

Module M31

Analyse II

Fonctions de Plusieurs Variables

C.-A. PILLET

La dernière version de ce document est disponible à l'adresse suivante :

<http://math.univ-tln.fr/~pillet/data/pdf/m31.pdf>

Conventions typographiques Le symbole \square marque la fin d'une démonstration (QED) alors que \triangleleft marque la fin d'un exemple. Les hyperliens vers le web sont en violet, comme celui ci-dessus. Les liens internes à ce documents sont en rouge, comme les têtes de paragraphe dans la table des matières qui suit.

Table des matières

1 Fonctions de plusieurs variables	7
1.1 Topologie de \mathbb{R}^n	7
1.1.1 Points et vecteurs, produit scalaire, norme et distance	7
1.1.2 Boules, voisinages, ensembles ouverts, ensembles fermés	9
1.1.3 Adhérence, intérieur, bord	12
1.1.4 Suites dans \mathbb{R}^n	14
1.1.5 Continuité	16
1.1.5.1 Limites de fonctions	17
1.1.5.2 Fonctions continues sur \mathbb{R}^n	18
1.1.5.3 Continuité uniforme	20
1.2 Fonctions différentiables	20
1.2.1 Courbes différentiables	20
1.2.2 Champs scalaires	22
1.2.2.1 Dérivées partielles	26
1.2.2.2 Dérivées directionnelles	29
1.2.2.3 Différentielle	30
1.2.2.4 Plan tangent	31
Approximation linéaire.	34
1.2.2.5 Règles de différentiation	36
1.2.2.6 Fonctions de classe C^n	36
1.2.2.7 Formule de Taylor	39
1.2.2.8 Extrema	43
Formes quadratiques dans \mathbb{R}^2 .	44
1.2.2.9 Le gradient	47
1.2.3 Champs vectoriels	50
1.2.3.1 Changements de coordonnées	54
1.2.3.2 Dérivées partielles d'un champ vectoriel	56
1.2.3.3 Champs vectoriels différentiables	57
1.2.3.4 Laplacien et fonctions harmoniques	59
1.2.4 Transformations	59
1.2.4.1 Jacobien d'une transformation	60

1.2.4.2	Fonctions composées, changement de variables	63
	Intermédiaire : le calcul des opérateurs différentiels.	65
1.2.4.3	Le théorème de la fonction inverse	71
1.2.4.4	Le théorème de la fonction implicite	73
2	Intégrales multiples	75
2.1	Sommes et intégrales de Riemann	75
2.2	Intégrales itérées	78
2.3	Intégrales doubles	79
2.3.1	Intégrales sur un rectangle	79
2.3.2	Intégrales sur un domaine simple	81
2.3.3	Changement de variables	85
2.3.4	Calcul d'aires	91
2.4	Intégrales triples	91
2.4.1	Intégrales sur des quadrilatères	91
2.4.2	Intégrales sur des domaines simples	92
2.4.3	Changement de variables	97
2.4.4	Calcul de volumes	100
3	Equations différentielles	102
3.1	Equations linéaires	104
3.1.1	Le problème homogène	105
3.1.2	Le problème inhomogène	110
3.1.3	Le cas d'ordre 1	113
3.1.4	Le cas d'ordre 2	114
3.1.5	Equations à coefficients constants	119
3.2	Equations non linéaires	128
3.2.1	Equations séparables	128
3.2.2	Deux théorèmes généraux	135
3.2.3	Analyse qualitative	138
3.2.4	Intégrales premières	144
3.2.4.1	EDO exacte	146
3.2.4.2	Facteur intégrant	149
3.2.5	Chaos	152

Introduction

L'objectif de ce cours d'analyse est de vous familiariser avec les bases du calcul différentiel et intégral des fonctions de plusieurs variables. Il forme un tout indissociable avec le cours d'analyse du premier semestre (M11). Ceci signifie en particulier que pour suivre efficacement ce cours vous devez avoir compris le contenu du M11 dont nous ferons un usage fréquent. Je vous invite donc à ressortir vos notes du premier semestre et à les consulter chaque fois que vous rencontrerez un problème. Certaines parties de ce cours, notamment dans les chapitres 1 et 3, feront également usage de notions de base d'algèbre linéaire que vous devez acquérir dans un cours parallèle (M21 pour les matheux, M24 pour les MASS). Là encore, n'hésitez pas à revoir vos notes au moment voulu. Réciproquement, les notions acquises dans ce cours vous seront nécessaires, dans la suite de vos études, pour aborder sereinement des cours comme les M31 M32 M52 M61 M62 M63 Q41 Q51 ... (pour les matheux) ou les M36 M48 M54 ... (pour les MASS).

Avant de commencer, voici quelques conseils qui pourraient vous aider dans votre apprentissage des Mathématiques.

1. Vous devez dès maintenant prendre conscience du fait que la segmentation des enseignements de mathématiques en modules MXY est artificielle et ne correspond à aucune réalité. Les Mathématiques, en tout cas au niveau de la Licence, forment un tout dont vous devrez parvenir, au terme de 3 années d'étude, à avoir une vision globale. Pour vous donner un exemple de cette intrication, des concepts d'algèbre linéaire (M21/M24) vont jouer un rôle central dans ce cours d'analyse (M22). En particulier les notions d'indépendance linéaire et d'application linéaire. Ce sont des notions intimement liées à la structure d'espace vectoriel. Un tel espace est un groupe commutatif pour l'addition. La structure de ces groupes est étudiée en détail dans un cours d'algèbre (M53). Une bonne compréhension des espaces vectoriels sera pour vous un atout important pour comprendre cette structure. D'ailleurs vous retrouverez ces mêmes groupes lorsque vous étudierez les systèmes différentiels (M31).
2. Ne cherchez pas à apprendre des recettes pour résoudre des problèmes types. Si cette méthode a peut-être fonctionné pour vous dans vos études secondaires, elle est vouée à l'échec à l'université. Vos cours n'y sont pas conçus comme des livres de recettes mais comme un apprentissage du métier de mathématicien. Pour apprendre ce métier vous devez chercher à comprendre non seulement chaque étape d'une démonstration, chaque pas d'un calcul, ce qui est généralement assez facile, mais aussi la stratégie sur laquelle repose cette démonstration où ce calcul. Cet exercice est plus difficile et demande du temps et de la réflexion. Il nécessite d'avoir non seulement une vision locale, pas à pas, de l'argument, mais aussi de la manière avec laquelle ces pas successifs s'enchaînent pour arriver à la conclusion à partir des hypothèses. Ce n'est qu'en pratiquant régulièrement cet exercice que vous apprendrez à construire et à formuler vos propres démonstrations et à mener à bien des calculs de plus en plus difficiles. Voici un petit exemple pour illustrer mon propos. La plupart d'entre vous se souviennent certainement de la recette pour résoudre l'équation du second degré

$$ax^2 + bx + c = 0.$$

Cette recette est tellement utile qu'il est effectivement approprié de la connaître par cœur. Mais vous souvenez-vous de comment on l'obtient? La stratégie consiste à compléter le carré. On l'appelle parfois "le truc des babyloniens" (bien que ces derniers usaient apparemment de "trucs" plus sophistiqués, regardez [ici](#) si cette histoire vous intéresse). On écrit

$$0 = ax^2 + bx + c = a \left(x^2 + 2 \frac{b}{2a} x \right) + c = a \left[\left(x + \frac{b}{2a} \right)^2 - \left(\frac{b}{2a} \right)^2 \right] + c,$$

pour obtenir

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2},$$

et par conséquence

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (1)$$

Cette stratégie peut être très utile dans d'autres problèmes. Par exemple dans le calcul de l'intégrale gaussienne (l'exemple 2.8 de ce cours) ou pour étudier les formes quadratiques en deux variables (voir le paragraphe *formes quadratiques*, page 44). On doit donc admettre que le truc des babyloniens est plus important que la recette (1) et qu'en tout état de cause on doit le connaître et savoir le mettre en œuvre de manière appropriée.

3. Bien entendu, les TD sont conçus pour vous aider dans cette démarche. Cependant ils ne seront efficaces que si vous faites un réel effort personnel pour résoudre les problèmes par vous même. Attendre patiemment que votre professeur vous en donne la solution ne vous apportera rien ! Demandez à vos enseignants de vous fournir les fiches de TD à l'avance pour pouvoir y travailler chez vous.
4. Faites des dessins ! Cherchez à visualiser des objets qui vous semblent abstraits. Un petit dessin est souvent une source d'inspiration. J'ai fait un effort particulier en rédigeant ces notes pour les accompagner de nombreuses illustrations. J'espère qu'elles vous aideront à comprendre des concepts parfois abstraits mais surtout qu'elles vous donneront des idées sur la façon d'utiliser des dessins dans votre travail personnel.
5. Lisez régulièrement votre cours. Avant chaque leçon, révisez les leçons précédentes. Si vous ne les avez pas assimilées, vous perdrez votre temps car vous serez incapable de comprendre ce qui suit et qui est construit sur ce qui précède. Relisez votre cours et si vous ne comprenez toujours pas, posez des questions. Vos professeurs seront ravis d'y répondre.
6. Enfin, vous devez concevoir la pratique des Mathématiques comme une activité créatrice, et surtout vous devez y trouver du plaisir !

Le cours est organisé de la façon suivante.

Dans le chapitre 1 nous aborderons le thème central du cours, les fonctions de plusieurs variables. Après une brève introduction nécessaire à la topologie élémentaire des espace euclidiens \mathbb{R}^n et à l'étude des fonctions de \mathbb{R} dans \mathbb{R}^n , nous définirons les notions de dérivées partielles et directionnelles d'une fonction de \mathbb{R}^n dans \mathbb{R} . Le concept de fonction différentiable sera alors introduit et nous en étudierons en détails les propriétés les plus élémentaires et les applications les plus importantes. Nous élargirons finalement le cadre de la discussion aux fonctions de \mathbb{R}^n dans \mathbb{R}^m .

Le chapitre 2 sera dédié au calcul intégral en plusieurs variables. Il sera orienté sur les aspects géométriques (calcul d'aires, de volumes) et calculatoires (réduction des intégrales multiples à des intégrales itérées) de "l'intégration". Les étudiants poursuivant leur cursus en Licence de Mathématiques retrouverons ce thème en 3ème année dans un cours qui lui sera exclusivement dédié et dans lequel il aborderons des aspects plus théoriques.

Finalement une courte introduction aux équations différentielles ordinaires sera l'objet du chapitre 3. Cette introduction se limitera aux équations scalaires. Les systèmes d'équations différentielles seront traités dans un autre cours, le M31. Nous traiterons en détails les équations linéaires du premier ordre ainsi que les équations linéaires à coefficients constants. Nous introduirons des méthodes importantes comme celle de la variation des constantes. Nous terminerons par l'étude des équations différentielles non linéaires du premier ordre.

Chapitre 1

Fonctions de plusieurs variables

1.1 Topologie de \mathbb{R}^n

Pour pouvoir définir la notion de limite d'une fonction de plusieurs variables, sa continuité ou sa différentiabilité nous avons besoin de quelques éléments de topologie. Nous nous contenterons d'abord de quelques notions très simples et absolument nécessaires pour la suite. Ceux d'entre vous qui poursuivront leurs cursus en Licence de Mathématiques retrouveront la topologie en 3^{ème} année dans un cours qui lui sera entièrement consacré (et qui, lui, ne sera pas aussi simple!).

Le mot topologie est formé des racines grecques *topos* → *lieu* et *logos* → *étude*. La topologie (on l'appelait au XIX^{ème} siècle *analysis situs* → *l'étude des lieux*) est un vaste domaine des mathématiques qui touche une variété d'autres sujets comme la géométrie, l'analyse ou l'algèbre. En ce qui nous concerne, elle étudie les ensembles de points et leurs propriétés invariantes par une classe de transformations dites continues.

1.1.1 Points et vecteurs, produit scalaire, norme et distance

Ce paragraphe est un bref rappel de notions élémentaires. Son but principal est de fixer les notations que nous utiliserons dans ce cours. Référez-vous à votre cours d'algèbre linéaire pour plus de détails.

Les points de l'espace affine euclidien \mathbb{R}^n seront notés

$$x = (x_1, x_2, \dots, x_n), \quad y = (y_1, y_2, \dots, y_n), \quad u = (u_1, u_2, \dots, u_n), \dots \quad (1.1)$$

Lorsque $n = 2$ ou $n = 3$ on utilisera aussi les notations

$$P = (x, y), \quad Q = (x, y, z), \quad x = (a, b), \quad x = (a, b, c), \quad y = (u, v), \dots$$

Les vecteurs de l'espace vectoriel euclidien \mathbb{R}^n seront notés en colonnes

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad u = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix},$$

Cependant, pour des raisons purement typographiques, nous utiliserons également la notation en ligne (1.1).

Les principales opérations sur les points et les vecteurs sont les suivantes :

1. La somme des deux vecteurs x et y est le vecteur

$$x + y = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}.$$

L'élément neutre pour l'addition des vecteurs est le vecteur nul qu'on dénote par 0 et dont toutes les composantes sont nulles.

2. Le produit du vecteur x par le scalaire $\alpha \in \mathbb{R}$ est le vecteur

$$\alpha x = \alpha \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}.$$

3. Le translaté du point $P = (x_1, x_2, \dots, x_n)$ par le vecteur $h = (h_1, h_2, \dots, h_n)$ est le point

$$P + h = (x_1 + h_1, x_2 + h_2, \dots, x_n + h_n).$$

4. Si P et Q sont deux points, \overrightarrow{PQ} est le vecteur avec lequel il faut translater P pour arriver à Q , c'est-à-dire $Q = P + \overrightarrow{PQ}$.

5. Le **produit scalaire** des deux vecteurs $x = (x_1, x_2, \dots, x_n)$ et $y = (y_1, y_2, \dots, y_n)$ est le nombre réel

$$x \cdot y = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

6. La **norme** du vecteur x est

$$\|x\| = \sqrt{x \cdot x} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}.$$

7. La **distance** entre les deux points $P = (x_1, x_2, \dots, x_n)$ et $Q = (y_1, y_2, \dots, y_n)$ est

$$d(P, Q) = \|\overrightarrow{PQ}\| = \sqrt{(y_1 - x_1)^2 + (y_2 - x_2)^2 + \dots + (y_n - x_n)^2}.$$

On dit que la fonction $d: \mathbb{R}^n \times \mathbb{R}^n \rightarrow [0, \infty[$ est la **métrique** de l'espace \mathbb{R}^n .

Les propriétés importantes de ces opérations sont résumées dans le résultat suivant.

Théorème 1.1 *Le produit scalaire est*

(i) *symétrique* : $x \cdot y = y \cdot x$,

(ii) *(bi)linéaire* : $(\alpha x + \beta y) \cdot z = \alpha x \cdot z + \beta y \cdot z$,

(iii) *défini positif* : $x \cdot x \geq 0$, et $x \cdot x = 0$ si et seulement si $x = 0$.

La norme satisfait les propriétés

(i) *de positivité* : $\|x\| \geq 0$ et $\|x\| = 0$ si et seulement si $x = 0$,

(ii) *d'homogénéité* : $\|\alpha x\| = |\alpha| \|x\|$,

(iii) *ainsi que l'inégalité du triangle* : $\|x + y\| \leq \|x\| + \|y\|$.

La métrique est

(i) *symétrique* : $d(P, Q) = d(Q, P)$,

(ii) *positive* : $d(P, Q) \geq 0$ et $d(P, Q) = 0$ si et seulement si $P = Q$,

(iii) *et satisfait l'inégalité du triangle* : $d(P, Q) \leq d(P, R) + d(R, Q)$.

Toutes ces propriétés sont évidentes, sauf peut-être l'inégalité du triangle satisfaite par la norme et la métrique. On remarque tout d'abord que l'inégalité pour la métrique suit de celle pour la norme et de la relation de Chasles $\overrightarrow{PQ} = \overrightarrow{PR} + \overrightarrow{RQ}$,

$$d(P, Q) = \|\overrightarrow{PQ}\| = \|\overrightarrow{PR} + \overrightarrow{RQ}\| \leq \|\overrightarrow{PR}\| + \|\overrightarrow{RQ}\| = d(P, R) + d(R, Q)$$

L'inégalité du triangle pour la norme est une conséquence de l'**inégalité de Cauchy-Schwarz**

$$|x \cdot y| \leq \|x\| \|y\|.$$

En effet, la linéarité permet d'écrire

$$\|x + y\|^2 = (x + y) \cdot (x + y) = x \cdot (x + y) + y \cdot (x + y),$$

et à l'aide de la symétrie nous pouvons convertir cette identité en

$$\|x + y\|^2 = (x + y) \cdot x + (x + y) \cdot y.$$

En utilisant encore une fois la linéarité et la symétrie on obtient

$$\|x + y\|^2 = x \cdot x + y \cdot x + x \cdot y + y \cdot y = x \cdot x + 2x \cdot y + y \cdot y = \|x\|^2 + 2x \cdot y + \|y\|^2.$$

L'inégalité de Cauchy-Schwarz nous donne finalement

$$\|x + y\|^2 \leq \|x\|^2 + 2\|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2,$$

et on obtient l'inégalité du triangle en prenant la racine carrée de chaque côté de cette inégalité. L'inégalité de Cauchy-Schwarz quant à elle suit du fait que

$$(x + \alpha y) \cdot (x + \alpha y) = \|x + \alpha y\|^2 \geq 0,$$

pour tout $\alpha \in \mathbb{R}$. En utilisant les propriétés de symétrie et de linéarité du produit scalaire on peut en effet en déduire que

$$f(\alpha) = \|x\|^2 + 2\alpha x \cdot y + \alpha^2 \|y\|^2 \geq 0,$$

pour tout $\alpha \in \mathbb{R}$. La fonction $f(\alpha)$ est un polynôme du second degré non négatif. Elle admet donc au plus une racine ce qui implique que son discriminant

$$(x \cdot y)^2 - \|x\|^2 \|y\|^2,$$

est négatif ou nul, c'est-à-dire que

$$(x \cdot y)^2 \leq \|x\|^2 \|y\|^2.$$

L'inégalité de Cauchy-Schwarz suit immédiatement de cette inégalité en prenant la racine carrée de ses deux membres.

1.1.2 Boules, voisinages, ensembles ouverts, ensembles fermés

Pour $\delta > 0$, $U_\delta(x_0)$ désigne la **boule** de rayon δ centrée en $x_0 \in \mathbb{R}^n$,

$$U_\delta(x_0) = \{x \in \mathbb{R}^n \mid d(x, x_0) < \delta\}.$$

Notez bien l'inégalité stricte, $d(x, x_0) < \delta$, dans cette définition, elle est très importante. $U_\delta(x_0)$ est donc l'ensemble des points de \mathbb{R}^n dont la distance à x_0 est *strictement* inférieure à δ .

Un ensemble $V \subset \mathbb{R}^n$ est un **voisinage** du point $x_0 \in \mathbb{R}^n$ s'il existe une boule $U_\delta(x_0)$ entièrement contenue dans V , c'est-à-dire s'il existe un réel $\delta > 0$ tel que, pour tout $x \in \mathbb{R}^n$, $d(x, x_0) < \delta$ implique $x \in V$.

Un ensemble est **ouvert** si c'est un voisinage de chacun de ses points. Autrement dit, $O \subset \mathbb{R}^n$ est ouvert si pour tout $x_0 \in O$ il existe une boule $U_\delta(x_0)$ telle que $U_\delta(x_0) \subset O$.

Un ensemble $F \subset \mathbb{R}^n$ est **fermé** si son complément $F^c = \mathbb{R}^n \setminus F$ est ouvert. F est donc fermé si, pour tout $x_0 \notin F$ il existe un réel $\delta > 0$ tel que, pour tout $x \in \mathbb{R}^n$, $d(x, x_0) < \delta$ implique que $x \notin F$.

Exemple 1.1 Si $n = 1$, $U_\delta(x_0) =]x_0 - \delta, x_0 + \delta[$ est un intervalle ouvert. Si $x_0 \in]a, b[$, alors $a < x_0 < b$ et $x_0 - a > 0$, $b - x_0 > 0$. On en déduit que

$$\delta = \min\left(\frac{x_0 - a}{2}, \frac{b - x_0}{2}\right) > 0.$$

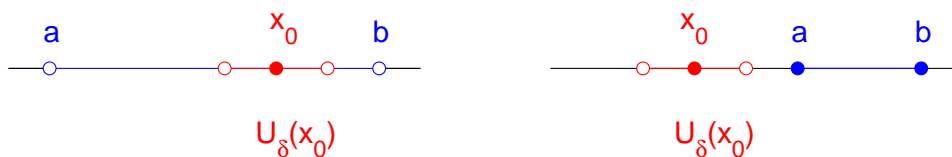
Notez que δ est la moitié de la distance entre x_0 et le bord de l'intervalle $]a, b[$ (voir la partie gauche de la figure ci-dessous). Comme il en résulte que

$$x_0 - \delta \geq x_0 - \frac{x_0 - a}{2} = \frac{x_0 + a}{2} > \frac{x_0 + x_0}{2} = x_0,$$

et

$$x_0 + \delta \leq x_0 + \frac{b - x_0}{2} = \frac{x_0 + b}{2} < \frac{x_0 + x_0}{2} = x_0,$$

on a $U_\delta(x_0) =]x_0 - \delta, x_0 + \delta[\subset]a, b[$ c'est-à-dire que si $d(x, x_0) < \delta$ alors $x \in]a, b[$. Nous avons montré que $]a, b[$ est un voisinage de x_0 . Comme ceci est vrai pour tout $x_0 \in]a, b[$, on en déduit que $]a, b[$ est un ensemble ouvert. On montre de façon analogue que $] - \infty, a[$ et $]b, +\infty[$ sont ouverts.



Si $x_0 \notin [a, b]$, on a soit $x_0 < a$, soit $x_0 > b$. Dans le premier cas $x_0 \in]-\infty, a[$ et comme cet intervalle est ouvert il existe $\delta > 0$ tel que $U_\delta(x_0) \subset]-\infty, a[\subset \mathbb{R} \setminus [a, b]$ (voir la partie droite de la figure ci-dessus). On montre de manière similaire que dans le second cas il existe $\delta > 0$ tel que $U_\delta(x_0) \subset]b, +\infty[\subset \mathbb{R} \setminus [a, b]$. On en déduit que l'intervalle $[a, b]$ est un ensemble fermé. Il en est de même des intervalles $] -\infty, a]$ et $[b, +\infty[$.

L'intervalle $[a, b[$ n'est ni ouvert, ni fermé. En effet, ce n'est pas un voisinage du point a car, pour tout $\delta > 0$, $U_\delta(a) =]a - \delta, a + \delta[\not\subset [a, b[$. On en conclut que $[a, b[$ n'est pas ouvert. De même, le complément $\mathbb{R} \setminus [a, b[=]-\infty, a[\cup]b, +\infty[$ n'est pas ouvert car ce n'est pas un voisinage du point b . On en déduit que $[a, b[$ n'est pas fermé. \triangleleft

Remarque 1.2 L'exemple précédent montre qu'un sous-ensemble de \mathbb{R}^n peut être ni ouvert ni fermé. Il peut aussi être à la fois ouvert et fermé. C'est le cas de \mathbb{R}^n et de l'ensemble vide.

