end{aligned}$$

Ainsi, on retrouve bien que

$$\iint_\Omega \operatorname{rot} F(x,y)\,dx\,dy = \oint_{\widehat{\gamma}} F\cdot\overrightarrow{dx},$$

qui est la formule de Green-Riemann. On vient donc de montrer que la formule de Green-Riemann pouvait se déduire de la formule de Stokes.

Passons à présent à la formule de Green-Ostrogradski (ou peut-être Gauss). On considère un domaine Ω de $\mathbb{R}^3$ qui est bordé par une surface S (penser par exemple à une boule qui est bordée par une sphère). On dit alors que S est orientée par la normale sortante si la normale ne pointe pas vers Ω. La formule de Green-Ostrogradski relie une intégrale triple sur Ω avec un flux sur S.

Théorème 4.5.2 (Formule de Green-Ostrogradski) *Soit Ω un domaine de $\mathbb{R}^3$ dont le bord est une surface paramétrée S de $\mathbb{R}^3$ orientée par normale sortante de Ω. Soit $F:\mathbb{R}^3\to\mathbb{R}^3$ un champ de vecteurs C^1. Alors, on a*

$$\iiint_\Omega \operatorname{div} F\,dx\,dy\,dz = \iint_S F\cdot\overrightarrow{dS}.$$

Exemple 4.5.2 Si F est un champ constant, alors $\operatorname{div} F = 0$ et donc $\iint_S F\cdot\overrightarrow{dS} = 0$: le flux d'un champ constant à travers une surface "fermée" (c'est-à-dire qui borde un domaine de $\mathbb{R}^3$) est toujours nul.

En fait, si on pense que F est le champ de vecteur des vitesses en chaque point d'un liquide, le flux de F est l'opposé la dérivée par rapport au temps de la quantité de liquide qui se trouve dans Ω, ou si vous préférez, la différence entre ce qui sort et ce qui rentre par unité de temps. La condition "$\operatorname{div} F = 0$" est une manière de coder que le liquide est "incompessible" (pensez à de l'eau), et dans ce cas le théorème du flux (Green-Ostrogradski) dit que le flux est nul (il en rentre autant qu'il en sort). Ceci ne doit pas vous surprendre (et en fait c'est parce qu'on veut que le flux soit nul que "$\operatorname{div} F = 0$" est une bonne manière de coder l'incompressibilité).

Exemple 4.5.3 Si $F(x,y,z) = (x,y,z)$, alors $\operatorname{div} F = 3$, et donc

$$\iiint_\Omega \operatorname{div} F\,dx\,dy\,dz = 3\,\text{Volume}(\Omega).$$

On a donc

$$\text{Volume}(\Omega) = \frac{1}{3}\iint_S F\cdot\overrightarrow{dS}.$$

Dans le cas de la boule unité, ceci donne un volume de $\frac{4\pi}{3}$, puisque $F\cdot n = 1$.

Notez qu'on a pris la formule avec $F(x,y,z) = (x,y,z)$, mais c'est par un simple souci d'invariance par rotation. on pouvait prendre aussi $F(x,y,z) = (x,0,0)$, et alors on obtenait $\text{Volume}(\Omega) = \iint_S F\cdot\overrightarrow{dS} = \iint_S x n_1 dS$, où x est la première coordonnée d'un point de S et n_1 la première coordonnée de la normale unitaire en ce point.

Remarque 4.5.2 On fera attention à ce que la formule de Stokes s'applique à des surfaces S avec un bord, alors que la formule de Green-Ostrogradski s'applique à une surface sans bord.

Pour conclure, on remarque que les formules

$$\begin{cases} \displaystyle\int_a^b f(t)\,dt = f(b) - f(a), \quad f : \mathbb{R} \to \mathbb{R}, \\ \displaystyle\iint_S \operatorname{rot} F \cdot \overrightarrow{dS} = \oint_{\widehat{\gamma}} F \cdot \overrightarrow{dx}, \quad F : \mathbb{R}^3 \to \mathbb{R}^3, \\ \displaystyle\iiint_\Omega \operatorname{div} F\,dx\,dy\,dz = \iint_S F \cdot \overrightarrow{dS}, \quad F : \mathbb{R}^3 \to \mathbb{R}^3 \end{cases}$$

possèdent des structures similaires : dans le terme de gauche, on a l'intégrale d'une quantité s'exprimant à l'aide de dérivées d'une fonction sur un domaine, et à droite on a une quantité s'exprimant à l'aide des valeurs de cette fonction au bord de ce domaine. C'est pourquoi on peut interpréter les formules de Stokes et de Green-Ostrogradski comme des généralisations de la formule $\int_a^b f'(t)\,dt = f(b) - f(a)$.

GD : Encore quelques commentaires sur le théorème 4.5.2, qui est encore malheureusement livré sans démonstration.