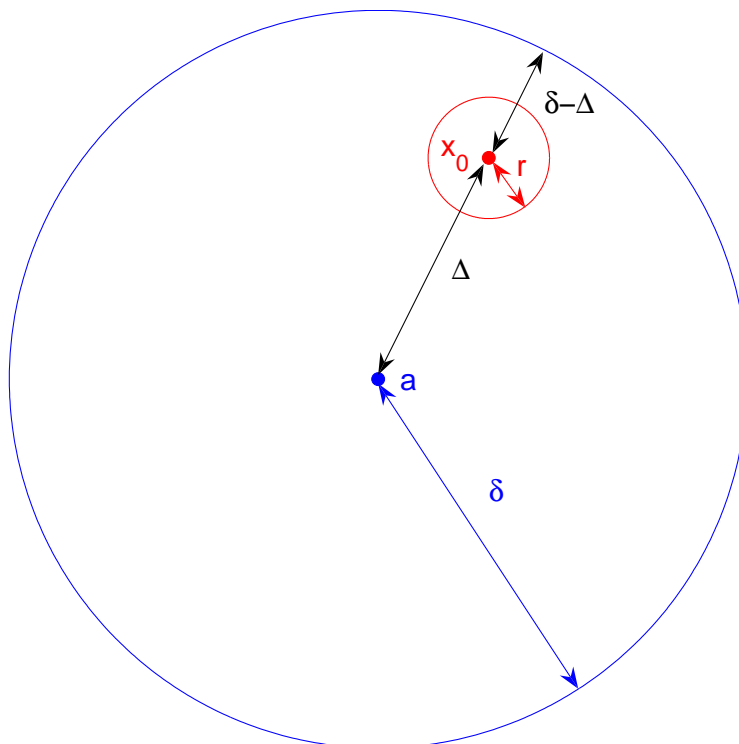
Exemple 1.2 Pour tous $a \in \mathbb{R}^n$ et $\delta > 0$, la boule $U_\delta(a)$ est un sous-ensemble ouvert de \mathbb{R}^n . Pour le démontrer, fixons un point $x_0 \in U_\delta(a)$ et posons $\Delta = d(x_0, a)$. Comme $\Delta < \delta$, on a

$$r = \frac{\delta - \Delta}{2} > 0.$$

Notez que r est la moitié de la distance entre x_0 et le bord de la boule $U_\delta(a)$ (voir la figure ci-dessous). Pour tout $x \in \mathbb{R}^n$ tel que $d(x, x_0) < r$ on a, en vertu de l'inégalité du triangle,

$$d(x, a) \leq d(x, x_0) + d(x_0, a) < r + \Delta = \frac{\delta - \Delta}{2} + \Delta = \frac{\delta + \Delta}{2} < \frac{\delta + \delta}{2} = \delta,$$

c'est-à-dire $x \in U_\delta(a)$. Nous avons donc montré que $U_r(x_0) \subset U_\delta(a)$, autrement dit que $U_\delta(a)$ est un voisinage de x_0 . Comme ceci est vrai pour tout $x_0 \in U_\delta(a)$, on en conclut que la boule $U_\delta(a)$ est ouverte.



On montre de façon très similaire que $\{x \in \mathbb{R}^n \mid d(x, a) \leq \delta\}$ est un ensemble fermé, c'est-à-dire que son complément $\{x \in \mathbb{R}^n \mid d(x, a) > \delta\}$ est ouvert. \triangleleft

Exemple 1.3 Le rectangle $]a, b[\times]c, d[$ est ouvert alors que le rectangle $[a, b] \times [c, d]$ est fermé. Le rectangle $]a, b[\times [c, d]$ n'est ni ouvert, ni fermé. \triangleleft

Les propriétés les plus importantes des ensembles ouverts/fermés sont les suivantes.

Théorème 1.3 1. Si J est un ensemble non vide et si, pour tout $j \in J$, O_j est un sous-ensemble ouvert de \mathbb{R}^n , alors leur réunion

$$\bigcup_{j \in J} O_j = \{x \in \mathbb{R}^n \mid x \in O_j \text{ pour un } j \in J\},$$

est un ensemble ouvert.

2. Si O_1, O_2, \dots, O_N sont des sous-ensembles ouverts de \mathbb{R}^n , leur intersection

$$\bigcap_{j=1}^N O_j = \{x \in \mathbb{R}^n \mid x \in O_j \text{ pour tout } j \in \{1, 2, \dots, N\}\},$$

est un ensemble ouvert.

3. Si F_1, F_2, \dots, F_N sont des sous-ensembles fermés de \mathbb{R}^n , alors leur réunion

$$\bigcup_{j=1}^N F_j = \{x \in \mathbb{R}^n \mid x \in F_j \text{ pour un } j \in \{1, 2, \dots, N\}\},$$

est un ensemble fermé.

4. Si J est un ensemble non vide et si, pour tout $j \in J$, F_j est un sous-ensemble fermé de \mathbb{R}^n , alors leur intersection

$$\bigcap_{j \in J} F_j = \{x \in \mathbb{R}^n \mid x \in F_j \text{ pour tout } j \in J\},$$

est un ensemble fermé.

Démonstration. 1. Si $x \in \bigcup_{j \in J} O_j$, il existe $j \in J$ tel que $x \in O_j$. Comme O_j est ouvert, il existe une boule $U_\delta(x)$ contenue dans O_j . On en déduit que $U_\delta(x)$ est aussi contenu dans la réunion $\bigcup_{j \in J} O_j$, ce qui montre que cette dernière est un ensemble ouvert.

2. Si $x \in O_1 \cap O_2 \cdots \cap O_N$, alors $x \in O_j$ pour tout $j \in \{1, 2, \dots, N\}$. Il existe donc des boules $U_{\delta_j}(x)$ telles que $U_{\delta_j}(x) \subset O_j$. Posons $\delta = \min(\delta_1, \delta_2, \dots, \delta_N)$. Comme $U_\delta(x) \subset U_{\delta_j}(x)$ pour tout $j \in \{1, 2, \dots, N\}$, on en déduit que $U_\delta(x) \subset O_1 \cap O_2 \cdots \cap O_N$ ce qui était à démontrer.

3. Pour tout $j \in \{1, 2, \dots, N\}$, F_j^c est ouvert. Par la propriété 2, $F_1^c \cap F_2^c \cdots \cap F_N^c$ est ouvert et son complément

$$(F_1^c \cap F_2^c \cdots \cap F_N^c)^c = F_1 \cup F_2 \cdots \cup F_N,$$

est fermé. On démontre la propriété 4 avec la même méthode. \square



L'intersection d'un nombre infini d'ensembles ouverts n'est pas nécessairement ouvert. Par exemple $U_{1/n}(a)$ est ouvert pour tout entier $n > 0$, mais

$$\bigcap_{n=1}^{\infty} U_{1/n}(a) = \{a\},$$

n'est pas ouvert. De même la réunion d'un nombre infini d'ensembles fermés n'est pas nécessairement fermé.

1.1.3 Adhérence, intérieur, bord

Un point $x \in \mathbb{R}^n$ est **intérieur** à $A \subset \mathbb{R}^n$ si A est un voisinage de x . L'ensemble de tous les points intérieurs à A est appelé l'**intérieur** de A et noté $\overset{\circ}{A}$. Comme tout point dont A est un voisinage est un élément de A , on a $\overset{\circ}{A} \subset A$.

Un point $x \in \mathbb{R}^n$ est **adhérent** à l'ensemble $A \subset \mathbb{R}^n$ si, pour tout $\delta > 0$, $U_\delta(x) \cap A$ n'est pas vide. L'ensemble de tous les points adhérents à A est appelé **adhérence** de A et noté \overline{A} . Comme tout point $x \in A$ est clairement adhérent à A , on a $A \subset \overline{A}$.

$x \in \mathbb{R}^n$ est un **point d'accumulation** de $A \subset \mathbb{R}^n$ si, pour tout $\delta > 0$, $U_\delta(x)$ contient un point de A distinct de x lui-même. Il est clair qu'un point d'accumulation de A est adhérent à A .

$x \in \mathbb{R}^n$ est un **point frontière** de $A \subset \mathbb{R}^n$ si, pour tout $\delta > 0$, $U_\delta(x)$ a une intersection non vide avec A et avec son complément. La **frontière** de A , qu'on appelle aussi **bord** de A est l'ensemble des points frontière de A . On dénote cet ensemble par ∂A .

$x \in \mathbb{R}^n$ est un **point isolé** de $A \subset \mathbb{R}^n$ si $x \in A$ et si il existe $\delta > 0$ tel que $U_\delta(x) \cap A = \{x\}$. Un point isolé de A est adhérent à A mais n'est pas un point d'accumulation de A . Tout point de A est soit isolé, soit un point d'accumulation de A .

Exemple 1.4 Tous les points de l'ensemble $A = \{(1/n, 0) \in \mathbb{R}^2 \mid n = 1, 2, \dots\}$ sont isolés. $(0, 0)$ est le seul point d'accumulation de A . L'adhérence de A est $\overline{A} = A \cup \{(0, 0)\}$ et son intérieur est vide $\overset{\circ}{A} = \emptyset$. Le bord de A est $\partial A = \overline{A}$.



◁

Exemple 1.5 La boule $A = U_r(a) \subset \mathbb{R}^n$ étant ouverte, on a $\overset{\circ}{A} = A$.

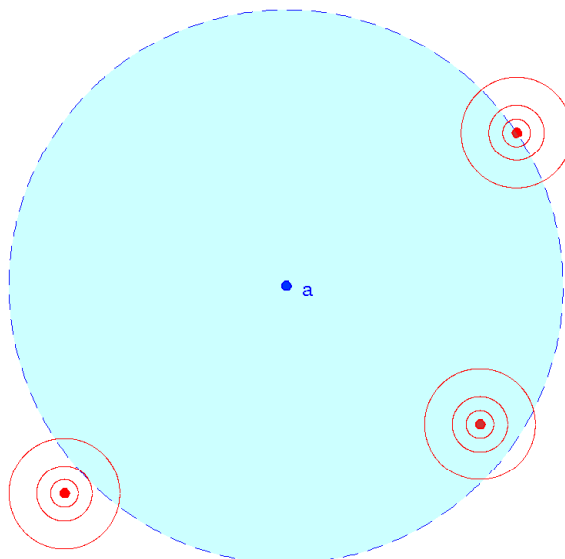
Si $d(x, a) = \Delta > r$, alors $U_\delta(x) \cap A = \emptyset$ pour tout $\delta < \Delta - r$ et par conséquence $x \notin \overline{A}$ et $x \notin \partial A$.

Si $d(x, a) = \Delta < r$, alors $x \in A \subset \overline{A}$ mais $U_\delta(x) \cap A^c = \emptyset$ pour tout $\delta < r - \Delta$ et donc $x \notin \partial A$.

Finalement, si $d(x, a) = r$, $U_\delta(x) \cap A \neq \emptyset$ et $U_\delta(x) \cap A^c \neq \emptyset$ pour tout $\delta > 0$. Par conséquence $x \in \overline{A}$ et $x \in \partial A$. On a donc

$$\overline{A} = \{x \in \mathbb{R}^n \mid d(x, a) \leq r\}, \quad \partial A = \{x \in \mathbb{R}^n \mid d(x, a) = r\}.$$

A n'a pas de point isolés et \overline{A} est l'ensemble de tous ses points d'accumulation.



Théorème 1.4 1. \bar{A} est fermé.

2. A est fermé $\Leftrightarrow A^c \cap \bar{A} = \emptyset \Leftrightarrow A = \bar{A}$.

3. \bar{A} est l'union de A et de ses points d'accumulation.

4. \bar{A} est le plus petit ensemble fermé contenant A .

5. $\overset{\circ}{A}$ est ouvert.

6. A est ouvert si et seulement si $A = \overset{\circ}{A}$.

7. $\overset{\circ}{A}$ est le plus grand ensemble ouvert contenu dans A .

8. $\partial A = \bar{A} \setminus \overset{\circ}{A} = \bar{A} \cap \bar{A}^c$.

Nous donnons ici la démonstration de ces résultats qui est à prendre comme un excellent exercice de logique.

Démonstration. 1. Nous devons montrer que \bar{A}^c est ouvert. Pour ce faire nous allons supposer qu'il ne l'est pas. Dans ce cas il existe $x \in \bar{A}^c$ tel que, pour tout $\delta > 0$, $U_{\delta/2}(x) \cap \bar{A}$ n'est pas vide. Soit y un élément de cette intersection. D'une part, comme $y \in \bar{A}$, $U_{\delta/2}(y) \cap A$ n'est pas vide. D'autre part, comme $d(x, y) < \delta/2$, l'inégalité du triangle $d(z, x) \leq d(z, y) + d(y, x) < d(z, y) + \delta/2$ nous dit que $U_{\delta/2}(y) \subset U_{\delta}(x)$. On en déduit que $U_{\delta}(x) \cap A$ n'est pas vide. Comme ceci est vrai pour tout $\delta > 0$, il en résulte que $x \in \bar{A}$, ce qui est une contradiction.

2. Nous montrerons la chaîne d'implications suivante :

$$A \text{ est fermé} \Rightarrow A^c \text{ est ouvert} \Rightarrow A^c \cap \bar{A} = \emptyset \Rightarrow A = \bar{A} \Rightarrow A \text{ est fermé,}$$

qui est logiquement équivalente à la proposition 2. La première implication de cette chaîne résulte de la définition d'un ensemble fermé.

Supposons A^c ouvert. Pour tout $x \in A^c$ il existe donc $\delta > 0$ tel que $U_{\delta}(x) \subset A^c$. On en déduit que $U_{\delta}(x) \cap A = \emptyset$ et donc que $x \notin \bar{A}$. Il en résulte bien que $A^c \cap \bar{A} = \emptyset$.

Comme $\bar{A} \setminus A = A^c \cap \bar{A}$, il suit de $A^c \cap \bar{A} = \emptyset$ et du fait que $A \subset \bar{A}$ que $A = \bar{A}$.

La dernière implication suit de la propriété 1.

3. Nous avons déjà remarqué que les éléments de A et ses points d'accumulation sont adhérents à A . Il reste donc à montrer que tout $x \in \bar{A} \setminus A$ est un point d'accumulation de A . Comme $x \in \bar{A}$, pour tout $\delta > 0$ la boule $U_{\delta}(x)$ contient un point de A . Comme $x \notin A$, ce point est distinct de x . On peut donc conclure que x est bien un point d'accumulation de A .

4. Soit F un ensemble fermé contenant A et $x \in \bar{A}$. Pour tout $\delta > 0$, $U_{\delta}(x)$ contient un point de A . Il contient donc un point de F . On en déduit que $x \in \bar{F}$. F étant fermé, la propriété 2 montre que $\bar{F} = F$. Il en découle que $x \in F$ ce qui nous permet de conclure que $\bar{A} \subset F$. Tout ensemble fermé contenant A contient donc \bar{A} . Comme ce dernier est fermé et contient A , on en déduit que c'est le plus petit ensemble ayant cette propriété.

5. Si $x \in \overset{\circ}{A}$, A est un voisinage de x . Il existe donc $\delta > 0$ tel que $U_{\delta}(x) \subset A$. Pour tout $y \in U_{\delta/2}(x)$, l'inégalité du triangle montre que $U_{\delta/2}(y) \subset U_{\delta}(x) \subset A$. On en déduit que A est un voisinage de y et que par conséquence $U_{\delta/2}(x) \subset \overset{\circ}{A}$.

6. Si A est ouvert, c'est un voisinage de chacun de ses points. On en déduit que $A \subset \overset{\circ}{A}$. Comme on sait d'autre part que $\overset{\circ}{A} \subset A$ on en conclut que $\overset{\circ}{A} = A$. La réciproque est une conséquence immédiate de la propriété 5.

7. Soit $O \subset A$ un ensemble ouvert. Comme O est un voisinage de chacun de ses points, on en déduit immédiatement que $O \subset \overset{\circ}{A}$. Tout ensemble ouvert contenu dans A est donc aussi contenu dans $\overset{\circ}{A}$. Comme ce dernier est un ensemble ouvert contenu dans A , c'est bien le plus grand ensemble possédant cette propriété.

8. Nous allons démontrer la chaîne d'inclusions

$$\partial A \subset \bar{A} \cap \bar{A}^c \subset \bar{A} \setminus \overset{\circ}{A} \subset \partial A,$$

qui est équivalente à la proposition 8.

Si $x \in \partial A$ alors, pour tout $\delta > 0$, on a d'une part $U_\delta(x) \cap A \neq \emptyset$ et donc $x \in \bar{A}$ et d'autre part $U_\delta(x) \cap A^c \neq \emptyset$ et donc $x \in \bar{A}^c$. On en déduit bien que $\partial A \subset \bar{A} \cap \bar{A}^c$.

Si $x \in \bar{A} \cap \bar{A}^c$, alors d'une part $x \in \bar{A}$ et d'autre part, pour tout $\delta > 0$, $U_\delta(x) \cap A^c \neq \emptyset$, ce qui implique que $U_\delta(x) \not\subset A$. On en déduit que A n'est pas un voisinage de x , c'est-à-dire que $x \notin \overset{\circ}{A}$. Il en résulte que $\bar{A} \cap \bar{A}^c \subset \bar{A} \setminus \overset{\circ}{A}$.

Si $x \in \bar{A} \setminus \overset{\circ}{A}$ on a d'une part $x \in \bar{A}$ et donc, pour tout $\delta > 0$, $U_\delta(x) \cap A \neq \emptyset$. D'autre part, A n'est pas un voisinage de x , c'est-à-dire que pour tout $\delta > 0$, $U_\delta(x) \not\subset A$ et donc $U_\delta(x) \cap A^c \neq \emptyset$. \square

1.1.4 Suites dans \mathbb{R}^n

Une suite dans \mathbb{R}^n est une fonction

$$\begin{aligned} \mathbb{N} &\rightarrow \mathbb{R}^n \\ k &\mapsto x_k, \end{aligned}$$

qu'on dénote par $(x_k)_{k \in \mathbb{N}}$. Une suite dans \mathbb{R}^n est donc définie par les n suites numériques des coordonnées du point $x_k \in \mathbb{R}^n$ que nous noterons

$$x_k = (x_{1,k}, x_{2,k}, \dots, x_{n,k}).$$

Si $A \subset \mathbb{R}^n$ et $x_k \in A$ pour tout $k \in \mathbb{N}$, nous dirons que $(x_k)_{k \in \mathbb{N}}$ est une suite dans A .

La suite $(x_k)_{k \in \mathbb{N}}$ dans \mathbb{R}^n est **convergente** et **converge** vers la **limite** $a \in \mathbb{R}^n$ si

$$\lim_{k \rightarrow \infty} d(x_k, a) = 0,$$

c'est-à-dire si le point x_k s'approche aussi près qu'on le désire du point limite a lorsque $k \rightarrow \infty$. On écrit alors

$$\lim_{k \rightarrow \infty} x_k = a.$$

Remarquez qu'on peut reformuler ceci en termes de boules de la manière suivante : la suite $(x_k)_{k \in \mathbb{N}}$ converge vers le point a si et seulement si, pour tout $\delta > 0$ il existe $K \in \mathbb{N}$ tel que $x_k \in U_\delta(a)$ pour tout $k > K$. Comme toutes les boules $U_\delta(a)$ sont des voisinages de a et, réciproquement, tout voisinage de a contient une telle boule, on peut aussi dire que la suite $(x_k)_{k \in \mathbb{N}}$ converge vers le point a si et seulement si, pour tout voisinage V de a il existe $K \in \mathbb{N}$ tel que $x_k \in V$ pour tout $k > K$. La limite d'une suite convergente est unique. En effet, si

$$\lim_{k \rightarrow \infty} x_k = a \quad \text{et} \quad \lim_{k \rightarrow \infty} x_k = b,$$

alors l'inégalité du triangle $d(a, b) \leq d(a, x_k) + d(x_k, b)$, valable pour tout $k \in \mathbb{N}$, implique

$$d(a, b) \leq \lim_{k \rightarrow \infty} d(a, x_k) + d(x_k, b) = 0,$$

ce qui nous permet de conclure que $a = b$.

Exemple 1.6 La suite

$$x_k = \left(1 + \frac{1}{k}, \frac{2}{k^2} \right),$$

converge vers le point $a = (1, 0)$ car la suite numérique

$$d(x_k, a) = \sqrt{\frac{1}{k^2} + \frac{4}{k^4}},$$

tend vers 0. \triangleleft

Comme le montre cet exemple, on doit connaître la limite de la suite pour pouvoir appliquer la définition de la convergence que nous venons de donner puisqu'il est nécessaire de calculer la suite numérique $d(x_k, a)$. En pratique on calcule la limite d'une suite dans \mathbb{R}^n à l'aide du résultat suivant qui ramène ce problème au calcul des limites de n suites numériques.

Théorème 1.5 La suite $(x_k)_{k \in \mathbb{N}}$ dans \mathbb{R}^n est convergente si et seulement si les n suites numériques $(x_{1,k})_{k \in \mathbb{N}}$, $(x_{2,k})_{k \in \mathbb{N}}$, ..., $(x_{n,k})_{k \in \mathbb{N}}$ des coordonnées de x_k sont toutes convergentes. De plus

$$\lim_{k \rightarrow \infty} x_k = (a_1, a_2, \dots, a_n),$$

où

$$a_j = \lim_{k \rightarrow \infty} x_{j,k},$$

pour $j = 1, 2, \dots, n$.

Démonstration. Pour tout $j \in \{1, 2, \dots, n\}$ on a

$$|x_{j,k} - a_j|^2 \leq |x_{1,k} - a_1|^2 + |x_{2,k} - a_2|^2 + \dots + |x_{n,k} - a_n|^2.$$

En prenant la racine carrée des deux côtés de cette inégalité on obtient

$$|x_{j,k} - a_j| \leq d(x_k, a).$$

On en déduit que si $d(x_k, a)$ tend vers 0 lorsque $k \rightarrow \infty$, il en va de même de $|x_{j,k} - a_j|$. Ceci montre bien que

$$\lim_{k \rightarrow \infty} x_k = a \quad \Rightarrow \quad \lim_{k \rightarrow \infty} x_{j,k} = a_j.$$

Réciproquement, si

$$\lim_{k \rightarrow \infty} x_{j,k} = a_j,$$

pour tout $j \in \{1, 2, \dots, n\}$, alors $|x_{j,k} - a_j|$ tend vers 0 lorsque $k \rightarrow \infty$ pour tout $j \in \{1, 2, \dots, n\}$ et par conséquence

$$d(x_k, a) = \sqrt{|x_{1,k} - a_1|^2 + |x_{2,k} - a_2|^2 + \dots + |x_{n,k} - a_n|^2},$$

tend également vers 0. □

Exemple 1.7 La suite dans \mathbb{R}^3 définie par

$$x_k = \left(k \sin \frac{1}{k}, \frac{k}{\sqrt{1+k^2}}, e^{-k} \right),$$

pour $k \geq 1$ converge vers $(1, 1, 0)$. En effet, on vérifie aisément que

$$\lim_{k \rightarrow \infty} k \sin \frac{1}{k} = 1, \quad \lim_{k \rightarrow \infty} \frac{k}{\sqrt{1+k^2}} = 1, \quad \lim_{k \rightarrow \infty} e^{-k} = 0.$$

◁

Exemple 1.8 La suite dans \mathbb{R}^2 définie par

$$x_k = (k, e^{-k}),$$

ne converge pas puisque la suite de ses premières coordonnées $x_{1,k} = k$ n'est pas convergente. ◁

Dans le paragraphe précédent, nous avons vu que pour vérifier qu'un ensemble $A \subset \mathbb{R}^n$ est fermé on doit montrer que son complément est ouvert. Le théorème suivant nous permet une approche alternative.

Théorème 1.6 Un point a est adhérent à l'ensemble $A \subset \mathbb{R}^n$ si et seulement si il existe une suite $(x_k)_{k \in \mathbb{N}}$ dans A qui converge vers a .

Démonstration. Supposons que $a \in \bar{A}$. Pour tout $k \in \mathbb{N}$ on pose $\delta_k = (1+k)^{-1}$ en remarquant que $\delta_k \rightarrow 0$ lorsque $k \rightarrow \infty$. Comme, pour tout $k \in \mathbb{N}$, $U_{\delta_k}(a) \cap A$ n'est pas vide, il existe $x_k \in A$ tel que $d(x_k, a) < \delta_k$. La suite $(x_k)_{k \in \mathbb{N}}$ ainsi obtenue converge bien vers a .

Supposons maintenant que la suite $(x_k)_{k \in \mathbb{N}}$, telle que $x_k \in A$ pour tout k , converge vers a . Comme nous l'avons remarqué au début de ce paragraphe, pour tout $\delta > 0$ il existe $K \in \mathbb{N}$ tel que $x_k \in U_\delta(a)$ pour tout $k > K$. Comme $x_k \in A$, on peut en déduire que $U_\delta(a) \cap A$ n'est pas vide et donc que a est adhérent à A . □

Remarque 1.7 En modifiant légèrement ces arguments, on montre aisément que a est un point d'accumulation de A si et seulement si il existe une suite dans $A \setminus \{a\}$ qui converge vers a .

Corollaire 1.8 *Un ensemble $A \subset \mathbb{R}^n$ est fermé si et seulement si, pour toute suite convergente $(x_k)_{k \in \mathbb{N}}$ dans A on a $\lim_{k \rightarrow \infty} x_k \in A$.*

Démonstration. Supposons que A soit fermé. Le théorème 1.6 nous permet de conclure que si $(x_k)_{k \in \mathbb{N}}$ est une suite dans A telle que $\lim_{k \rightarrow \infty} x_k = a$ alors $a \in \bar{A}$. Comme le théorème 1.4 nous dit que $\bar{A} = A$, on a bien $a \in A$.

Réciproquement, si toute suite convergente dans A converge vers un point de A , le théorème 1.6 nous permet de conclure que $\bar{A} \subset A$. Comme A est toujours inclus dans \bar{A} , on en déduit que $A = \bar{A}$ et le théorème 1.4 nous permet de conclure que A est fermé. \square

Exemple 1.9 Une droite est un sous-ensemble fermé du plan. En effet, une droite du plan est décrite par

$$\mathcal{D} = \{(x_1, x_2) \in \mathbb{R}^2 \mid Ax_1 + Bx_2 = C\},$$

où A, B et C sont des constantes. Si $(x_k)_{k \in \mathbb{N}}$ est une suite convergente dans \mathcal{D} alors

$$Ax_{1,k} + Bx_{2,k} = C,$$

pour tout $k \in \mathbb{N}$. Si $\lim_{k \rightarrow \infty} x_k = a = (a_1, a_2)$, le théorème 1.5 nous permet de dire que $a_1 = \lim_{k \rightarrow \infty} x_{1,k}$ et $a_2 = \lim_{k \rightarrow \infty} x_{2,k}$. On en déduit que

$$Aa_1 + Ba_2 = \lim_{k \rightarrow \infty} Ax_{1,k} + Bx_{2,k} = C,$$

c'est-à-dire que $a \in \mathcal{D}$. Le corollaire 1.8 nous permet de conclure que \mathcal{D} est fermé. Un raisonnement analogue montre que tout sous-espace affine de \mathbb{R}^n est fermé. \triangleleft

1.1.5 Continuité

Si $A \subset \mathbb{R}^n$, une fonction f définie sur A est une relation qui, à chaque point $x \in A$ associe un nombre réel $f(x)$. Comme un point $x \in A$ est déterminé par ses n coordonnées $x = (x_1, x_2, \dots, x_n)$, on peut voir f comme une fonction de n variables réelles. C'est pourquoi on peut écrire $f(x) = f(x_1, x_2, \dots, x_n)$. Voici quelques exemples de telles fonctions.

Exemple 1.10 La fonction

$$f(x_1, x_2) = \sqrt{x_1^2 + x_2^2},$$

est définie sur \mathbb{R}^2 tout entier. On peut aussi la décrire par $f(x) = d(x, (0, 0)) = \|x\|$.

La fonction

$$g(u, v, w) = \frac{uv}{\sqrt{(w-u)(w-v)}}$$

est définie sur $A = \{(u, v, w) \in \mathbb{R}^3 \mid (w-u)(w-v) > 0\}$.

La fonction

$$h(x, y, z, t) = \frac{e^{x^2+y^2+z^2-t^2}}{(x-1)(y-1)(z-1)(t+1)},$$

est définie sur $\mathbb{R}^4 \setminus \{(1, 1, 1, -1)\}$. \triangleleft

1.1.5.1 Limites de fonctions

Définition 1.9 Soit f une fonction définie sur $A \subset \mathbb{R}^n$ et a un point d'accumulation de A . On dit que $f(x)$ converge lorsque x tend vers a si il existe $L \in \mathbb{R}$ tel que, pour toute suite $(x_k)_{k \in \mathbb{N}}$ dans $A \setminus \{a\}$ convergeant vers a , on a

$$\lim_{k \rightarrow \infty} f(x_k) = L.$$

Dans ce cas on dit que L est la *limite* de f lorsque x tend vers a et on écrit

$$\lim_{x \rightarrow a} f(x) = L.$$

Remarque 1.10 L'existence et la valeur de cette limite peut dépendre de A , c'est pourquoi on écrit parfois

$$\lim_{A \ni x \rightarrow a} f(x) = L,$$

pour bien spécifier l'ensemble A considéré dans cette limite. L'exemple suivant illustre cette problématique.

Exemple 1.11 Considérons la fonction définie sur \mathbb{R}^2 par

$$f(x, y) = \begin{cases} 1 & \text{si } x > 0, \\ 0 & \text{si } x \leq 0. \end{cases}$$

Les suites $x_k = (1/k, 0)$ et $y_k = (-1/k, 0)$ convergent toutes deux vers $(0, 0)$. La suite $f(x_k) = 1$ converge vers 1 alors que la suite $f(y_k) = 0$ converge vers 0. On en conclut que f ne converge pas lorsque x tend vers $(0, 0)$. Cependant, si on restreint f au demi-plan $A = \{(x, y) \in \mathbb{R}^2 \mid x > 0\}$ on a $f(x) = 1$ pour tout $x \in A$ et donc $\lim_{k \rightarrow \infty} f(x_k) = 1$ pour toute suite $(x_k)_{k \in \mathbb{N}}$ dans A convergeant vers $(0, 0) \in \partial A$. Dans ce cas on peut écrire

$$\lim_{A \ni x \rightarrow (0,0)} f(x) = 1.$$

On montre de manière similaire que si $B = \{(x, y) \in \mathbb{R}^2 \mid x < 0\}$, alors

$$\lim_{B \ni x \rightarrow (0,0)} f(x) = 0.$$

◁

Théorème 1.11 $\lim_{x \rightarrow a} f(x) = L$ si et seulement si, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $0 < d(x, a) < \delta$ et $x \in A$ impliquent que $|f(x) - L| < \varepsilon$.

Démonstration. Nous démontrons la nécessité de la condition formulée dans le théorème par contraposition. Nous supposons donc que $\lim_{x \rightarrow a} f(x) = L$ mais qu'il existe $\varepsilon > 0$ tel que, pour tout $\delta > 0$ il existe $x \in A$ avec $0 < d(x, a) < \delta$ et $|f(x) - L| \geq \varepsilon$. Dans ce cas, pour tout $k \in \mathbb{N}$ il existe un $x_k \in A \setminus \{a\}$ tel que $d(x_k, a) < (1 + k)^{-1}$ et $|f(x_k) - L| \geq \varepsilon$. On en déduit que la suite $(x_k)_{k \in \mathbb{N}}$ dans $A \setminus \{a\}$ converge vers a , mais que $f(x_k)$ ne converge pas vers L , une contradiction avec notre hypothèse de départ.

Pour démontrer que la condition du théorème est suffisante, on remarque que si cette condition est satisfaite et si la suite $(x_k)_{k \in \mathbb{N}}$ dans $A \setminus \{a\}$ converge vers a , alors pour tout $\varepsilon > 0$ il existe $K \in \mathbb{N}$ tel que $0 < d(x_k, a) < \delta$ pour tout $k > K$, ce qui implique que $|f(x_k) - L| < \varepsilon$. On en conclut que $f(x_k)$ converge vers L . □

Remarque 1.12 On peut paraphraser ce théorème en disant que $f(x)$ converge vers L lorsque x tend vers a si et seulement si la distance entre $f(x)$ et L peut être rendue aussi petite qu'on veut en rapprochant suffisamment le point x du point a .

1.1.5.2 Fonctions continues sur \mathbb{R}^n

Définition 1.13 Une fonction f définie sur $A \subset \mathbb{R}^n$ est *continue* en $a \in A$ si a est un point isolé de A ou si

$$\lim_{x \rightarrow a} f(x) = f(a).$$

On dit que f est continue sur $X \subset A$ si elle est continue en chaque point de X .

Exemple 1.12 La fonction définie par

$$f(x, y) = \begin{cases} 0 & \text{si } (x, y) = (0, 0), \\ \frac{\sin x \sin y}{\sqrt{x^2 + y^2}} & \text{sinon,} \end{cases}$$

est continue sur \mathbb{R}^2 . Pour le prouver, nous devons montrer qu'elle est continue en chaque point $a \in \mathbb{R}^2$. Considérons tout d'abord le cas particulier $a = (0, 0)$. Pour montrer que f est continue en ce point nous devons montrer que si $(x_k)_{k \in \mathbb{N}}$ est une suite qui converge vers $(0, 0)$, alors la suite numérique $f(x_k)$ converge vers 0. Comme $|\sin x| \leq |x|$ pour tout $x \in \mathbb{R}$, on a

$$|f(x, y)| = \left| \frac{\sin x \sin y}{\sqrt{x^2 + y^2}} \right| \leq \frac{|x| |y|}{\sqrt{x^2 + y^2}},$$

pour tout $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Comme d'autre part

$$|x|^2 - 2|x| |y| + |y|^2 = (|x| - |y|)^2 \geq 0,$$

on a l'inégalité

$$|x| |y| \leq \frac{1}{2} (|x|^2 + |y|^2) = \frac{1}{2} (x^2 + y^2), \quad (1.2)$$

pour tout $(x, y) \in \mathbb{R}^2$. On en déduit que

$$|f(x, y)| \leq \frac{1}{2} \frac{x^2 + y^2}{\sqrt{x^2 + y^2}} = \frac{1}{2} \sqrt{x^2 + y^2}.$$

En remarquant que cette inégalité, que nous avons démontrée pour $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$, est trivialement satisfaite pour $(x, y) = (0, 0)$, nous pouvons écrire que

$$|f(x_k)| \leq \frac{1}{2} d(x_k, (0, 0)),$$

et en conclure que $f(x_k)$ tend bien vers 0 si la suite x_k converge vers $(0, 0)$.

Considérons maintenant le cas $a = (u, v) \neq (0, 0)$. Si la suite (x_k, y_k) converge vers a , alors

$$\lim_{k \rightarrow \infty} x_k = u, \quad \lim_{k \rightarrow \infty} y_k = v,$$

et par conséquence


$$\lim_{k \rightarrow \infty} \sin x_k \sin y_k = \sin u \sin v, \quad \lim_{k \rightarrow \infty} \sqrt{x_k^2 + y_k^2} = \sqrt{u^2 + v^2}.$$

Comme $\sqrt{u^2 + v^2} \neq 0$, on en conclut que

$$\lim_{k \rightarrow \infty} f(x_k, y_k) = \lim_{k \rightarrow \infty} \frac{\sin x_k \sin y_k}{\sqrt{x_k^2 + y_k^2}} = \frac{\sin u \sin v}{\sqrt{u^2 + v^2}} = f(u, v).$$

◁

Remarque 1.14 A l'aide du théorème 1.11 on peut dire que f est continue en a si et seulement si, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $d(x, a) < \delta$ implique $|f(x) - f(a)| < \varepsilon$. Autrement dit, pour toute boule $U_\varepsilon(f(a))$ il existe une boule $U_\delta(a)$ telle que $f(x) \in U_\varepsilon(f(a))$ pour tout $x \in U_\delta(a)$.



Si f est continue en $\mathbf{a} = (a_1, a_2, \dots, a_n)$, il en résulte immédiatement que la fonction

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

est continue en $t = a_j$, et ceci pour tout $j \in \{1, 2, \dots, n\}$. Cependant la continuité de ces fonctions n'est pas suffisante pour assurer la continuité de f en \mathbf{a} , comme le montre l'exemple suivant.

Exemple 1.13 Considérons la fonction f définie par

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

En remarquant que $f(x, 0) = 0$ et $f(0, y) = 0$ on en conclut que les deux fonctions d'une variable $t \mapsto f(t, 0)$ et $t \mapsto f(0, t)$ sont continues en $t = 0$. Considérons, pour θ fixé, la suite $x_k = (k^{-1} \cos \theta, k^{-1} \sin \theta)$ dans \mathbb{R}^2 . On a $\lim_{k \rightarrow \infty} x_k = 0$. Cependant, comme $f(x_k) = 2 \cos \theta \sin \theta = \sin 2\theta$ on remarque que

$$\lim_{k \rightarrow \infty} f(x_k) = \sin 2\theta,$$

dépend de θ , et donc du choix de la suite x_k . Ceci montre que f n'a pas de limite lorsque $x \rightarrow (0, 0)$. f n'est donc pas continue en $(0, 0)$. ◁

En pratique, on peut souvent mettre à profit les règles élémentaires suivantes pour étudier la continuité d'une fonction :

1. Les fonctions "coordonnées" $X_j(x_1, x_2, \dots, x_n) = x_j$ sont continues sur \mathbb{R}^n .
2. Si f et g sont deux fonctions continues en $\mathbf{a} \in \mathbb{R}^n$, alors les fonctions

$$\begin{aligned} (f + g)(x) &= f(x) + g(x), \\ (fg)(x) &= f(x)g(x), \end{aligned}$$

sont continues en \mathbf{a} . Si en plus $g(\mathbf{a}) \neq 0$, alors la fonction

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)},$$

est continue en \mathbf{a} .

3. Si f_1, f_2, \dots, f_m sont des fonctions continues en $\mathbf{a} \in \mathbb{R}^n$ et si g est une fonction continue en $\mathbf{b} = (f_1(\mathbf{a}), f_2(\mathbf{a}), \dots, f_m(\mathbf{a})) \in \mathbb{R}^m$ la fonction $F(x) = g(f_1(x), f_2(x), \dots, f_m(x))$ est continue en \mathbf{a} .

Exemple 1.14 La fonction


$$f(x_1, x_2, x_3) = e^{x_1^2 + x_2^2} \cos(x_3) + \frac{(x_1^2 - x_3^2)}{x_1^2 + x_2^2 + x_3^2},$$

est continue en $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$. ◁

Théorème 1.15 Si f est continue sur \mathbb{R}^n et $A \subset \mathbb{R}$ est ouvert (resp. fermé) alors

$$f^{-1}(A) = \{x \in \mathbb{R}^n \mid f(x) \in A\},$$

est ouvert (resp. fermé).



L'image, par une fonction continue, d'un ensemble ouvert (resp. fermé) n'est pas nécessairement un ensemble ouvert (resp. fermé).

1.1.5.3 Continuité uniforme

Définition 1.16 Une fonction f est *uniformément continue* sur $A \subset \mathbb{R}^n$ si, pour tout $\varepsilon > 0$, il existe un $\delta > 0$ tel que $x, x' \in A$ et $d(x, x') < \delta$ impliquent $|f(x) - f(x')| < \varepsilon$.

Autrement dit, f est uniformément continue sur A si

$$\lim_{\delta \rightarrow 0} \sup \{|f(x) - f(x')| \mid x, x' \in A, d(x, x') < \delta\} = 0.$$

Exemple 1.15 La fonction

$$f(x) = \frac{1}{1 + x_1^2 + x_2^2},$$

est uniformément continue sur \mathbb{R}^2 . En effet, le développement

$$f(x) - f(x') = \frac{x_1'^2 + x_2'^2 - x_1^2 - x_2^2}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)} = \frac{(x_1' + x_1)(x_1' - x_1) + (x_2' + x_2)(x_2' - x_2)}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)},$$

permet d'obtenir l'estimation

$$|f(x) - f(x')| \leq \frac{|x_1| + |x_1'|}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)} |x_1 - x_1'| + \frac{|x_2| + |x_2'|}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)} |x_2 - x_2'|.$$

Comme

$$\begin{aligned} \frac{|x_1| + |x_1'|}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)} &= \frac{|x_1|}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)} + \frac{|x_1'|}{(1 + x_1^2 + x_2^2)(1 + x_1'^2 + x_2'^2)} \\ &\leq \frac{|x_1|}{(1 + x_1^2)} + \frac{|x_1'|}{(1 + x_1'^2)} \\ &\leq \frac{1}{2} + \frac{1}{2} \leq 1, \end{aligned}$$

on déduit aisément que

$$|f(x) - f(x')| \leq |x_1 - x_1'| + |x_2 - x_2'| \leq d(x, x').$$

On a donc

$$\sup \{|f(x) - f(x')| \mid x, x' \in \mathbb{R}^2, d(x, x') < \delta\} \leq \delta,$$

et par conséquence

$$\lim_{\delta \rightarrow 0} \sup \{|f(x) - f(x')| \mid x, x' \in \mathbb{R}^2, d(x, x') < \delta\} = 0.$$

◁

Définition 1.17 1. Un sous-ensemble $B \subset \mathbb{R}^n$ est *borné* s'il est contenu dans une boule $U_r(x)$.
2. Un sous-ensemble $K \subset \mathbb{R}^n$ est *compact* s'il est fermé et borné.

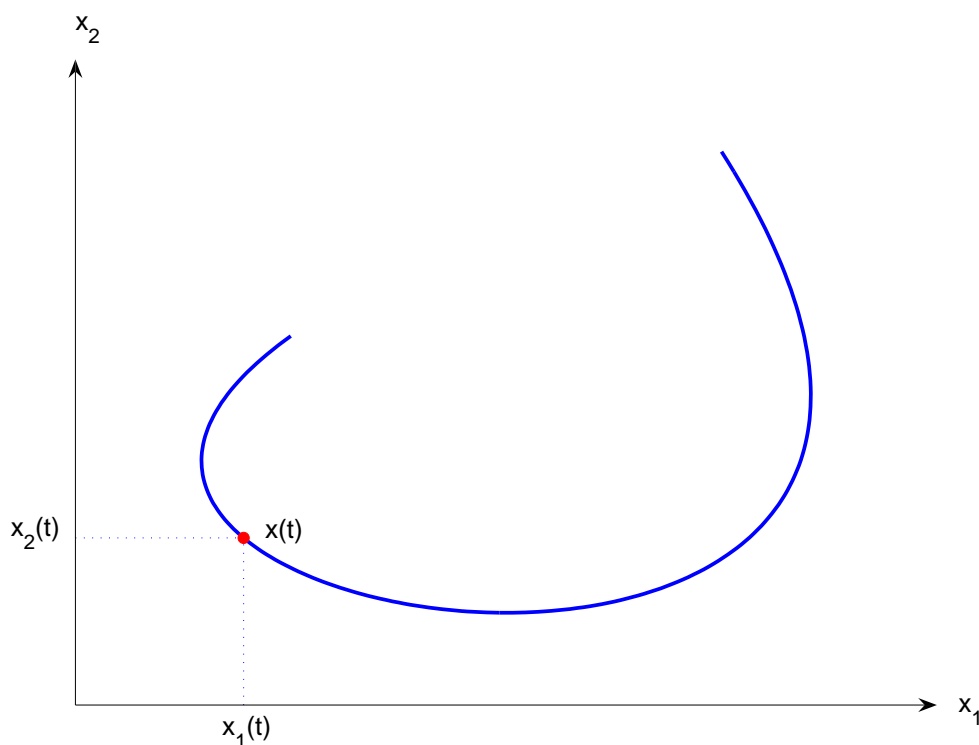
Nous admettrons le résultat suivant sans démonstration.

Théorème 1.18 Toute fonction continue sur un ensemble compact y est uniformément continue.

1.2 Fonctions différentiables

1.2.1 Courbes différentiables

Une fonction $x : I =]a, b[\rightarrow \mathbb{R}^2$, est déterminée par les deux fonctions $x_1 : I \rightarrow \mathbb{R}$ et $x_2 : I \rightarrow \mathbb{R}$ telles que $x(t) = (x_1(t), x_2(t))$.



La fonction x est **différentiable** en $t \in I$ si la limite

$$x'(t) = \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h},$$

existe. Comme

$$\frac{x(t+h) - x(t)}{h} = \left(\frac{x_1(t+h) - x_1(t)}{h}, \frac{x_2(t+h) - x_2(t)}{h} \right),$$

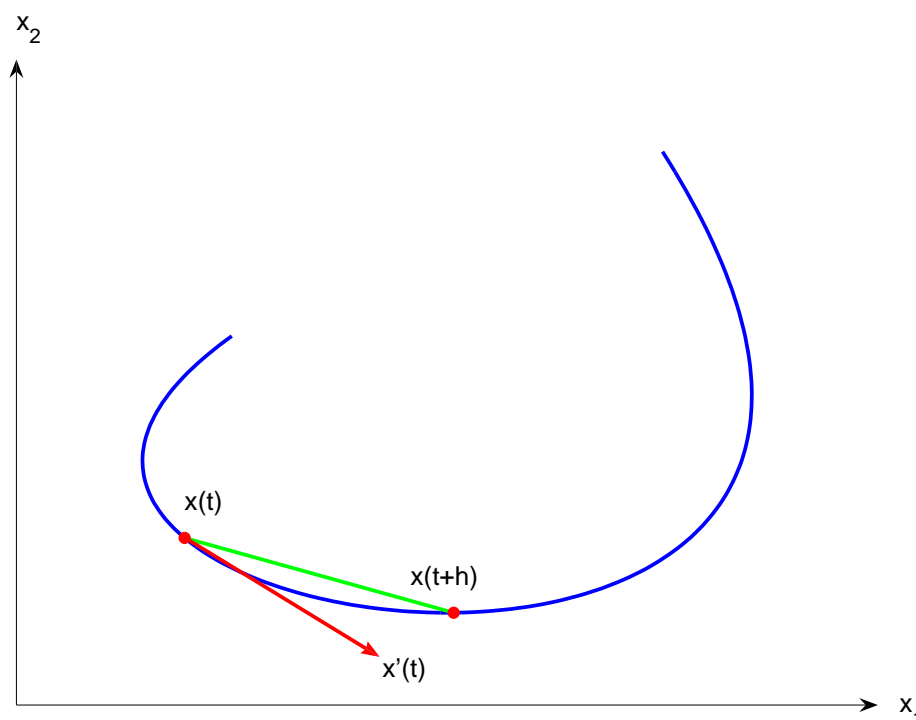
on en déduit que x est différentiable en t si et seulement si les deux fonctions x_1 et x_2 sont différentiables en t . Dans ce cas on a

$$\frac{dx}{dt}(t) = x'(t) = (x'_1(t), x'_2(t)).$$

Définition 1.19 Une **courbe (paramétrée) différentiable** dans le plan \mathbb{R}^2 est une fonction différentiable

$$\begin{aligned} x :]a, b[&\rightarrow \mathbb{R}^2 \\ t &\mapsto x(t). \end{aligned}$$

L'ensemble des points $\{x(t) \mid t \in]a, b[\} \subset \mathbb{R}^2$ est la **trace** de cette courbe. La dérivée $x'(t)$ est le **vecteur vitesse** (ou simplement la vitesse) de la courbe au point $x(t)$.