D'abord, on a utilisé les produits vectoriels pour vous permettre de calculer $\iint_S F \cdot \overrightarrow{dS}$ plus facilement. Mais en fait, sous la forme

$$\iiint_\Omega \operatorname{div} F\,dx\,dy\,dz = \iint_S (F \cdot n)\,dS,$$

où n est encore la normale sortante à Ω, la formule est vraie dans $\mathbb{R}^n$ en toutes dimensions. Par contre, je ne vous ai donné ici aucun moyen concret de calculer les intégrales sur des hypersurfaces S dans $\mathbb{R}^n$, où $n \geq 4$.

Pour des domaines assez simples (et contrairement à la fomule de Stokes qui est plus difficile à comprendre), le théorème 4.5.2 n'est pas si compliqué que ça. On note d'abord qu'il suffit de démontrer lorsque F a une seule composante, donc par exemple $F = (0, 0, f)$ pour une fonction (scalaire) f. Alors il prend la forme suivante :

$$\iiint_\Omega \frac{\partial f}{\partial z}\,dx\,dy\,dz = \iint_S f(w) n_3(w)\,dS(w), \tag{4.1}$$

où n_3 est la composante verticale de n. Par exemple, pour le domaine Ω un peu plus simple compris entre deux graphes, disons de g_1 et $g_2 \geq g_1$ définies sur le même domaine V, on peut procéder comme suit. Par Fubini on a

$$\begin{aligned} \iiint_\Omega \frac{\partial f}{\partial z}(x,y,z)\,dx\,dy\,dz &= \iint_V \left\{ \int_{g_1(x,y)}^{g_2(x,y)} \frac{\partial f}{\partial z}(x,y,z)\,dz \right\} dx\,dy \\ &= \iint_V \big[f(x,y,g_2(x,y)) - f(x,y,g_1(x,y)) \big]\,dx\,dy \end{aligned} \tag{4.2}$$

alors que pour $\iint_S f(w)n_3(w)\,dS(w)$, la partie verticale de la surface ne contribue pas (puisque n_3 y est nulle), et qu'il reste $\iint_S f(w)n_3(w)\,dS(w) = I_1 + I_2$, où pour $j = 1, 2$, I_j coresponde à la partie S_j de S qui est dans le graphe de g_j. Ensuite, pour J_2,

$$J_2 = \iint_{S_j} f(w)n_3(w)\,dS(w) = \iint_V f(x,y,g_2(x,y))n_3(x,y,g_2(x,y))\sqrt{1+\frac{\partial g_2}{\partial x}^2+\frac{\partial g_2}{\partial y}^2}\,dx\,dy. \tag{4.3}$$

où pour gagner de la place je n'ai pas écrit que les dérivées partielles sont calculées en (x, y). Et maintenant, miracle, notre vecteur normal initial était $N(x, y) = (-\frac{\partial g_2}{\partial x}, -\frac{\partial g_2}{\partial y}, 1)$, donc $n_3(x, y, g_2(x, y)) = N_3(x, y)/||N(x, y)|| = 1/||N(x, y)||$, qui est l'inverse de la racine qu'on trouve dans l'intégrale de (4.3). Du coup $J_2 = \iint_V f(x, y, g_2(x, y))dxdy$.

Et pour J_1, on a le même calcul, sauf que maintenant la normale sortante est orioentée vers le bas, donc en fait est de la forme $-N_3(x, y)/||N(x, y)|| = -1/||N(x, y)||$, ce qui donne le bon signe. On ajoute et on trouve comme en (4.2).

Encore une remarque : quand $\Omega = \Omega_1 \cup \Omega_2$ avec deux gentils domaines disjoints mais qui ont une partie de frontière commune, la formule pour Ω se déduit de celles pour Ω_1 et Ω_2, parce que $\iiint_\Omega \operatorname{div} F\,dx\,dy\,dz = \iiint_{\Omega_1} \operatorname{div} F\,dx\,dy\,dz + \iiint_{\Omega_2} \operatorname{div} F\,dx\,dy\,dz$, et que quand on somme les deux flux, la partie commune des deux frontières donne deux contribution opposées, parce que F est le même, mais les normales extérieures sont opposées.

C'est bien heureux que notre formule ait cette propriété, sinon on aurait pu s'inquiéter, mais ça aide aussi à démontrer le théorème en se ramenant à des cas plus simples, comme celui ci-dessus.

Exemple 4.5.4 Le théorème 4.5.2 est à la base de plein d'autres exemples de formules dites d'intégration par parties. Ainsi, pour un domaine Ω suffisament simple dans $\mathbb{R}^n$, on a

$$\iiint_\Omega \Delta f\,dx\,dy\,dz = \iint_S \nabla f \cdot \overrightarrow{dS} = \iint_S \nabla f \cdot n\,dS$$

pour tout f à valeurs scalaires, et de classe C^2 sur un voisinage de Ω.

Appliquer la formule avec $F = \nabla f$ en notant que $\operatorname{div} \nabla f = \Delta f$. La formule où on prend $g = f^2$ est assez utile aussi.