Lorsqu'il n'est pas nul, le vecteur vitesse $x'(t)$ dirige la tangente à la courbe au point $x(t)$ et le vecteur normalisé

$$T(t) = \frac{x'(t)}{\|x'(t)\|},$$

est le **vecteur tangent** à la courbe en $x(t)$.

Exemple 1.16 La trace de la fonction $x:]0, \pi[\rightarrow \mathbb{R}^2$ définie par $x(t) = (\cos t, \sin t)$ est un demi cercle de rayon 1 centré à l'origine. Cette courbe est différentiable, son vecteur vitesse

$$x'(t) = (-\sin t, \cos t),$$

ne s'annule jamais. Sa tangente au point $x(t)$ est donnée par $\{x(t) + \lambda x'(t) \mid \lambda \in \mathbb{R}\}$, soit

$$\{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 = \cos t - \lambda \sin t, x_2 = \sin t + \lambda \cos t, \lambda \in \mathbb{R}\}.$$

L'élimination du paramètre λ permet d'obtenir l'équation de cette tangente

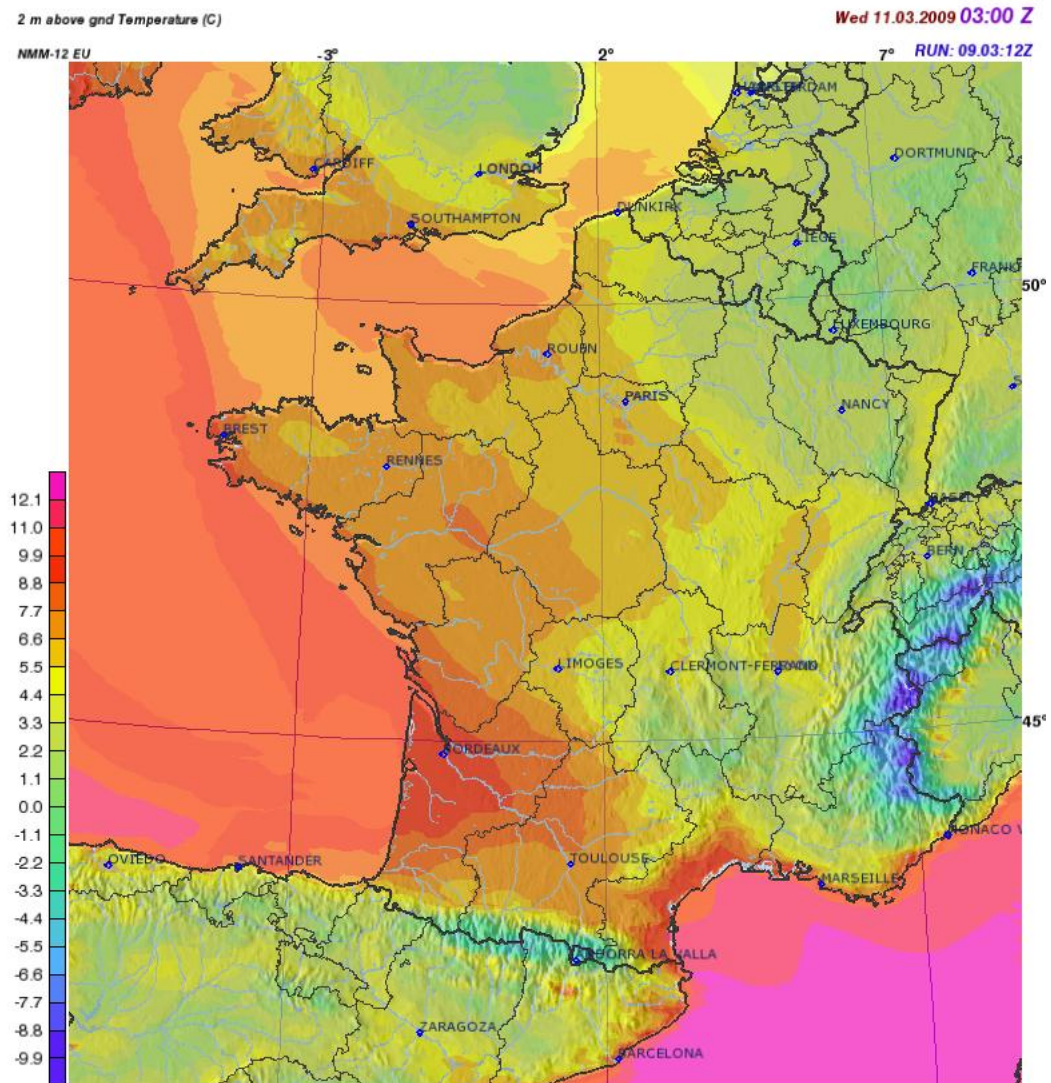
$$x_1 \cos t + x_2 \sin t = 1.$$

◁

Remarque 1.20 On généralise facilement ces constructions aux fonctions $x: I \rightarrow \mathbb{R}^n$ pour $n > 2$.

1.2.2 Champs scalaires

Un **champ scalaire** sur une région Ω du plan affine \mathbb{R}^2 est une fonction $f: \Omega \rightarrow \mathbb{R}$ c'est-à-dire une relation qui, à chaque point $x = (x_1, x_2) \in \Omega$ associe une valeur $f(x) = f(x_1, x_2) \in \mathbb{R}$. On dira que Ω est le **domaine** de la fonction f qu'on notera $D(f)$. Dans les applications, un champ scalaire décrit souvent les variations d'un paramètre en fonction de la position dans l'espace physique. Exemple : la température au niveau du sol en fonction de la latitude et de la longitude est un champ scalaire. La figure suivante en donne une représentation codée par une échelle de couleurs.



(c) 2009 by meteoblue TM - www.meteoblue.com

D'autres possibilités sont envisageables. Exemple : dans une mole de gaz parfait la loi d'Avogadro

$$pV = RT,$$

lie la pression p , le volume V et la température absolue T (R , la constante des gaz parfait, ne varie pas). Cette relation permet d'exprimer la pression comme une fonction de deux "variables" : la température et le volume

$$p = f(T, V) = \frac{RT}{V}. \tag{1.3}$$

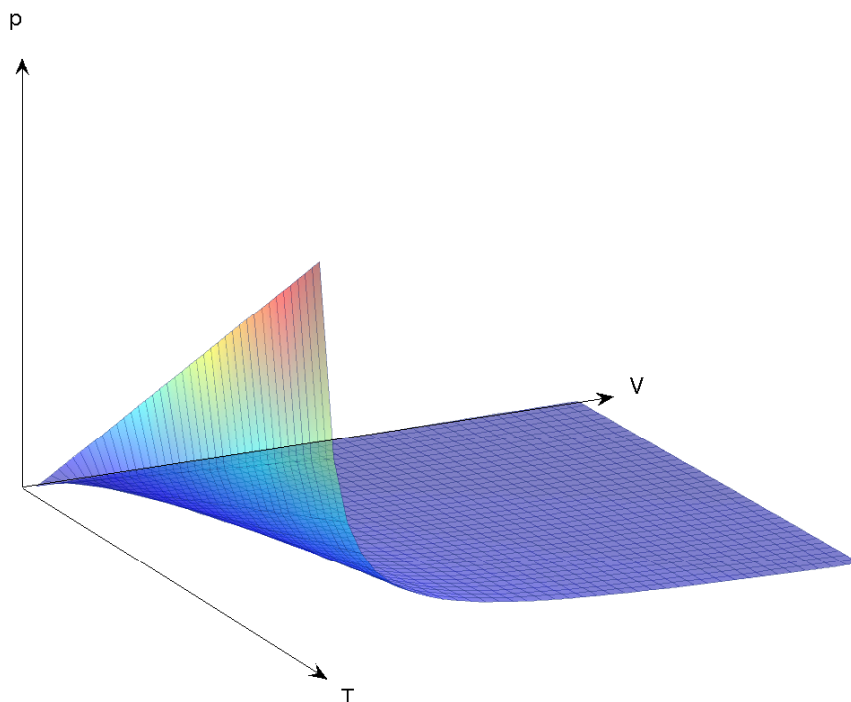
Ici, le domaine "naturel" de la fonction f est $D(f) = \{(T, V) \in \mathbb{R}^2 \mid T > 0 \text{ et } V > 0\}$ car la température absolue et la pression sont des quantités positives.

Le **graphe** d'une fonction $f : D(f) \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ est l'ensemble

$$G(f) = \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid (x_1, x_2) \in D(f) \text{ et } x_3 = f(x_1, x_2)\}.$$

Cette surface dans \mathbb{R}^3 nous donne une représentation géométrique de la fonction f .

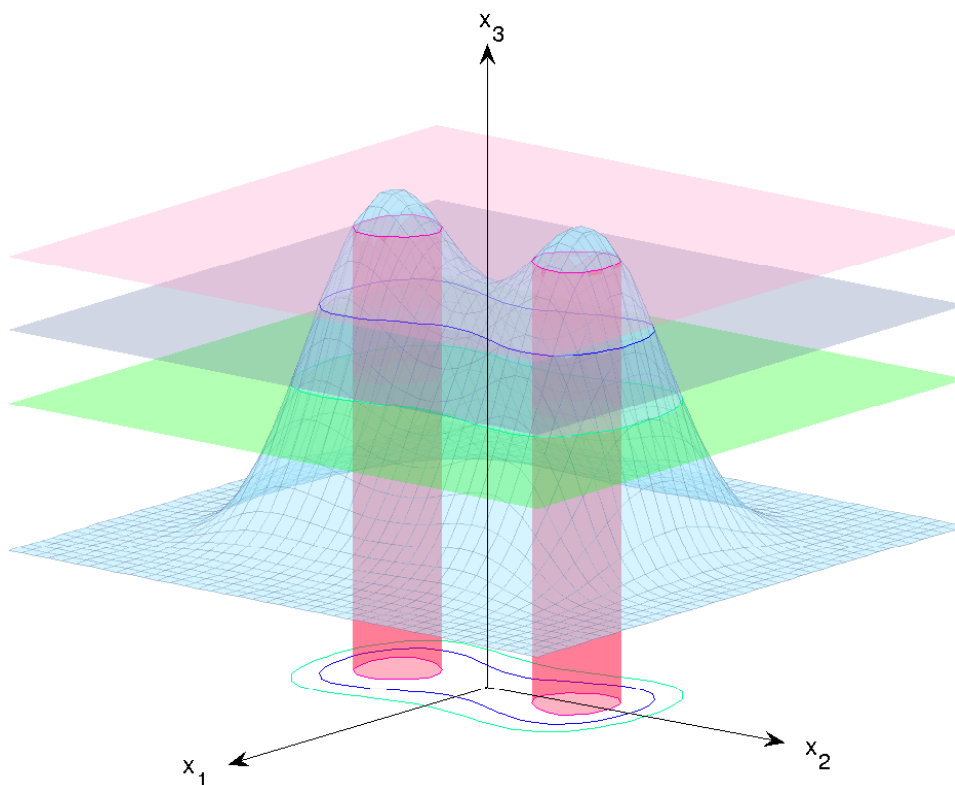
La figure suivante représente le graphe de la fonction $f(T, V) = RT/V$ (en fait seulement une partie de ce graphe!).



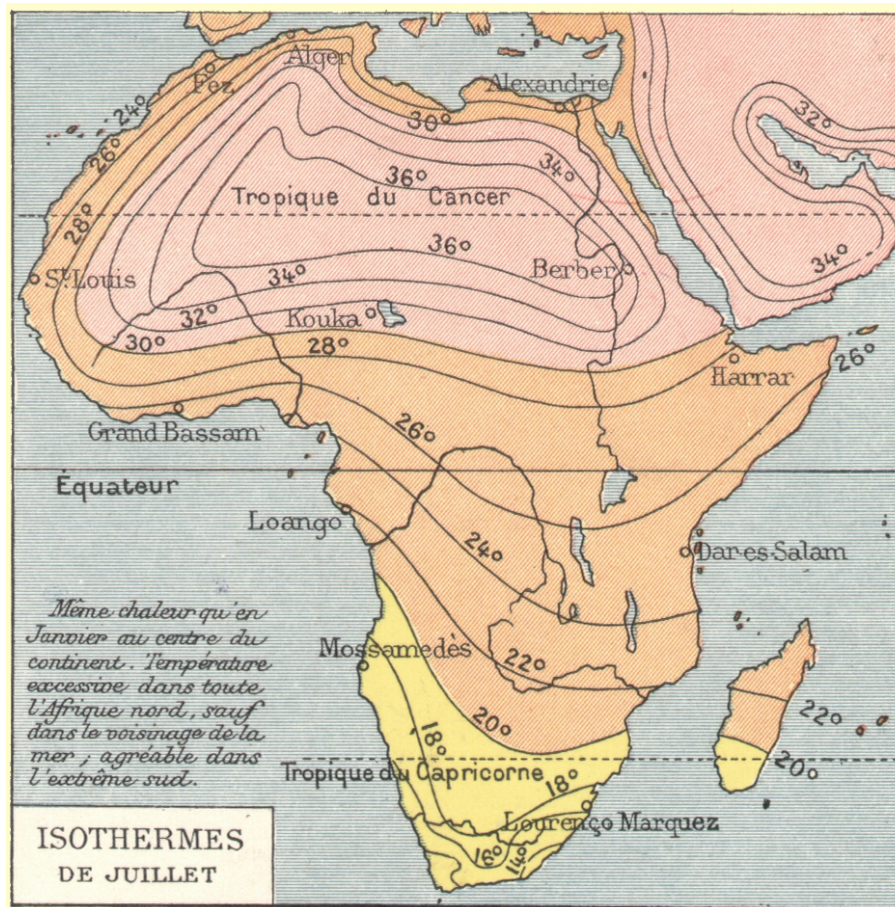
Une autre façon de représenter une fonction de 2 variable est de tracer ses **courbes de niveau**

$$N_f(\alpha) = \{(x_1, x_2) \in \mathbb{R}^2 \mid f(x_1, x_2) = \alpha\}.$$

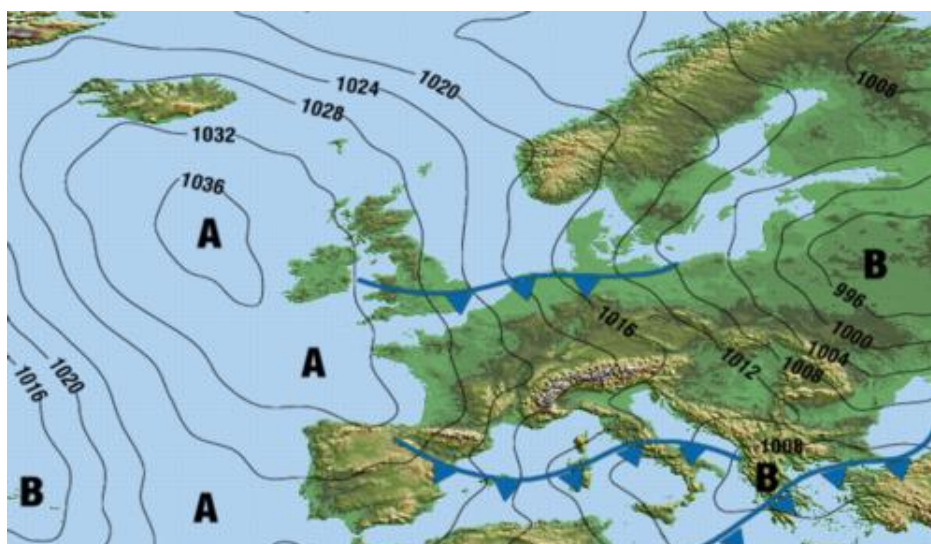
Géométriquement, la courbe de niveau $N_f(\alpha)$ est la projection sur le plan x_1, x_2 de l'intersection du graphe $G(f)$ avec le plan d'équation $x_3 = \alpha$, comme suggéré dans la figure suivante.



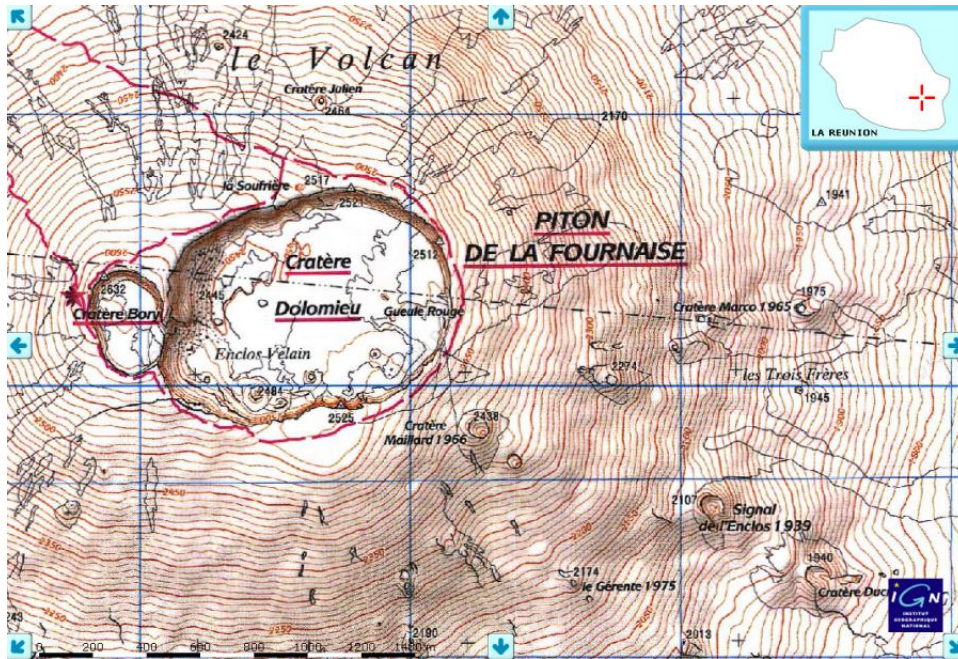
Exemples : si f est la température en fonction de la position ses courbes de niveaux sont les isothermes.



Si f détermine la pression, ses courbes de niveau sont les isobares.



Finalement, si f représente la hauteur d'un point de la surface terrestre, ses courbes de niveau sont celles apparaissant sur les cartes topographiques.



S'il est plus difficile de représenter graphiquement des fonctions de plus de 2 variables, il est facile de définir un champ scalaire sur un domaine $\Omega \subset \mathbb{R}^n$ avec $n > 2$. Comme nous le verrons, le passage d'une variable $f(x)$ à deux variables $f(x_1, x_2)$ requiert des concepts nouveaux alors que le passage de deux variables à n variables n'est souvent qu'une question de notation.

1.2.2.1 Dérivées partielles

Définition 1.21 Soit $f : D(f) \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ une fonction de deux variables. On dit que f admet en $x = (a, b) \in D(f)$ une *dérivée partielle* par rapport à la première variable si la limite

$$\frac{\partial f}{\partial x_1}(a, b) = \lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h}, \tag{1.4}$$

existe.

On remarquera que la limite (1.4) n'est autre que la dérivée en a , au sens usuel du terme, de la fonction d'une seule variable $x \mapsto f(x, b)$, la seconde variable restant fixée à sa valeur b . On définit de façon analogue la dérivée partielle de f par rapport à sa seconde variable. De nombreuses notations usuelles existent pour ces dérivées partielles, on écrit par exemple

$$\frac{\partial f}{\partial x_1}(a, b) = \left. \frac{\partial f}{\partial x_1} \right|_{x=(a,b)} = \left(\frac{\partial}{\partial x_1} f \right) (a, b) = (\partial_{x_1} f)(a, b) = f_{x_1}(a, b) = f_1(a, b),$$

pour la dérivée partielle par rapport à la première variable. Les dérivées partielles d'ordre supérieur sont définies par induction. Par exemple, si (1.4) existe, sa dérivée partielle par rapport à la seconde variable est définie par

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) = \left. \frac{\partial}{\partial x_2} \frac{\partial f}{\partial x_1} \right|_{x=(a,b)} = \lim_{k \rightarrow 0} \frac{1}{k} \left(\frac{\partial f}{\partial x_1}(a, b+k) - \frac{\partial f}{\partial x_1}(a, b) \right).$$

On note qu'en général l'ordre dans lequel on effectue les dérivées est important

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) \neq \frac{\partial^2 f}{\partial x_1 \partial x_2}(a, b),$$

voir cependant le théorème 1.32 ci-dessous à ce sujet. Les notations pour les dérivées d'ordre supérieur sont nombreuses, il faut en particulier bien noter comment elles encodent l'ordre des dérivées,

$$\begin{aligned}\frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b) &= (\partial_{x_2} \partial_{x_1} f)(a, b) = f_{x_1 x_2}(a, b) = f_{12}(a, b), \\ \frac{\partial^3 f}{\partial x_2 \partial x_1^2}(a, b) &= (\partial_{x_1} \partial_{x_1}^2 f)(a, b) = f_{x_1 x_1 x_2}(a, b) = f_{112}(a, b).\end{aligned}$$

Exemple 1.17 La fonction

$$f(x, y) = \begin{cases} 0 & \text{si } (x, y) = (0, 0); \\ \frac{xy^3}{x^2 + y^2} & \text{sinon,} \end{cases}$$

admet des dérivées partielles en tout point de \mathbb{R}^2 . Pour $(x, y) \neq (0, 0)$ on a

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

et

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

Pour calculer les dérivées partielles en $(0, 0)$, on utilise la définition

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0,$$

et

$$\frac{\partial f}{\partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{f(0, h) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 - 0}{h} = 0.$$

Nous pouvons maintenant calculer les dérivées partielles mixtes du second ordre en $(0, 0)$ comme suit

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial x}(0, h) - \frac{\partial f}{\partial x}(0, 0) \right) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{h^5}{h^4} - 0 \right) = 1,$$

alors que

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{h \rightarrow 0} \frac{1}{h} \left(\frac{\partial f}{\partial y}(h, 0) - \frac{\partial f}{\partial y}(0, 0) \right) = \lim_{h \rightarrow 0} \frac{1}{h} (0 - 0) = 0.$$

◁

Exemple 1.18 Les dérivées partielles de la fonction $f(T, V) = RT/V$ sont

$$\frac{\partial f}{\partial T}(T, V) = \frac{R}{V}, \quad \frac{\partial f}{\partial V}(T, V) = -\frac{RT}{V^2}.$$

Ses 4 dérivées secondes sont


$$\frac{\partial^2 f}{\partial T^2}(T, V) = 0, \quad \frac{\partial^2 f}{\partial V \partial T}(T, V) = -\frac{R}{V^2} = \frac{\partial^2 f}{\partial T \partial V}(T, V), \quad \frac{\partial^2 f}{\partial V^2}(T, V) = 2\frac{RT}{V^3},$$

et ses 8 dérivées d'ordre 3 sont

$$\begin{aligned}\frac{\partial^3 f}{\partial T^3}(T, V) &= \frac{\partial^3 f}{\partial V \partial T^2}(T, V) = \frac{\partial^3 f}{\partial T \partial V \partial T}(T, V) = \frac{\partial^3 f}{\partial T^2 \partial V}(T, V) = 0, \\ \frac{\partial^3 f}{\partial V^3}(T, V) &= -6\frac{RT}{V^4}, \quad \frac{\partial^3 f}{\partial T \partial V^2}(T, V) = \frac{\partial^3 f}{\partial V \partial T \partial V}(T, V) = \frac{\partial^3 f}{\partial V^2 \partial T}(T, V) = 2\frac{R}{V^3}.\end{aligned}$$

◁

On se souviendra qu'une fonction d'une variable différentiable en un point est continue en ce point.



Si f admet des dérivées partielles en (a, b) , elle n'est pas nécessairement continue en ce point.

Exemple 1.19 La fonction f de l'exemple 1.13 n'est pas continue en $(0, 0)$. Cependant, il suit du fait que $f(x, 0) = f(0, y) = 0$ qu'elle admet en ce point des dérivées partielles

$$\frac{\partial f}{\partial x}(0, 0) = 0, \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

◁

Exemple 1.20 Le lagrangien d'un point matériel de masse m se mouvant sur une droite sous l'influence d'une force conservatrice de potentiel V est la différence entre son énergie cinétique et son énergie potentielle. C'est une fonction de deux variables : sa position q et sa vitesse \dot{q} ,

$$L(q, \dot{q}) = \frac{1}{2}m\dot{q}^2 - V(q).$$

Le mouvement du point matériel est décrit par l'équation d'Euler-Lagrange

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) \right) = \frac{\partial L}{\partial q}(q, \dot{q}). \quad (1.5)$$

Les dérivées partielles du lagrangien sont

$$\frac{\partial L}{\partial \dot{q}}(q, \dot{q}) = m\dot{q}, \quad \frac{\partial L}{\partial q}(q, \dot{q}) = V'(q).$$

Sous l'équation (1.5) se cache donc l'équation de Newton

$$m\ddot{q} = -V'(q).$$

◁

Exemple 1.21 L'entropie thermodynamique d'un gaz est une fonction du nombre N de molécules qui le compose, de l'énergie cinétique totale U de ces molécules et du volume V occupé par ce gaz. Les forces électromagnétiques agissant entre les molécules étant de courte portée, il est raisonnable de les négliger lorsque la distance moyenne entre les molécules est grande, c'est-à-dire lorsque le gaz est suffisamment dilué. En mécanique statistique on démontre que dans cette approximation l'entropie est

$$S(U, V, N) = k_B N \left(\log \left(\frac{V}{N} \right) + \frac{3}{2} \log \left(\frac{U}{N} \right) \right),$$

où k_B est la constante de Boltzmann, une constante fondamentale qui fixe l'échelle des températures absolues. Il est d'usage de mesurer cette température en degrés Kelvin (K°), et dans ce cas $k_B = 1.3806503 \cdot 10^{-23}$ Joule / Kelvin. Il découle des principes fondamentaux de la thermodynamique que la température T et la pression p du gaz peuvent s'exprimer à l'aide des dérivées partielles de l'entropie

$$\frac{1}{T} = \frac{\partial S}{\partial U}(U, V, N), \quad p = \frac{\partial S}{\partial V}(U, V, N).$$

En calculant ces dérivées partielles on obtient les relations

$$\frac{1}{T} = k_B \frac{3}{2} \frac{N}{U}, \quad p = k_B \frac{N}{V}. \quad (1.6)$$

En notant que le nombre de molécules N est égal au produit du nombre de mole n par le nombre d'Avogadro N_A on obtient, à partir de la seconde relation (1.6), l'équation d'état

$$pV = n(k_B N_A)T.$$

En comparant cette dernière avec la loi d'Avogadro (voir (1.3) où nous avons supposé $n = 1$) on obtient la relation $R = k_B N_A$ entre la constante phénoménologique des gaz parfait R et les deux constantes fondamentales k_B et N_A . La première relation (1.6) montre que la température du gaz est une mesure de l'énergie cinétique moyenne

$$u = \frac{U}{N} = \frac{3}{2} k_B T,$$

des molécules qui le compose. ◁

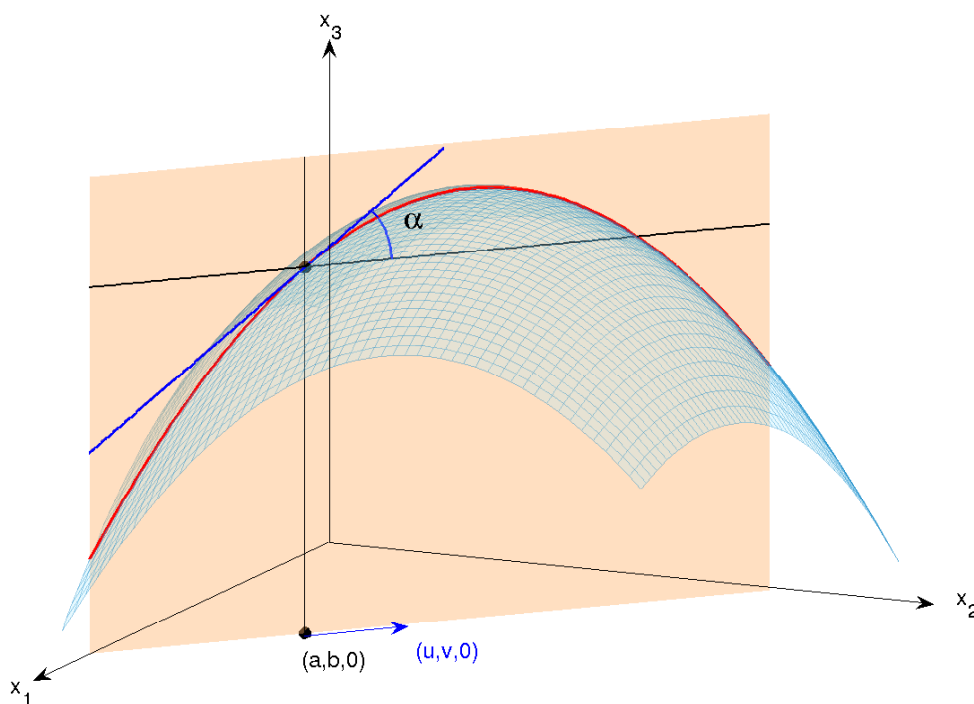
1.2.2.2 Dérivées directionnelles

Les dérivées partielles d'une fonction de deux variables décrivent le taux de variation de cette fonction lorsqu'une de ses variable varie, l'autre restant constante. De manière plus générale, on peut s'intéresser au taux de variation de f lorsque ses arguments (x_1, x_2) varient dans une direction fixée.

Définition 1.22 Soit $(u, v) \in \mathbb{R}^2$ un vecteur unité ($\|(u, v)\| = 1$). La *dérivée directionnelle* de f en $(a, b) \in D(f)$ selon la direction (u, v) est définie par

$$(D_{(u,v)} f)(a, b) = \left. \frac{d}{dt} f(a + tu, b + tv) \right|_{t=0} = \lim_{t \rightarrow 0} \frac{f(a + tu, b + tv) - f(a, b)}{t}. \quad (1.7)$$

Pour interpréter géométriquement cette quantité, on note que la courbe définie par la formule $t \mapsto x(t) = (a + tu, b + tv, f(a + tu, b + tv))$ est l'intersection du graphe de f avec le plan vertical engendré par les deux vecteurs $(u, v, 0)$, $(0, 0, 1)$ et passant par le point $(a, b, 0)$.



L'existence de la dérivée directionnelle (1.7) est donc équivalente à l'existence d'une tangente à cette courbe en $(a, b, f(a, b))$. La valeur de cette dérivée est la pente de la tangente (voir l'illustration ci-dessus)

$$(D_{(u,v)} f)(a, b) = \tan \alpha.$$

On dira aussi que c'est la *pente du graphe* de f en (a, b) dans la direction (u, v) .

On remarquera finalement les deux cas particuliers

$$(D_{(1,0)} f)(a, b) = \frac{\partial f}{\partial x_1}(a, b), \quad (D_{(0,1)} f)(a, b) = \frac{\partial f}{\partial x_2}(a, b).$$

Exemple 1.22 La dérivée directionnelle de la fonction $f(T, V) = RT/V$ en (T, V) selon la direction (u, v) est

$$(D_{(u,v)}f)(T, V) = \left. \frac{d}{dt} \frac{R(T + tu)}{V + tv} \right|_{t=0} = \frac{R}{V}u - \frac{RT}{V^2}v.$$

◁

Exemple 1.23 Soit f la fonction définie par

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

Ses dérivées directionnelles en $(0, 0)$ sont

$$\begin{aligned} (D_{(u,v)}f)(0, 0) &= \lim_{t \rightarrow 0} \frac{f(tu, tv) - f(0, 0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{(tu)(tv)^2}{t((tu)^2 + (tv)^2)} \\ &= \frac{uv^2}{u^2 + v^2} = f(u, v). \end{aligned}$$

La fonction f admet donc en $(0, 0)$ des dérivées dans toutes les directions (u, v) .

◁

1.2.2.3 Différentielle

Définition 1.23 La fonction f est *différentiable* en un point (a, b) intérieur à $D(f)$ si il existe des constantes A et B telles que

$$\lim_{(h,k) \rightarrow (0,0)} \frac{f(a+h, b+k) - [f(a, b) + Ah + Bk]}{\|(h, k)\|} = 0. \tag{1.8}$$

f est *différentiable sur un ouvert* $O \subset D(f)$ si elle est *différentiable en chaque point* $(a, b) \in O$.

Nous montrerons un peu plus bas que si f est différentiable en (a, b) les constantes A et B dans l'expression (1.8) sont déterminées de manière unique, ce qui justifie la définition suivante.

Définition 1.24 Si f est différentiable en (a, b) , la *différentielle* de f en (a, b) est l'application linéaire

$$\begin{aligned} df_{(a,b)} &: \mathbb{R}^2 &\rightarrow &\mathbb{R} \\ (u, v) &\mapsto &Au + Bv, \end{aligned}$$

où les constantes A et B sont déterminées par (1.8).

La différentielle de f en (a, b) est également appelée *application linéaire tangente* à f en (a, b) ou simplement *dérivée* de f en (a, b) .

Exemple 1.24 $f(x, y) = xy$ est différentiable sur \mathbb{R}^2 . Pour le vérifier, notons que pour tout point $(a, b) \in \mathbb{R}^2$ on a $f(a+h, b+k) - f(a, b) = ak + bh + hk$. En posant $A = b$ et $B = a$ on obtient, à l'aide de l'inégalité 1.2,

$$\frac{f(a+h, b+k) - [f(a, b) + Ah + Bk]}{\|(h, k)\|} = \frac{hk}{\sqrt{h^2 + k^2}} \leq \frac{1}{2} \sqrt{h^2 + k^2},$$

d'où l'on déduit facilement (1.8). La différentielle de f en (a, b) est donc donnée par

$$df_{(a,b)}(u, v) = bu + av.$$

◁

Théorème 1.25 Si f est différentiable en (a, b) , elle est continue en ce point.

Démonstration. Comme $Ah + Bk$ et $\|(h, k)\|$ tendent tous deux vers 0 lorsque $(h, k) \rightarrow (0, 0)$, il suit immédiatement de (1.8) que

$$\lim_{(h,k) \rightarrow (0,0)} f(a+h, b+k) = f(a, b).$$

□

Lemme 1.26 Si $t \mapsto x(t) \in \mathbb{R}^2$ est une courbe différentiable et si f est différentiable en $x(t_0)$ alors $f(x(t))$ est différentiable en $t = t_0$ et

$$\left. \frac{df(x(t))}{dt} \right|_{t=t_0} = df_{x(t_0)}(x'(t_0))$$

Démonstration. Soient $x(t) = (x_1(t), x_2(t))$ une courbe différentiable et f différentiable en $x(t_0)$. On peut écrire

$$\frac{f(x(t)) - f(x(t_0))}{t - t_0} = F(t) + G(t)H(t),$$

où

$$F(t) = A \frac{x_1(t) - x_1(t_0)}{t - t_0} + B \frac{x_2(t) - x_2(t_0)}{t - t_0},$$

$$G(t) = \frac{f(x(t)) - [f(x(t_0)) + (A(x_1(t) - x_1(t_0)) + B(x_2(t) - x_2(t_0)))]}{\|x(t) - x(t_0)\|},$$

$$H(t) = \frac{\|x(t) - x(t_0)\|}{t - t_0}.$$

Lorsque $t \rightarrow t_0$ on a $F(t) \rightarrow Ax'_1(t_0) + Bx'_2(t_0)$ et

$$|H(t)| = \left\| \frac{x(t) - x(t_0)}{t - t_0} \right\| \rightarrow \|x'(t_0)\|,$$

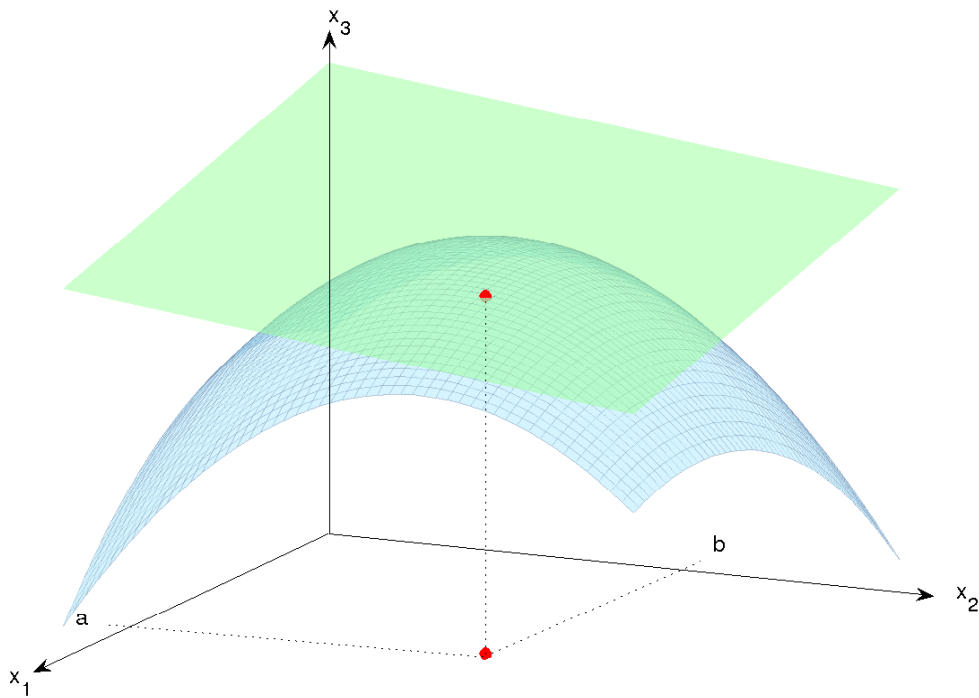
et donc $H(t)$ reste borné alors que $G(t) \rightarrow 0$ par (1.8). On obtient bien

$$\left. \frac{df(x(t))}{dt} \right|_{t=t_0} = Ax'_1(t_0) + Bx'_2(t_0). \tag{1.9}$$

□

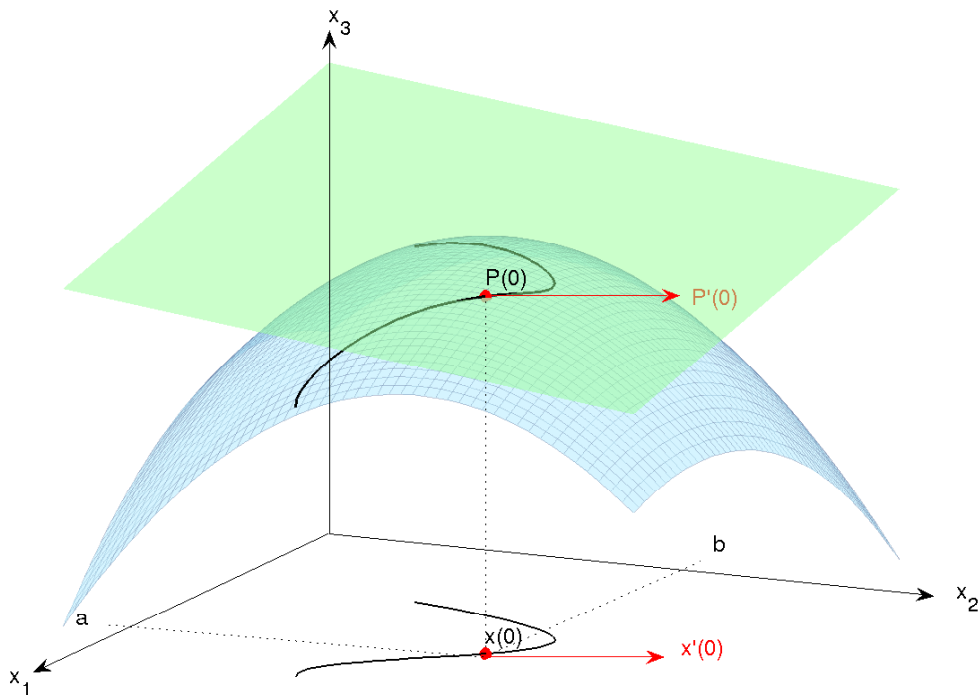
1.2.2.4 Plan tangent

La dérivée d'une fonction d'une variable détermine la droite tangente à son graphe. Qu'en est-il des fonctions de 2 variables? Qu'est-ce que le plan tangent au graphe de f en (a, b) ?



Pour mieux comprendre sa nature, considérons la courbe différentiable $t \mapsto x(t) = (x_1(t), x_2(t))$ dans le plan x_1x_2 passant par (a, b) . Sans restreindre la généralité, nous pouvons supposer que $x(0) = (x_1(0), x_2(0)) = (a, b)$. Alors $t \mapsto P(t) = (x_1(t), x_2(t), x_3(t))$ où $x_3(t) = f(x_1(t), x_2(t))$ est une courbe sur le graphe de f . Si f est différentiable en (a, b) , le lemme 1.26 montre que cette courbe est elle aussi différentiable en $t = 0$. Son vecteur vitesse en $P(0)$ est $P'(0) = (x'_1(0), x'_2(0), x'_3(0))$ où

$$x'_3(0) = df_{(a,b)}(x'_1(0), x'_2(0)) = Ax'_1(0) + Bx'_2(0). \quad (1.10)$$



La courbe $P(t)$ étant sur le graphe de f , $P'(0)$ doit être tangent à ce graphe en $P(0)$. En fait, le **plan tangent** au graphe de f en $(a, b, f(a, b))$, si il existe, est engendré par les vecteurs vitesse de toutes les courbes différentiables sur le graphe de f passant par ce point.

Soit

$$x_3 = f(a, b) + \alpha(x_1 - a) + \beta(x_2 - b),$$

l'équation du plan tangent (ceci revient à dire que $\mathbf{n} = (-\alpha, -\beta, 1)$ est un vecteur normal à ce plan). Nous voulons déterminer ses *coefficients directeurs* α et β . Commençons par remarquer que le vecteur vitesse $P'(0)$ appartient au plan tangent si et seulement si $P'(0) \perp \mathbf{n}$,

$$\mathbf{n} \cdot P'(0) = -\alpha x_1'(0) - \beta x_2'(0) + x_3'(0) = 0,$$

c'est-à-dire

$$x_3'(0) = \alpha x_1'(0) + \beta x_2'(0).$$

En comparant cette expression avec (1.10) on conclut que les coefficients directeurs α et β ne sont autres que les coefficients A et B de la différentielle $df_{(a,b)}$. Pour déterminer ces coefficients il suffit de considérer une classe particulière de courbes : $t \mapsto \mathbf{x}(t) = (a + tu, b + tv)$ (des segments de droites!). Dans ce cas $x_1'(0) = u$, $x_2'(0) = v$ et

$$x_3'(0) = \left. \frac{d}{dt} f(a + tu, b + tv) \right|_{t=0} = (D_{(u,v)} f)(a, b),$$

est la dérivée directionnelle de f en (a, b) selon (u, v) . Nous avons ainsi montré que si f est différentiable en (a, b) alors ses dérivées directionnelles en (a, b) existent et sont déterminées par la différentielle

$$(D_{(u,v)} f)(a, b) = df_{(a,b)}(u, v).$$

En considérant $(u, v) = (1, 0)$, nous obtenons

$$A = df_{(a,b)}(1, 0) = (D_{(1,0)} f)(a, b) = \frac{\partial f}{\partial x_1}(a, b),$$

et $(u, v) = (0, 1)$ mène à

$$B = df_{(a,b)}(0, 1) = (D_{(0,1)} f)(a, b) = \frac{\partial f}{\partial x_2}(a, b).$$

Nous avons donc démontré le théorème suivant :

Théorème 1.27 Soit f une fonction différentiable en (a, b) .

i. f admet des dérivées partielles en (a, b) et sa différentielle est donnée par

$$df_{(a,b)}(u, v) = u \frac{\partial f}{\partial x_1}(a, b) + v \frac{\partial f}{\partial x_2}(a, b). \quad (1.11)$$

ii. Pour tout vecteur unité $(u, v) \in \mathbb{R}^2$ la dérivée directionnelle $(D_{(u,v)} f)(a, b)$ existe et est une fonction linéaire de (u, v) donnée par la formule

$$(D_{(u,v)} f)(a, b) = df_{(a,b)}(u, v). \quad (1.12)$$

iii. Le graphe de f admet en $(a, b, f(a, b))$ un plan tangent d'équation

$$x_3 = f(a, b) + (x_1 - a) \frac{\partial f}{\partial x_1}(a, b) + (x_2 - b) \frac{\partial f}{\partial x_2}(a, b). \quad (1.13)$$



L'existence des dérivées partielles et directionnelles de f est une condition nécessaire pour sa différentiabilité. Cette condition n'est cependant pas suffisante !

Exemple 1.25 La fonction de l'exemple 1.23 admet en $(0, 0)$ des dérivées directionnelles, cependant $(D_{(u,v)} f)(0, 0) = f(u, v)$ n'est pas une fonction linéaire de (u, v) ! f n'est donc pas différentiable en $(0, 0)$. f n'a pas de différentielle en (a, b) et son graphe n'admet de pas de plan tangent en $(0, 0, 0)$. \triangleleft

Approximation linéaire. On peut reformuler la condition (1.8) de la définition 1.23 de la manière suivante : f est différentiable en (a, b) si et seulement si il existe une fonction $\varepsilon(h, k)$ telle que

$$f(a + h, b + k) = f(a, b) + df_{(a,b)}(h, k) + \varepsilon(h, k) \|(h, k)\|, \tag{1.14}$$

et

$$\lim_{(h,k) \rightarrow (0,0)} \varepsilon(h, k) = 0. \tag{1.15}$$

En interprétant le terme $f(a, b) + df_{(a,b)}(h, k)$ comme une approximation de $f(a + h, b + k)$, la condition (1.15) exige que l'erreur $\varepsilon(h, k) \|(h, k)\|$ décroisse plus vite que $\|(h, k)\|$ lorsque $(h, k) \rightarrow 0$. On peut donc utiliser l'expression

$$f(x_1, x_2) \approx f(a, b) + df_{(a,b)}(x_1 - a, x_2 - b) = f(a, b) + (x_1 - a) \frac{\partial f}{\partial x_1}(a, b) + (x_2 - b) \frac{\partial f}{\partial x_2}(a, b),$$

comme une approximation de f pour des valeurs de (x_1, x_2) assez proches de (a, b) .

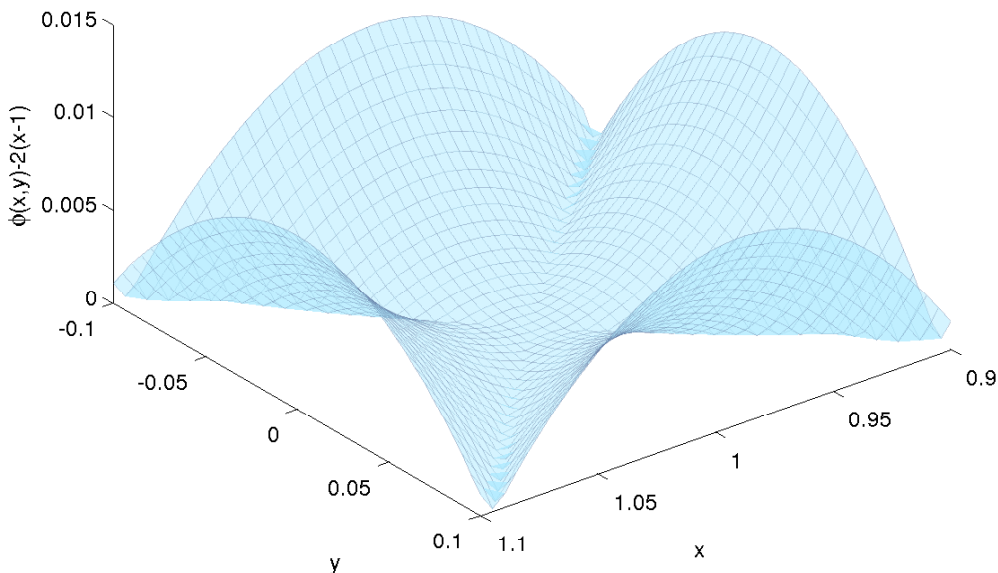
Exemple 1.26 En dimension 2, le potentiel électrostatique d'une charge ponctuelle située à l'origine $(0, 0)$ du plan est donné par $\phi(x, y) = \log(x^2 + y^2)$. Pour évaluer ce potentiel au voisinage du point $(1, 0)$ on peut faire l'approximation $\phi(x, y) \approx \phi(1, 0) + (x - 1)\phi_x(1, 0) + y\phi_y(1, 0)$. Comme

$$\phi_x(x, y) = \frac{2x}{x^2 + y^2}, \quad \phi_y(x, y) = \frac{2y}{x^2 + y^2},$$

on a

$$\phi(1, 0) = \log(1) = 0, \quad \phi_x(1, 0) = 2, \quad \phi_y(1, 0) = 0,$$

et donc $\phi(x, y) \approx 2(x - 1)$. La figure suivante montre l'erreur d'approximation $\phi(x, y) - 2(x - 1)$ dans un voisinage de taille 0.1 du point de référence $(1, 0)$.



◁

Exemple 1.27 Soient $P = (a, b, f(a, b))$, $Q = (a', b', f(a', b'))$ et $R = (a'', b'', f(a'', b''))$ trois points *non alignés* sur le graphe de f , $p' = (h', k') = (a' - a, b' - b)$, $p'' = (h'', k'') = (a'' - a, b'' - b)$. Il est aisé de montrer (faites le!) que l'équation du plan déterminé par ces trois points est

$$x_3 = f(a, b) + \alpha(x_1 - a) + \beta(x_2 - b),$$

où les coefficients directeurs sont donnés par

$$\alpha = \frac{k''(f(a', b') - f(a, b)) - k'(f(a'', b'') - f(a, b))}{h'k'' - h''k'},$$

$$\beta = \frac{h'(f(a'', b'') - f(a, b)) - h''(f(a', b') - f(a, b))}{h'k'' - h''k'}.$$

En utilisant l'expression (1.14) on peut donc écrire

$$f(a', b') - f(a, b) = h' \frac{\partial f}{\partial x_1}(a, b) + k' \frac{\partial f}{\partial x_2}(a, b) + \varepsilon' \|p'\|,$$

$$f(a'', b'') - f(a, b) = h'' \frac{\partial f}{\partial x_1}(a, b) + k'' \frac{\partial f}{\partial x_2}(a, b) + \varepsilon'' \|p''\|,$$

et en déduire que

$$\alpha = \frac{\partial f}{\partial x_1}(a, b) + \xi, \quad \beta = \frac{\partial f}{\partial x_2}(a, b) + \eta,$$

où

$$\xi = \frac{k''\varepsilon' \|p'\| - k'\varepsilon'' \|p''\|}{h'k'' - h''k'}, \quad \eta = \frac{h'\varepsilon'' \|p''\| - h''\varepsilon' \|p'\|}{h'k'' - h''k'}.$$

On remarque maintenant que le module $\Delta = |h'k'' - h''k'|$ du dénominateur de ces expressions n'est autre que la norme $\|P' \times P''\|$ du produit vectoriel des deux vecteurs

$$P' = (h', k', 0), \quad P'' = (h'', k'', 0),$$

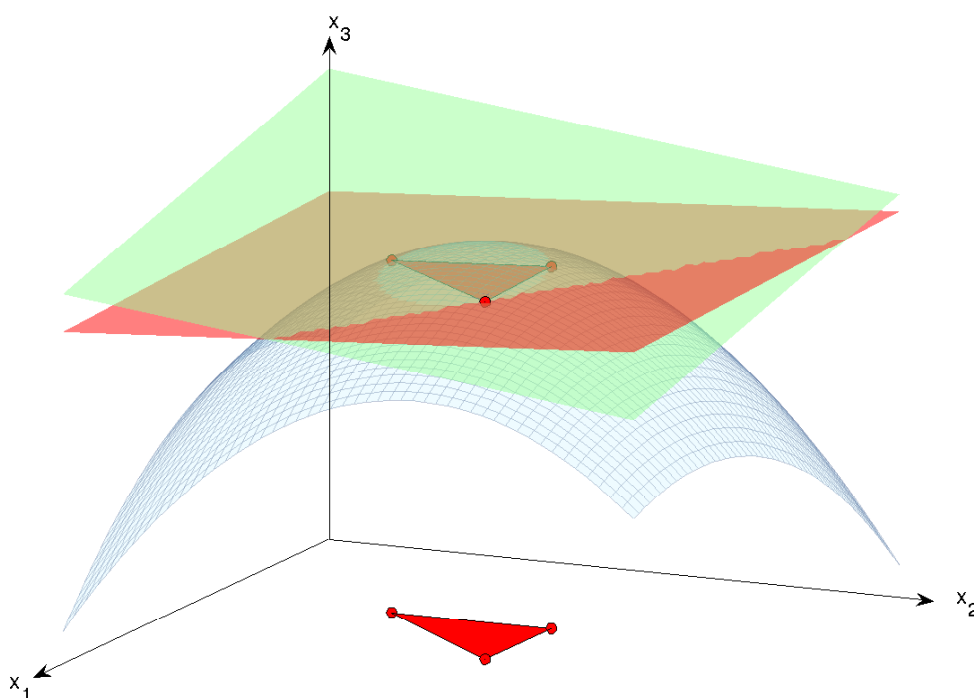
(vérifiez le!). Comme $\|P' \times P''\| = \|P'\| \|P''\| \sin \varphi = \|p'\| \|p''\| \sin \varphi$, φ étant l'angle entre p' et p'' , et $|k'| \leq \|p'\|$, $|k''| \leq \|p''\|$ on en déduit que

$$\|\xi\| \leq \frac{|k''|}{\|p''\| |\sin \varphi|} |\varepsilon'| + \frac{|k'|}{\|p'\| |\sin \varphi|} |\varepsilon''| \leq \frac{|\varepsilon'| + |\varepsilon''|}{|\sin \varphi|}.$$

On montre de manière analogue que $\|\eta\|$ satisfait la même inégalité. L'estimation d'erreur (1.15) montre si $p' \rightarrow (0, 0)$ et $p'' \rightarrow (0, 0)$ de façon telle que $|\sin \varphi| \geq c > 0$ pour une constante c (c'est-à-dire que les vecteurs p' et p'' restent linéairement indépendants!), alors $\xi \rightarrow 0$ et $\eta \rightarrow 0$. On interprète géométriquement le fait que

$$\alpha \rightarrow \frac{\partial f}{\partial x_1}(a, b) \quad \text{et} \quad \beta \rightarrow \frac{\partial f}{\partial x_2}(a, b),$$

en disant que, lorsque Q et R s'approchent de P le plan contenant P, Q et R s'approche du plan tangent au graphe de f en P .



1.2.2.5 Règles de différentiation

Les règles suivantes généralisent les règles usuelles concernant le calcul des dérivées d'une fonction d'une variable. Leurs démonstration constitue un bon exercice que je vous conseille fortement !

1. Si f et g sont différentiable en (a, b) et $\alpha, \beta \in \mathbb{R}$ alors $\alpha f + \beta g$ est aussi différentiable et

$$d(\alpha f + \beta g)_{(a,b)} = \alpha df_{(a,b)} + \beta dg_{(a,b)}.$$

2. Si f et g sont différentiables en (a, b) alors fg l'est également et

$$d(fg)_{(a,b)} = f(a, b)dg_{(a,b)} + g(a, b)df_{(a,b)}.$$

3. Si f et g sont différentiables en (a, b) et si $g(a, b) \neq 0$ alors f/g est différentiable en (a, b) et

$$d(f/g)_{(a,b)} = \frac{g(a, b)df_{(a,b)} - f(a, b)dg_{(a,b)}}{g(a, b)^2}.$$

4. Si f est différentiable en (a, b) et F est un fonction d'une variable différentiable en $f(a, b)$ alors $F \circ f$ est différentiable en (a, b) et

$$d(F \circ f)_{(a,b)} = (F' \circ f(a, b))df_{(a,b)}$$

5. Si $f(x_1, x_2) = F(x_1)$ et si F est différentiable en a , alors f est différentiable en (a, b) pour tout $b \in \mathbb{R}$ et $df_{(a,b)}(u, v) = F'(a)u$ pour tout $(u, v) \in \mathbb{R}^2$.

6. Si $t \mapsto x(t) \in \mathbb{R}^2$ est différentiable en t_0 et si f est différentiable en $x(t_0)$ alors $f \circ x$ est différentiable en t_0 et $(f \circ x)'(t_0) = df_{x(t_0)}(x'(t_0))$.

1.2.2.6 Fonctions de classe C^n

Définition 1.28 Une fonction f est de de classe C^n sur l'ouvert $O \subset D(f) \subset \mathbb{R}^2$ si toutes ses dérivées partielles

$$\frac{\partial^{n_1+n_2} f}{\partial x_1^{n_1} \partial x_2^{n_2}}(x_1, x_2) \quad \text{avec} \quad n_1 + n_2 \leq n,$$

existent et sont continues en chaque point $(x_1, x_2) \in O$. f est de classe C^∞ sur O si elle y est de classe C^n pour tout entier n .

Exemple 1.28 La fonction $f(x, y) = \sqrt{x^2 - y^2}$ est définie sur le domaine

$$D(f) = \{(x, y) \in \mathbb{R}^2 \mid -|x| \leq y \leq |x|\}.$$

Elle admet, en chaque point de l'ouvert

$$O = \{(x, y) \in \mathbb{R}^2 \mid -|x| < y < |x|\},$$

les dérivées partielles

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

Ces deux fonctions sont continues en chaque point de O . La fonction f est donc de classe C^1 sur O . Les dérivées secondes de f sont

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

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

et on en déduit que f est de classe C^2 sur O . En fait il est aisé de se convaincre qu'elle est de classe C^∞ sur O .

Comme le montre cet exemple, il est souvent plus facile de montrer qu'une fonction admet des dérivées partielles continues que de vérifier explicitement la condition (1.8). Le résultat suivant est utile dans beaucoup d'applications.

Théorème 1.29 *Toute fonction de classe C^1 sur l'ouvert O est différentiable sur O .*

Pour démontrer ce théorème, nous utiliserons le lemme suivant qui est une conséquence immédiate du théorème des accroissements finis pour les fonctions d'une variable :

Lemme 1.30 *Supposons que f admette des dérivées partielles en chaque point de la boule ouverte $U_\delta(x) = \{y \in \mathbb{R}^2 \mid \|x - y\| < \delta\}$. Alors, pour tout $y \in U_\delta(x)$ il existe $\xi, \xi' \in U_\delta(x)$ tels $\|\xi - x\| \leq \|y - x\|$, $\|\xi' - x\| \leq \|y - x\|$ et*

$$f(y_1, y_2) - f(x_1, x_2) = (y_1 - x_1) \frac{\partial f}{\partial x_1}(\xi) + (y_2 - x_2) \frac{\partial f}{\partial x_2}(\xi').$$

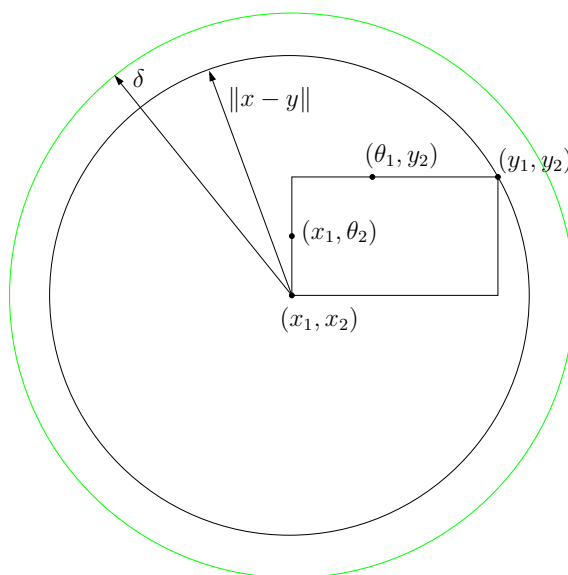
Démonstration. On peut écrire

$$f(y_1, y_2) - f(x_1, x_2) = (f(y_1, y_2) - f(x_1, y_2)) + (f(x_1, y_2) - f(x_1, x_2)),$$

et remarquer que, par le théorème des accroissements finis appliqué à la fonction $x \mapsto f(x, y_2)$

$$f(y_1, y_2) - f(x_1, y_2) = (y_1 - x_1) \frac{\partial f}{\partial x_1}(\theta_1, y_2),$$

pour un θ_1 entre x_1 et y_1 .



De même on peut trouver θ_2 entre x_2 et y_2 tel que

$$f(x_1, y_2) - f(x_1, x_2) = (y_2 - x_2) \frac{\partial f}{\partial x_2}(x_1, \theta_2).$$

On obtient le résultat recherché en posant $\xi = (\theta_1, y_2)$ et $\xi' = (x_1, \theta_2)$ et en remarquant (voir la figure ci-dessus) que $\|\xi - x\| \leq \|y - x\|$ et $\|\xi' - x\| \leq \|y - x\|$. \square

Démonstration du théorème 1.29 Soit f de classe C^1 sur O , $x = (a, b) \in O$ et $\delta > 0$ tel que $U_\delta(x) \subset O$. Pour tout $(h, k) \in \mathbb{R}^2$ tel que $\|(h, k)\| < \delta$ on peut appliquer le lemme précédent pour obtenir

$$f(a + h, b + k) - f(a, b) = h \frac{\partial f}{\partial x_1}(\xi) + k \frac{\partial f}{\partial x_2}(\xi'),$$

pour des $\xi, \xi' \in \mathcal{U}_\delta(x)$ tels que

$$\|\xi - x\| \leq \|(h, k)\|, \quad \|\xi' - x\| \leq \|(h, k)\|. \quad (1.16)$$

On obtient donc la relation

$$\begin{aligned} & \frac{1}{\|(h, k)\|} \left(f(a+h, b+k) - \left[f(a, b) + h \frac{\partial f}{\partial x_1}(a, b) + k \frac{\partial f}{\partial x_2}(a, b) \right] \right) = \\ & \frac{h}{\|(h, k)\|} \left(\frac{\partial f}{\partial x_1}(\xi) - \frac{\partial f}{\partial x_1}(x) \right) + \frac{k}{\|(h, k)\|} \left(\frac{\partial f}{\partial x_2}(\xi') - \frac{\partial f}{\partial x_2}(x) \right). \end{aligned} \quad (1.17)$$

Lorsque $(h, k) \rightarrow (0, 0)$ on a d'une part

$$\left| \frac{h}{\|(h, k)\|} \right| \leq 1 \quad \text{et} \quad \left| \frac{k}{\|(h, k)\|} \right| \leq 1,$$

et d'autre part, en vertu de (1.16), $\xi \rightarrow x$ et $\xi' \rightarrow x$. La continuité des dérivées partielles impliquant que

$$\lim_{\xi \rightarrow x} \left(\frac{\partial f}{\partial x_1}(\xi) - \frac{\partial f}{\partial x_1}(x) \right) = \lim_{\xi' \rightarrow x} \left(\frac{\partial f}{\partial x_2}(\xi') - \frac{\partial f}{\partial x_2}(x) \right) = 0,$$

on déduit de (1.17) que

$$\lim_{(h,k) \rightarrow (0,0)} \frac{1}{\|(h, k)\|} \left(f(a+h, b+k) - \left[f(a, b) + h \frac{\partial f}{\partial x_1}(a, b) + k \frac{\partial f}{\partial x_2}(a, b) \right] \right) = 0,$$

c'est-à-dire que f est différentiable en (a, b) . □

Exemple 1.29 La fonction $x \mapsto \|x\|$ est différentiable sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. En effet on a

$$\begin{aligned} \frac{\partial \|x\|}{\partial x_1}(x_1, x_2) &= \frac{\partial \sqrt{x_1^2 + x_2^2}}{\partial x_1}(x_1, x_2) = \frac{x_1}{\sqrt{x_1^2 + x_2^2}} = \frac{x_1}{\|x\|}, \\ \frac{\partial \|x\|}{\partial x_2}(x_1, x_2) &= \frac{\partial \sqrt{x_1^2 + x_2^2}}{\partial x_2}(x_1, x_2) = \frac{x_2}{\sqrt{x_1^2 + x_2^2}} = \frac{x_2}{\|x\|}, \end{aligned}$$

et ces deux fonctions sont continues sur $\mathbb{R}^2 \setminus \{(0, 0)\}$ (pourquoi?). ◁

Si la fonction f est différentiable, on peut faire un peu mieux que le lemme 1.30 :

Théorème 1.31 (Théorème des accroissements finis) *Si f est différentiable sur un domaine convexe O alors, pour tout $x, y \in O$ il existe ξ , dans le segment de droite reliant x et y ($\xi = x + \theta(y - x)$, pour un $\theta \in [0, 1]$) tel que*

$$f(y) - f(x) = (y_1 - x_1) \frac{\partial f}{\partial x_1}(\xi) + (y_2 - x_2) \frac{\partial f}{\partial x_2}(\xi). \quad (1.18)$$

Démonstration. Par définition, O est convexe si $x, y \in O$ implique que le segment reliant x et y est tout entier inclus dans O : $\xi(t) = x + t(y - x) \in O$ pour tout $t \in [0, 1]$. Il résulte du lemme 1.26 que $F(t) = f(\xi(t))$ est différentiable sur cet intervalle. Le théorème des accroissements finis pour une fonction d'une variable permet d'écrire

$$f(y) - f(x) = F(1) - F(0) = F'(\theta)(1 - 0) = F'(\theta),$$

pour un $\theta \in [0, 1]$. D'autre part, le lemme 1.26 montre que

$$F'(\theta) = df_{\xi(\theta)}(\xi'(\theta)) = df_{\xi(\theta)}(y - x) = (y_1 - x_1) \frac{\partial f}{\partial x_1}(\xi(\theta)) + (y_2 - x_2) \frac{\partial f}{\partial x_2}(\xi(\theta)),$$

ce qui achève la démonstration. □



La convexité de O est essentielle dans ce théorème.

Exemple 1.30 $f(x) = \|x\|$ est différentiable sur $\mathbb{R}^2 \setminus \{(0, 0)\}$. Considérons les points $x = (-1, 0)$ et $y = (1, 0)$. On notera que le segment $S = [-1, 1] \times \{0\}$ reliant ces points n'est pas inclus dans $\mathbb{R}^2 \setminus \{(0, 0)\}$. On a $f(y) - f(x) = 0$, $y_1 - x_1 = 2$ et $y_2 - x_2 = 0$. Si la formule des accroissements finis (1.18) était vérifiée pour un $\xi \in S \setminus \{(0, 0)\}$ on aurait

$$f(y) - f(x) = 0 = 2 \frac{\partial f}{\partial x_1}(\xi) + 0 \frac{\partial f}{\partial x_2}(\xi) = 2 \frac{\xi_1}{\|\xi\|} = \pm 2,$$

ce qui est absurde. ◁

Théorème 1.32 (Théorème de Clairaut) Si f est de classe C^2 sur l'ouvert O alors

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a, b) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b),$$

pour tout $(a, b) \in O$.

Démonstration. Soit $(a, b) \in O$ fixé. Comme O est ouvert, $(a + h, b + k) \in O$ si $\|(h, k)\|$ est assez petit. Dans un premier temps nous considérons un (h, k) fixé satisfaisant cette condition. En appliquant à la fonction $g(u) = f(u, b + k) - f(u, b)$ le théorème des accroissements finis en 1 variable nous obtenons

$$g(a + h) - g(a) = hg'(a + \theta h), \tag{1.19}$$

pour un $\theta \in [0, 1]$. Comme

$$g'(u) = \frac{\partial f}{\partial x_1}(u, b + k) - \frac{\partial f}{\partial x_1}(u, b),$$

nous pouvons appliquer encore une fois le théorème des accroissements finis pour obtenir

$$g'(a + \theta h) = \frac{\partial f}{\partial x_1}(a + \theta h, b + k) - \frac{\partial f}{\partial x_1}(a + \theta h, b) = k \frac{\partial^2 f}{\partial x_2 \partial x_1}(a + \theta h, b + \theta' k), \tag{1.20}$$

pour un $\theta' \in [0, 1]$. On déduit donc de (1.19) (1.20) que

$$\frac{1}{h} \left(\frac{f(a + h, b + k) - f(a + h, b)}{k} - \frac{f(a, b + k) - f(a, b)}{k} \right) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a + \theta h, b + \theta' k).$$

Par construction, pour tout (h, k) suffisamment proche de $(0, 0)$ cette dernière identité est satisfaite pour une paire $(\theta, \theta') \in [0, 1] \times [0, 1]$.

Nous pouvons donc prendre la limite $k \rightarrow 0$ dans cette identité. Compte tenu de l'existence des dérivées premières et de la continuité des dérivées secondes, nous obtenons dans cette limite

$$\frac{1}{h} \left(\frac{\partial f}{\partial x_2}(a + h, b) - \frac{\partial f}{\partial x_2}(a, b) \right) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a + \theta h, b).$$

La limite $h \rightarrow 0$ permet alors de conclure

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(a, b) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(a, b).$$

◻

1.2.2.7 Formule de Taylor

Comme pour les fonctions d'une variable, l'existence de dérivées d'ordre supérieur permet des approximations plus précises que l'approximation linéaire discutée plus haut.

Théorème 1.33 Soient f de classe C^n sur l'ouvert $O \subset \mathbb{R}^2$, $(a, b) \in O$ et $\delta > 0$ tel que $U_\delta(a, b) \subset O$. La formule de Taylor

$$f(a + h, b + k) = \sum_{n_1 + n_2 < n} \frac{\partial^{n_1 + n_2} f}{\partial x_1^{n_1} \partial x_2^{n_2}}(a, b) \frac{h^{n_1} k^{n_2}}{n_1! n_2!} + R_n(a, b; h, k),$$

(la somme porte sur tous les entiers non négatifs n_1, n_2 satisfaisant $n_1 + n_2 < n$) est vérifiée pour tout $(h, k) \in \mathcal{U}_\delta(0, 0)$ avec un reste donné par la formule intégrale

$$R_n(a, b; h, k) = \sum_{j=0}^n \left(\int_0^1 \Phi_n(s) \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a + sh, b + sk) ds \right) \frac{h^j k^{n-j}}{j!(n-j)!}, \quad (1.21)$$

où $\Phi_n(s) = n(1-s)^{n-1}$. Une autre expression de ce reste est donnée par la formule de Lagrange :

$$R_n(a, b; h, k) = \sum_{j=0}^n \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a + \theta h, b + \theta k) \frac{h^j k^{n-j}}{j!(n-j)!}, \quad (1.22)$$

pour un $\theta \in [0, 1]$ dépendant de a, b, h, k .

Remarque 1.34 La forme précise du reste $R_n(a, b; h, k)$ est peu pratique. Dans les applications il est souvent suffisant d'avoir une estimation de ce reste. Pour obtenir une telle estimation, on remarque tout d'abord que la continuité des dérivées d'ordre n et le fait que $\theta \in [0, 1]$ impliquent que

$$\lim_{(h,k) \rightarrow (0,0)} \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a + \theta h, b + \theta k) = \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a, b), \quad (1.23)$$

pour tout $j = 1, \dots, n$. On note ensuite que

$$\left| \frac{h^j k^{n-j}}{\|(h, k)\|^n} \right| = \left| \frac{h^2}{h^2 + k^2} \right|^{j/2} \left| \frac{k^2}{h^2 + k^2} \right|^{\frac{n-j}{2}} \leq 1.$$

Nous pouvons donc écrire, à partir de la formule de Lagrange (1.22),

$$R_n(a, b; h, k) = \sum_{j=0}^n \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a, b) \frac{h^j k^{n-j}}{j!(n-j)!} + \varepsilon_n(a, b; h, k) \|(h, k)\|^n,$$

où

$$\varepsilon_n(a, b; h, k) = \sum_{j=0}^n \left(\frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a + \theta h, b + \theta k) - \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a, b) \right) \frac{1}{j!(n-j)!} \frac{h^j k^{n-j}}{\|(h, k)\|^n},$$

admet l'estimation

$$|\varepsilon_n(a, b; h, k)| \leq \sum_{j=0}^n \left| \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a + \theta h, b + \theta k) - \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a, b) \right| \frac{1}{j!(n-j)!}. \quad (1.24)$$

Il suit de (1.23) que $\varepsilon_n(a, b; h, k) \rightarrow 0$ lorsque $(h, k) \rightarrow (0, 0)$.

Nous avons donc obtenu le corollaire suivant.

Corollaire 1.35 *Sous les hypothèses du théorème 1.33,*

$$f(a + h, b + k) = \sum_{n_1 + n_2 \leq n} \frac{\partial^{n_1 + n_2} f}{\partial x_1^{n_1} \partial x_2^{n_2}}(a, b) \frac{h^{n_1} k^{n_2}}{n_1! n_2!} + \varepsilon_n(a, b; h, k) \|(h, k)\|^n, \quad (1.25)$$

où $\varepsilon_n(a, b; h, k) \rightarrow 0$ lorsque $(h, k) \rightarrow 0$.

Remarque 1.36 En utilisant la forme intégrale du reste (1.21) et en suivant l'argument menant au corollaire 1.35 on peut montrer que la fonction $\varepsilon_n(a, b; h, k)$ est continue sur tout sous-ensemble $\Lambda \subset \mathbb{R}^4$ tel que $\{(a + h, b + k) \mid (a, b, h, k) \in \Lambda\} \subset O$.

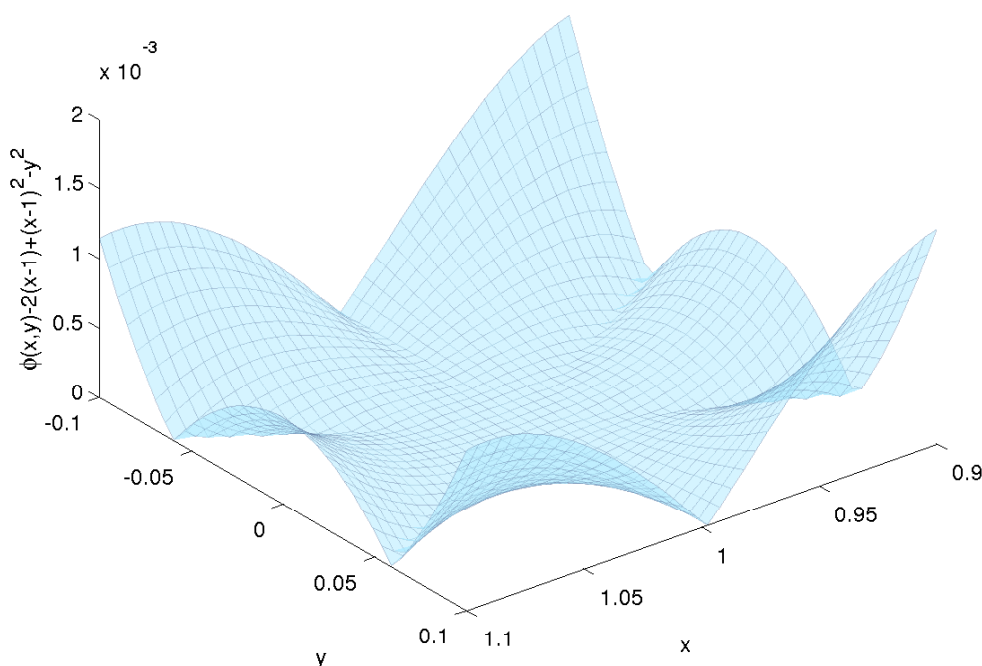
Exemple 1.31 Nous reprenons l'exemple 1.26. Les dérivées secondes de la fonction ϕ sont données par

$$\phi_{xx}(x, y) = -2 \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad \phi_{xy}(x, y) = \phi_{yx}(x, y) = -4 \frac{xy}{(x^2 + y^2)^2}, \quad \phi_{yy}(x, y) = 2 \frac{x^2 - y^2}{(x^2 + y^2)^2},$$

d'où l'on déduit $\phi_{xx}(1, 0) = -2$, $\phi_{xy}(1, 0) = \phi_{yx}(1, 0) = 0$, $\phi_{yy}(1, 0) = 2$. Le développement de Taylor à l'ordre 2 de la fonction ϕ en $(1, 0)$ est donc

$$\phi(x, y) \approx 2(x - 1) - (x - 1)^2 + y^2.$$

Comparez la figure suivante, illustrant l'erreur d'approximation, à celle de l'exemple 1.26 (notez bien l'échelle sur l'axe vertical !).



◁

Démonstration du théorème 1.33 Sous nos hypothèses $x(t) = (a + th, b + tk) \in O$ pour tout $t \in [0, 1]$. Montrons tout d'abord que la fonction $F(t) = f(x(t))$ est de classe C^n et ses dérivées sont données par la formule

$$F^{(n)}(t) = \sum_{j=0}^n \binom{n}{j} \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(x(t)) h^j k^{n-j}.$$

La démonstration se fait par induction. Pour $n = 1$ le lemme 1.26 montre qu'on a bien

$$F'(t) = \frac{d}{dt} f(x(t)) = df_{x(t)}(x'(t)) = h \frac{\partial f}{\partial x_1}(x(t)) + k \frac{\partial f}{\partial x_2}(x(t)).$$

Supposons la formule vérifiée pour $F^{(n)}(t)$, alors

$$F^{(n+1)}(t) = \sum_{j=0}^n \binom{n}{j} \frac{d}{dt} \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(x(t)) h^j k^{n-j},$$

et le lemme 1.26 nous donne

$$\frac{d}{dt} \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(x(t)) = h \frac{\partial^{n+1} f}{\partial x_1^{j+1} \partial x_2^{n-j}}(x(t)) + k \frac{\partial^{n+1} f}{\partial x_1^j \partial x_2^{n-j+1}}(x(t)).$$

Nous pouvons donc écrire $F^{(n+1)}(t)$ comme

$$\sum_{j=0}^n \binom{n}{j} \frac{\partial^{n+1} f}{\partial x_1^{j+1} \partial x_2^{n-j}}(x(t)) h^{j+1} k^{n-j} + \sum_{j=0}^n \binom{n}{j} \frac{\partial^{n+1} f}{\partial x_1^j \partial x_2^{n+1-j}}(x(t)) h^j k^{n+1-j}.$$

En changeant la variable de sommation dans le première somme, on arrive à

$$\sum_{j=1}^{n+1} \binom{n}{j-1} \frac{\partial^{n+1} f}{\partial x_1^j \partial x_2^{n+1-j}}(x(t)) h^j k^{n+1-j} + \sum_{j=0}^n \binom{n}{j} \frac{\partial^{n+1} f}{\partial x_1^j \partial x_2^{n+1-j}}(x(t)) h^j k^{n+1-j}.$$

En regroupant les termes avec $j = 1, 2, \dots, n$ on obtient l'expression

$$F^{(n+1)}(t) = \begin{cases} \binom{n}{0} \frac{\partial^{n+1} f}{\partial x_2^{n+1}}(x(t)) k^{n+1} \\ + \sum_{j=1}^n \left[\binom{n}{j-1} + \binom{n}{j} \right] \frac{\partial^{n+1} f}{\partial x_1^j \partial x_2^{n+1-j}}(x(t)) h^j k^{n+1-j} \\ + \binom{n}{n} \frac{\partial^{n+1} f}{\partial x_1^{n+1}}(x(t)) h^{n+1}. \end{cases}$$

Finalement, en remarquant que

$$\binom{n}{0} = 1 = \binom{n+1}{0}, \quad \binom{n}{n} = 1 = \binom{n+1}{n+1},$$

et

$$\binom{n}{j-1} + \binom{n}{j} = \frac{n!j}{j!(n+1-j)!} + \frac{n!(n+1-j)}{j!(n+1-j)!} = \frac{(n+1)!}{j!(n+1-j)!} = \binom{n+1}{j},$$

on obtient le résultat recherché

$$F^{(n+1)}(t) = \sum_{j=0}^{n+1} \binom{n+1}{j} \frac{\partial^{n+1} f}{\partial x_1^j \partial x_2^{n+1-j}}(x(t)) h^j k^{n+1-j}.$$

Appliquons maintenant la formule de Taylor en une variable à la fonction $F(t)$,

$$F(t) = \sum_{m=0}^{n-1} F^{(m)}(0) \frac{t^m}{m!} + R_n(t),$$

avec le reste

$$R_n(t) = \frac{t^n}{(n-1)!} \int_0^1 (1-s)^{n-1} F^{(n)}(s) ds.$$

En posant $t = 1$ on obtient le développement de la fonction f

$$f(a+h, b+k) = \sum_{m=0}^{n-1} \frac{1}{m!} \sum_{j=0}^m \binom{m}{j} \frac{\partial^m f}{\partial x_1^j \partial x_2^{m-j}}(a, b) h^j k^{m-j} + R_n(1).$$

Après un changement de variable de sommation, ce développement s'écrit

$$f(a+h, b+k) = \sum_{n_1+n_2 < n} \frac{\partial^{n_1+n_2} f}{\partial x_1^{n_1} \partial x_2^{n_2}}(a, b) \frac{h^{n_1} k^{n_2}}{n_1! n_2!} + R_n, \quad (1.26)$$

où le reste s'exprime comme

$$R_n = n \sum_{j=0}^n \left(\int_0^1 (1-s)^{n-1} \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a+sh, b+sk) ds \right) \frac{h^j k^{n-j}}{j!(n-j)!}.$$

En invoquant la formule de Lagrange pour le reste

$$R_n(t) = F^{(n)}(\theta) \frac{t^n}{n!},$$

pour un $\theta \in [0, t]$, on obtient la forme alternative suivante pour le reste dans (1.26)

$$R_n = \sum_{j=0}^n \frac{\partial^n f}{\partial x_1^j \partial x_2^{n-j}}(a + \theta h, b + \theta k) \frac{h^j k^{n-j}}{j!(n-j)!}.$$

□

1.2.2.8 Extrema

La détermination des extrema, maxima ou minima, d'une fonction est un problème important apparaissant dans de nombreuses applications. Comme dans le cas des fonctions d'une variable, le calcul différentiel s'avère un outil efficace pour ces questions.

Définition 1.37 Un point $x_0 \in D(f)$ est appelé *point critique* de f si toutes ses dérivées partielles du premier ordre s'annulent en ce point.

Exemple 1.32 Déterminons les points critiques de la fonction $f(x, y) = x^2 + \alpha xy + y^2$ en fonction du paramètre α . Les dérivées premières de f sont

$$f_x(x, y) = 2x + \alpha y, \quad f_y(x, y) = 2y + \alpha x.$$

Les points critiques de f sont donc les solutions du système homogène

$$\begin{aligned} 2x + \alpha y &= 0, \\ \alpha x + 2y &= 0. \end{aligned}$$

Le déterminant de ce système est $4 - \alpha^2$. Pour $\alpha \notin \{-2, 2\}$, l'unique solution est $(0, 0)$. Pour $\alpha = \pm 2$, les solutions forment une droite d'équation $y = \mp x$. La fonction f a donc un unique point critique $x = (0, 0)$ si $\alpha \notin \{-2, 2\}$, une infinité de points critiques $\{x = (u, u) \mid u \in \mathbb{R}\}$ si $\alpha = -2$ et une infinité de points critiques $\{x = (u, -u) \mid u \in \mathbb{R}\}$ si $\alpha = 2$. ◁

Définition 1.38 (i) Un point $x_0 \in D(f)$ est un *maximum local* de f si il existe un voisinage de x_0 , $U_\delta(x_0) \subset D(f)$, tel que $f(x) \leq f(x_0)$ pour tout $x \in U_\delta(x_0)$.

(ii) Un point $x_0 \in D(f)$ est un *minimum local* de f si il existe un voisinage de x_0 , $U_\delta(x_0) \subset D(f)$, tel que $f(x) \geq f(x_0)$ pour tout $x \in U_\delta(x_0)$.

Théorème 1.39 Si f admet des dérivées partielles sur l'ouvert O et si $x \in O$ est un maximum ou un minimum local de f , alors x est un point critique de f .

Définition 1.40 Un point critique de f qui n'est ni un maximum local ni un minimum local est appelé *point de selle* ou *col*.

Démonstration du théorème 1.39 Si $x_0 = (a, b) \in O$ est un minimum local de f , a est un minimum local de la fonction $x \mapsto f(x, b)$. La dérivée de cette fonction doit donc s'annuler en $x = a$, c'est-à-dire que

$$\frac{\partial f}{\partial x_1}(a, b) = 0.$$

En considérant la fonction $y \mapsto f(a, y)$ on montre de façon analogue que

$$\frac{\partial f}{\partial x_2}(a, b) = 0.$$

Le cas d'un maximum local se traite de manière similaire. □

Comme dans le cas des fonctions d'une seule variable, il est utile d'examiner les dérivées secondes pour déterminer la nature d'un point critique. La situation dans le cas de deux variables est cependant plus compliquée et nécessite une petite préparation.

Formes quadratiques dans \mathbb{R}^2 . Une **forme quadratique** dans \mathbb{R}^2 est une fonction de la forme

$$f(x_1, x_2) = Ax_1^2 + 2Bx_1x_2 + Cx_2^2,$$

où A, B et C sont des coefficients réels que nous supposons non tous nuls. Son **déterminant** est défini par

$$\Delta = AC - B^2.$$

1. Si $\Delta > 0$ alors $A \neq 0$ et on peut écrire, en complétant le carré,

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{A} ((Ax_1)^2 + 2(Ax_1)(Bx_2) + (Bx_2)^2) + \frac{1}{A} (AC - B^2)x_2^2 \\ &= \frac{1}{A} [(Ax_1 + Bx_2)^2 + \Delta x_2^2]. \end{aligned} \quad (1.27)$$

On en déduit immédiatement que $f(x_1, x_2)$ ne s'annule qu'en $(0, 0)$ et qu'en tout autre point le signe de $f(x_1, x_2)$ est celui de A . Si $A > 0$, on dit que f est **définie positive**. Dans ce cas $f(x_1, x_2) > 0$ pour tout $(x_1, x_2) \neq (0, 0)$. En particulier, $\lambda = \min\{f(\cos \varphi, \sin \varphi) \mid \varphi \in [0, 2\pi]\} > 0$ et on a

$$f(x_1, x_2) = f(r \cos \varphi, r \sin \varphi) = f(\cos \varphi, \sin \varphi)r^2 \geq \lambda r^2 = \lambda \|(x_1, x_2)\|^2.$$

Si $A < 0$, f est **définie négative** et $f(x_1, x_2) < 0$ pour tout $(x_1, x_2) \neq (0, 0)$. Dans ce cas on montre de manière analogue que

$$f(x_1, x_2) \leq -\lambda \|(x_1, x_2)\|^2,$$

où $\lambda = \min\{-f(\cos \varphi, \sin \varphi) \mid \varphi \in [0, 2\pi]\} > 0$.

2. Si $\Delta < 0$, la forme quadratique f est **indéfinie**. Si $A \neq 0$, la l'équation (1.27) ci-dessus montre que $f(x_1, x_2)$ prend aussi bien des valeurs positives que négatives. Si $A = 0$ on arrive à la même conclusion en observant que

$$f(x_1, x_2) = x_2(2Bx_1 + Cx_2).$$

3. Si $\Delta = 0$ la forme quadratique est **semi-définie**. Si $A \neq 0$ l'équation (1.27) devient

$$f(x_1, x_2) = \frac{1}{A} (Ax_1 + Bx_2)^2.$$

On remarque que f s'annule sur la droite d'équation $Ax_1 + Bx_2 = 0$. Elle possède, en dehors de cette dernière, le même signe que A . Si $A = 0$ on doit avoir $B = 0$ et $f(x_1, x_2) = Cx_2^2$. Dans ce cas f s'annule sur la droite d'équation $x_2 = 0$ et prend en dehors de cette droite le signe de C .

Revenons maintenant à l'étude de la fonction f au voisinage du point critique $x_0 = (a, b)$. Nous supposons que f est de classe C^2 dans ce voisinage. Dans ce cas, elle admet un développement de Taylor à l'ordre 2 et l'expression (1.25), avec $n = 2$, donne

$$f(a + h, b + k) = f(a, b) + \frac{1}{2} (Ah^2 + 2Bhk + Ck^2) + \varepsilon(h, k) \|(h, k)\|^2,$$

où

$$A = \frac{\partial^2 f}{\partial x_1^2}(a, b), \quad B = \frac{\partial^2 f}{\partial x_1 \partial x_2}(a, b), \quad C = \frac{\partial^2 f}{\partial x_2^2}(a, b), \quad (1.28)$$

et $\varepsilon(h, k) \rightarrow 0$ lorsque $(h, k) \rightarrow (0, 0)$.

Supposons que la forme quadratique $Ah^2 + 2Bhk + Ck^2$ soit définie positive. Il existe donc $\lambda > 0$ tel que $Ah^2 + 2Bhk + Ck^2 \geq \lambda \|(h, k)\|^2$. Comme $\varepsilon(h, k) \rightarrow 0$, $(\lambda/2 + \varepsilon(h, k)) > 0$ si (h, k) est assez proche de $(0, 0)$, on en déduit que

$$f(a + h, b + k) - f(a, b) \geq \left(\frac{\lambda}{2} + \varepsilon(h, k) \right) \|(h, k)\|^2 \geq 0,$$

c'est-à-dire que (a, b) est un minimum local de f .

De manière analogue, on montre que (a, b) est un maximum local de f si la forme quadratique $Ah^2 + 2Bhk + Ck^2$ est définie négative. Lorsqu'elle est indéfinie, $f(a + h, b + k) - f(a, b)$ admet aussi bien des

valeurs positives que négatives pour des petits (h, k) . Dans ce cas (a, b) est un point de selle. Lorsque la forme quadratique $Ah^2 + 2Bhk + Ck^2$ est semi-définie, il est impossible de conclure sur la base des dérivées secondes et une analyse plus détaillée, que nous n'aborderons pas ici, se révèle nécessaire.

En résumé, si $x_0 = (a, b)$ est un point critique de la fonction f et si A, B, C sont donnés par (1.28) la nature de ce point est déterminé par le tableau suivant.

$AC - B^2$	A	nature de (a, b)
+	+	minimum local
+	-	maximum local
-	\pm	point de selle
0		?

Exemple 1.33 Considérons la fonction $f(x, y) = (x^2 - y^2)e^{-(x^2+y^2)}$ qui est clairement de classe C^∞ . Ses dérivées premières sont

$$f_x(x, y) = 2x(1 - x^2 + y^2)e^{-(x^2+y^2)}, \quad f_y(x, y) = -2y(1 + x^2 - y^2)e^{-(x^2+y^2)}.$$

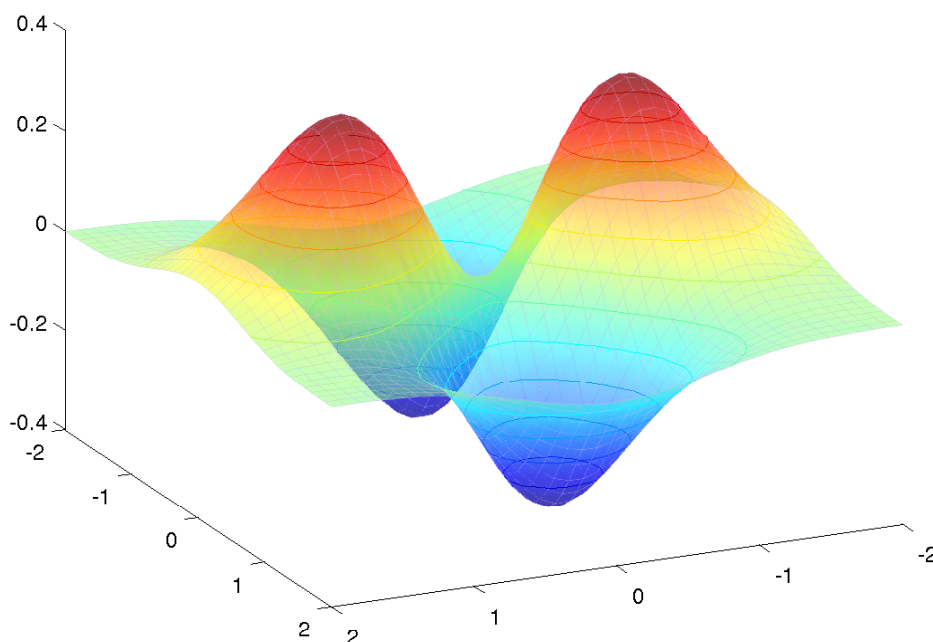
On a donc $f_x(x, y) = 0$ si et seulement si $x = 0$ ou $x^2 - y^2 = 1$ et $f_y(x, y) = 0$ si et seulement si $y = 0$ ou $x^2 - y^2 = -1$. On en déduit que f possède 5 points critiques $(0, 0), (0, \pm 1), (\pm 1, 0)$. Les dérivées secondes de f sont

$$\begin{aligned} f_{xx}(x, y) &= 2(2x^4 - 2x^2y^2 - 5x^2 + y^2 + 1)e^{-(x^2+y^2)}, \\ f_{xy}(x, y) = f_{yx}(x, y) &= 4xy(x^2 - y^2)e^{-(x^2+y^2)}, \\ f_{yy}(x, y) &= -2(2y^4 - 2x^2y^2 - 5y^2 + x^2 + 1)e^{-(x^2+y^2)}. \end{aligned}$$

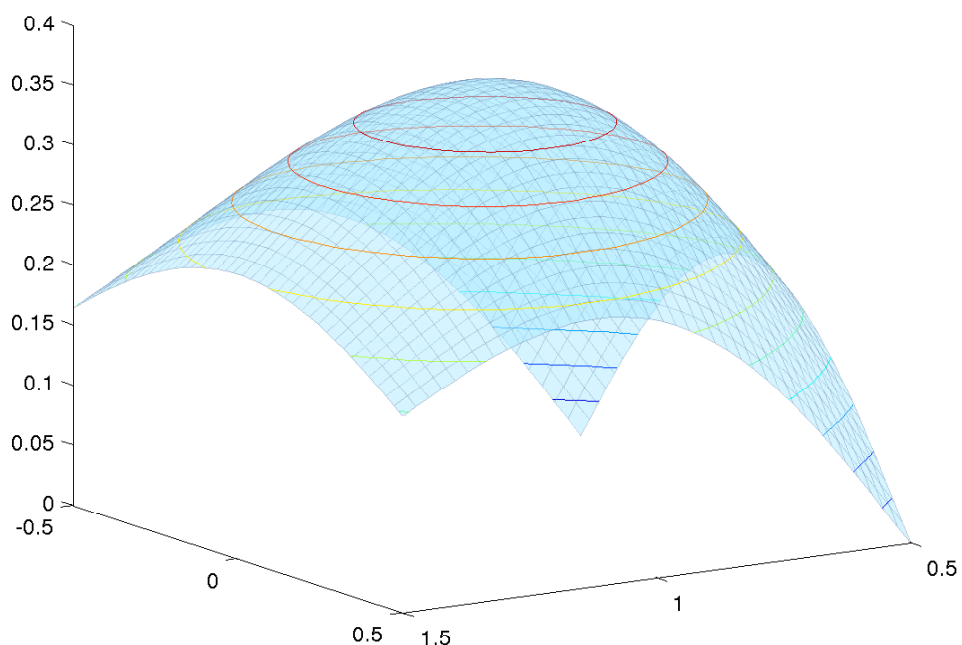
On a donc le tableau suivant

Point critique	$AC - B^2$	A	nature du point critique
$(0, 0)$	-4	2	point de selle
$(0, \pm 1)$	$16e^{-2}$	$4e^{-1}$	minimum local
$(\pm 1, 0)$	$16e^{-2}$	$-4e^{-1}$	maximum local

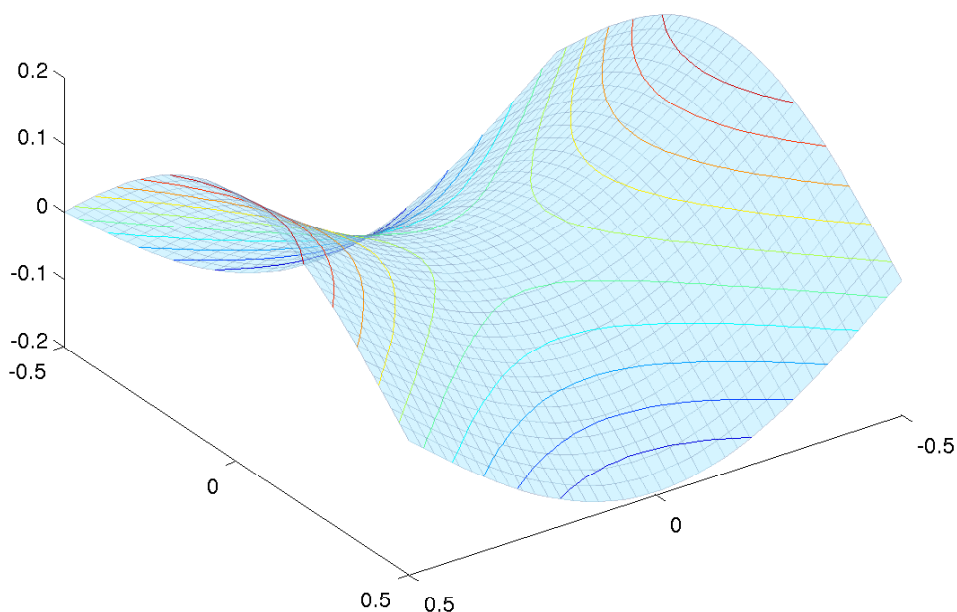
à comparer avec le graphe ci-dessous!



Observez, sur la figure suivante, la structure des courbes de niveau au voisinage d'un maximum local (ici le point $(1, 0)$). Elles ressemblent à des ellipses. Bien entendu on observerait le même type de courbes de niveau près d'un minimum local.



Le comportement des courbes de niveau est très différent au voisinage d'un point de selle. Dans ce cas, ces courbes s'apparentent à des hyperboles !



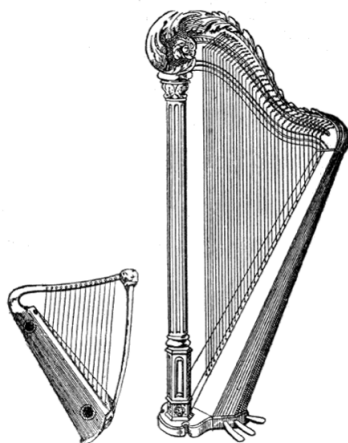
Cette différence se comprend mieux en considérant l'équation $Ah^2 + 2Bhk + Ck^2 = \pm 1$. Dans le cas du maximum local, c'est l'équation d'une ellipse alors que dans le cas d'un point de selle, c'est l'équation d'une hyperbole. \triangleleft

1.2.2.9 Le gradient

Définition 1.41 Le *gradient* de la fonction $f(x_1, x_2)$ en (a, b) est le vecteur

$$\nabla f(a, b) = \left(\frac{\partial f}{\partial x_1}(a, b), \frac{\partial f}{\partial x_2}(a, b) \right).$$

Le symbole ∇ se lit “nabla”, d’un mot grec désignant une harpe. On utilise également la notation $\text{grad } f$ pour le gradient de f .



Harpes, p. 984.

Si f est différentiable en (a, b) , on peut écrire sa différentielle et ses dérivées directionnelles comme un produit scalaire

$$(Df)_{(u,v)}(a, b) = df_{(a,b)}(u, v) = \nabla f(a, b) \cdot (u, v), \tag{1.29}$$

et son développement de Taylor d’ordre 1 comme

$$f(a + h, b + k) = f(a, b) + \nabla f(a, b) \cdot (h, k) + \varepsilon(h, k)\|(h, k)\|.$$

On remarquera en particulier qu’un point critique de f est caractérisé par le fait que son gradient s’anule

$$\nabla f(a, b) = 0 \iff \frac{\partial f}{\partial x_1}(a, b) = \frac{\partial f}{\partial x_2}(a, b) = 0.$$

Si f est de classe C^1 sur O , le gradient de f est défini en chaque point de O et ses deux composantes sont des fonctions continues sur O . Le champ de gradient de f est l’application

$$\begin{aligned} \nabla f : \quad O &\rightarrow \mathbb{R}^2 \\ (x_1, x_2) &\mapsto \nabla f(x_1, x_2), \end{aligned}$$

qui associe un vecteur à chaque point de O . C’est un premier exemple de champ vectoriel, objets sur lesquels nous reviendrons dans le paragraphe 1.2.3. Dans ce paragraphe, nous nous intéresserons aux propriétés géométriques du gradient.

Pour comprendre la signification géométrique du vecteur gradient on applique l’inégalité de Cauchy-Schwarz à la relation (1.29). On se souviendra que cette inégalité stipule que le produit scalaire de deux vecteurs x et y satisfait

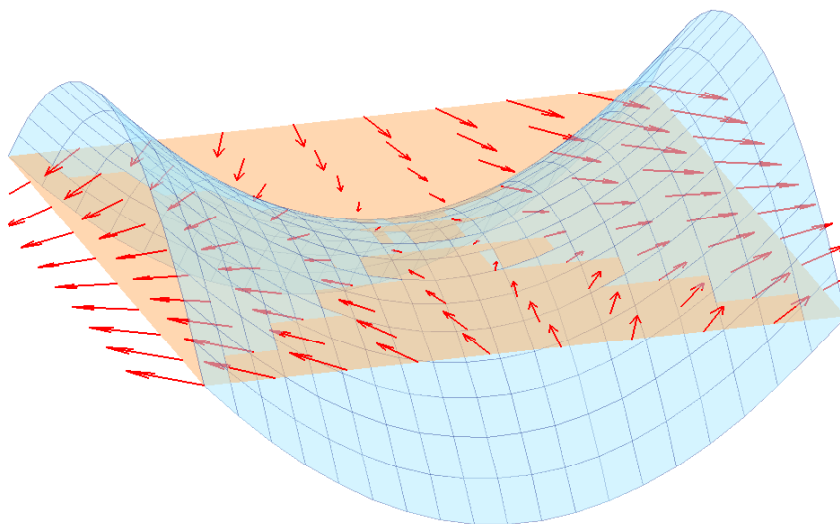
$$x \cdot y \leq \|x\|\|y\|,$$

et que cette inégalité est saturée, c’est-à-dire que $x \cdot y = \|x\|\|y\|$, si et seulement si x et y ont la même direction (il suffit pour s’en convaincre de considérer la formule $x \cdot y = \|x\|\|y\| \cos \angle(x, y)$ où $\angle(x, y)$ désigne l’angle entre x et y). On en déduit que si (u, v) est un vecteur unité la dérivée directionnelle de f selon cette direction satisfait

$$(Df)_{(u,v)}(a, b) \leq \|\nabla f(a, b)\|,$$

et qu'on a l'égalité si la direction (u, v) est celle du gradient $\nabla f(a, b)$. La norme du gradient de f en (a, b) est donc égale à la plus grande dérivée directionnelle de cette fonction en ce point. D'après la discussion du paragraphe 1.2.2.2, c'est aussi la plus forte pente du graphe de f au point (a, b) . La direction du gradient est, quant à elle, celle de la plus forte pente, c'est-à-dire la direction dans laquelle la fonction f croît le plus rapidement.

Exemple 1.34 La figure suivante montre l'allure du champ de gradient au voisinage du point de selle $(0, 0)$ de la fonction f de l'exemple 1.33.



◁

Si la courbe différentiable $x(t) = (x_1(t), x_2(t))$ est une courbe de niveau de la fonction f , alors $f(x(t))$ est une constante indépendante de t . On a donc

$$0 = \frac{d}{dt} f(x(t)) = df_{x(t)}(x'(t)) = \nabla f(x(t)) \cdot x'(t),$$

qui exprime le fait que le gradient ∇f est, en tout point $x(t)$ de la courbe de niveau, perpendiculaire à son vecteur vitesse $x'(t)$, lui-même tangent à la courbe. Le gradient de f est donc partout perpendiculaire à ses courbes de niveau.

Exemple 1.35 La fonction $f(x) = \|x\|^{-1}$ est différentiable en dehors de l'origine $(0, 0)$. Son gradient est

$$\nabla f(x) = -\frac{x}{\|x\|^3} = \left(-\frac{x_1}{\|x\|^3}, -\frac{x_2}{\|x\|^3} \right).$$

Les courbes de niveau de f sont des cercles centrés à l'origine

$$f(x) = c \iff \|x\| = \frac{1}{c}.$$

Elles sont paramétrées par

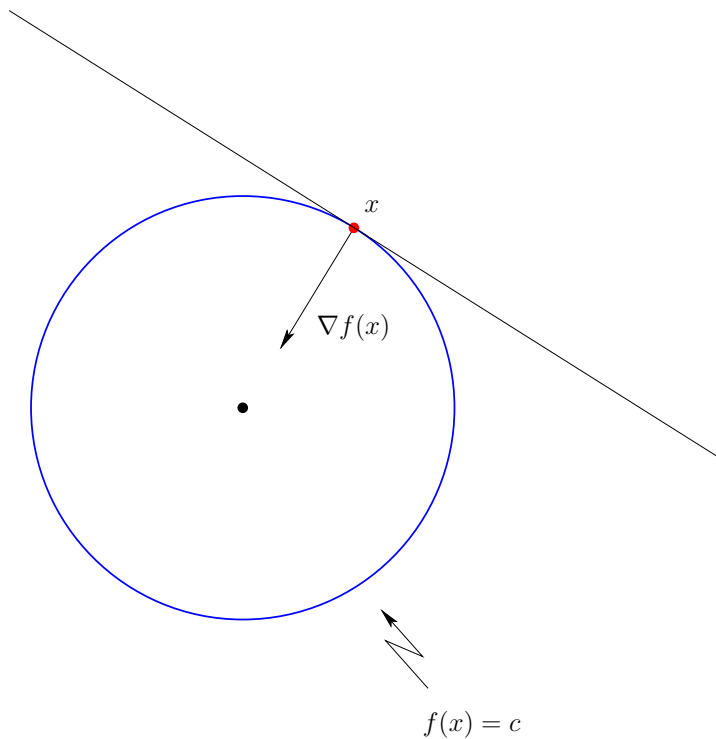
$$x(t) = (x_1(t), x_2(t)) = \left(\frac{1}{c} \cos t, \frac{1}{c} \sin t \right).$$

Le vecteur vitesse sur ces courbes est

$$x'(t) = (x'_1(t), x'_2(t)) = \left(-\frac{1}{c} \sin t, \frac{1}{c} \cos t \right).$$

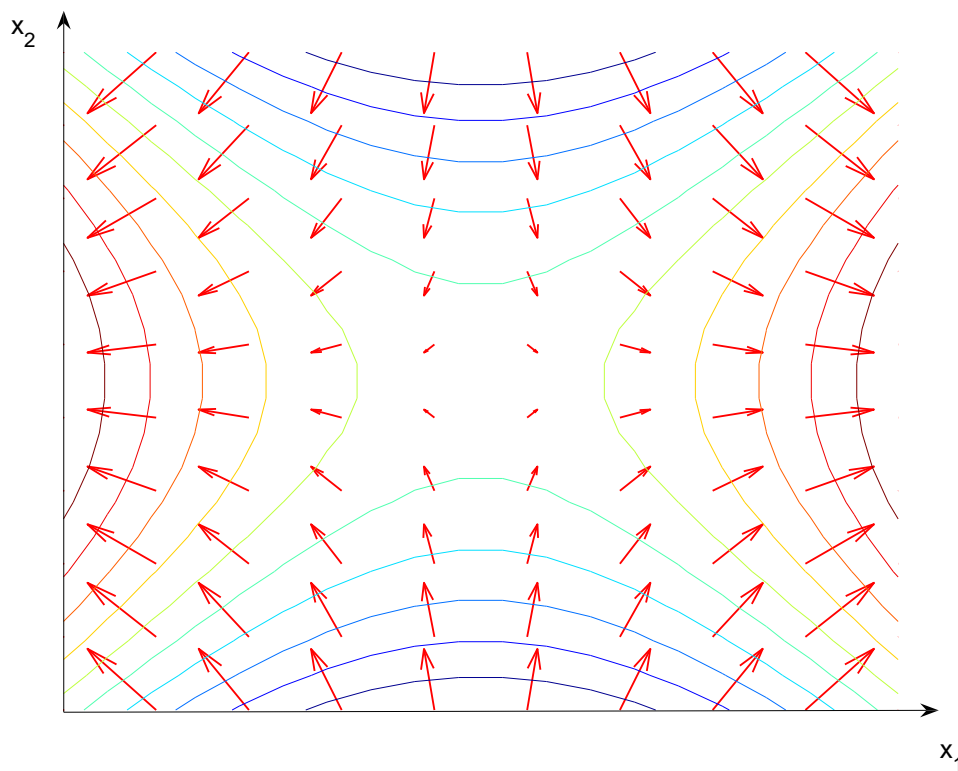
On vérifie bien que ce vecteur est perpendiculaire au gradient de f

$$\nabla f(x(t)) \cdot x'(t) = -\frac{x(t)}{\|x(t)\|^3} \cdot x'(t) = -(\cos t, \sin t) \cdot \left(-\frac{1}{c} \sin t, \frac{1}{c} \cos t \right) = 0.$$



△

Exemple 1.36 Sur la figure suivante sont représentées les courbes de niveau ainsi que le champ de gradient de la fonction f de l'exemple 1.33 au voisinage du point de selle $(0, 0)$.



△

1.2.3 Champs vectoriels

Un **champ vectoriel** sur $\Omega \subset \mathbb{R}^2$ est une fonction $X : \Omega \rightarrow \mathbb{R}^2$ qui, à chaque point $x = (x_1, x_2)$ de Ω associe un vecteur

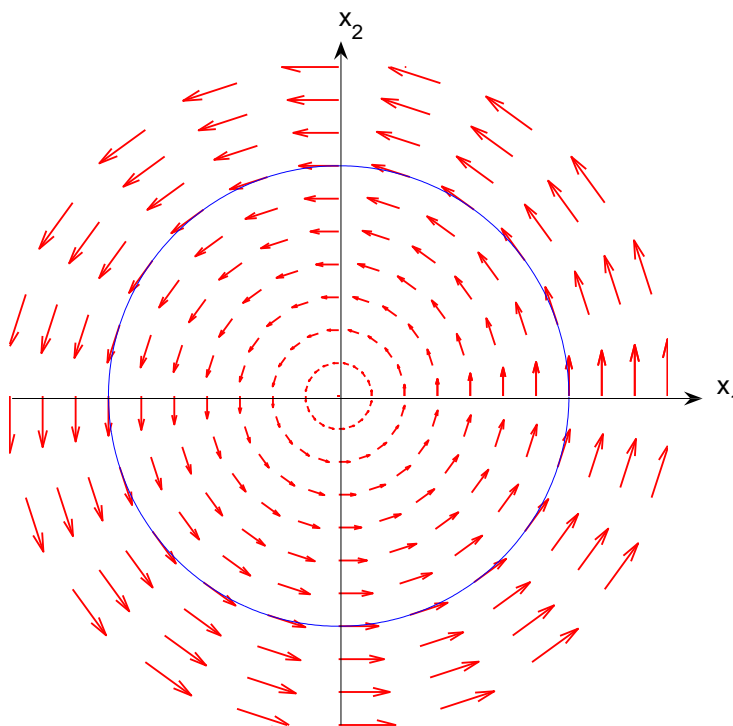
$$X(x) = \begin{pmatrix} X_1(x_1, x_2) \\ X_2(x_1, x_2) \end{pmatrix},$$

de \mathbb{R}^2 . Comme pour un champ scalaire, on dit que Ω est le **domaine** de X qu'on dénote $D(X)$. On définit de manière analogue un champ vectoriel sur un domaine de \mathbb{R}^3 ou de manière générale sur un domaine de \mathbb{R}^n . On parle alors d'un champ vectoriel de dimension n . Nous avons déjà rencontré un exemple de champs vectoriels dans le paragraphe 1.2.2.9 : le champ de gradient d'un champ scalaire. En voici d'autres.

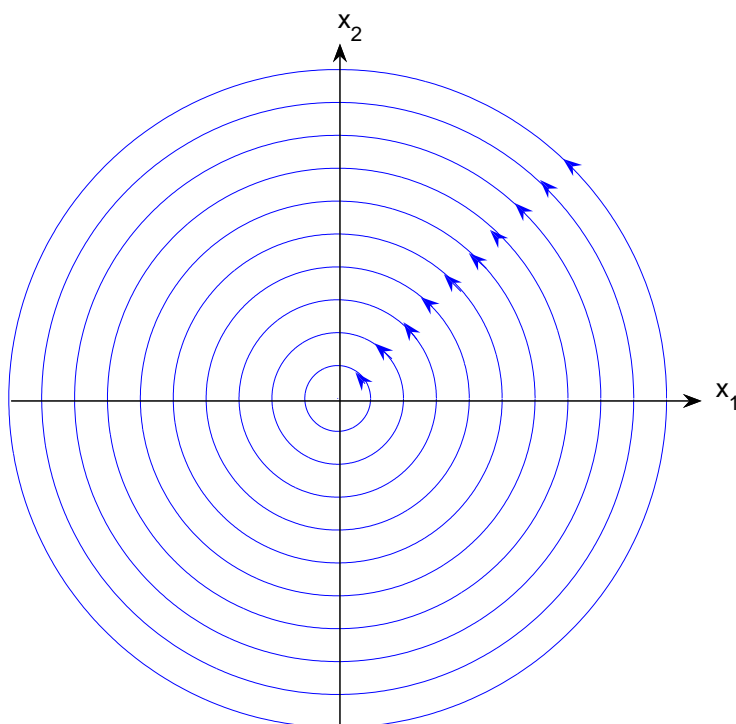
Exemple 1.37 La figure suivante est un portrait du champ vectoriel

$$X(x_1, x_2) = \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}, \quad (1.30)$$

défini pour $x = (x_1, x_2) \in D(X) = \mathbb{R}^2$. Ce portrait est obtenu en dessinant une flèche proportionnelle au vecteur $X(x)$ en un certain nombre de points $x \in D(X)$.



On notera sur cette figure que le vecteur $X(x)$ est partout tangent aux cercles centrés sur l'origine. On appelle **ligne de champ** une courbe qui, en chacun de ses points, est tangente au champ. Les cercles centrés sur l'origine sont donc les lignes de champ du champ (1.30). On représente parfois un champ vectoriel en dessinant uniquement ses lignes de champ.



Une telle représentation permet de visualiser la direction du champ, par contre elle ne contient aucune information sur son intensité, c'est-à-dire sur la longueur $\|X(x)\|$. Par exemple le champ

$$Y(x_1, x_2) = \frac{1}{x_1^2 + x_2^2} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \frac{X(x_1, x_2)}{x_1^2 + x_2^2}, \quad (1.31)$$

a les même lignes de champ que X , il est cependant très différent (en particulier, il est singulier en $x = (0, 0)$). \triangleleft

Les champs vectoriels jouent un rôle fondamental en géométrie différentielle, sujet que nous n'aborderons pas dans ce cours. Ils sont aussi omniprésents en physique.

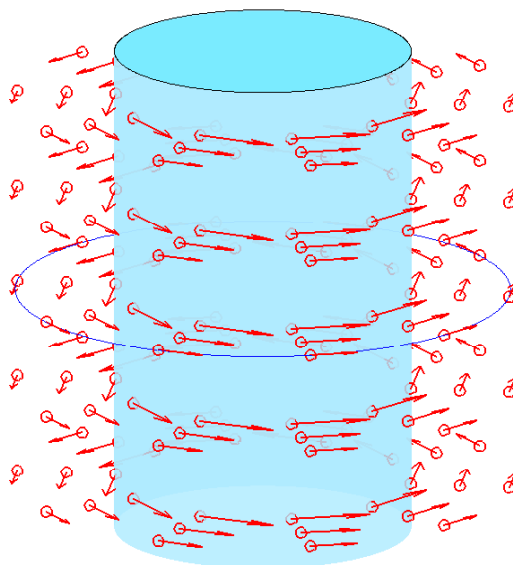
Exemple 1.38 Voici le champ magnétique engendré par un courant électrique stationnaire d'intensité I dans un conducteur cylindrique infiniment long, de rayon R , centré sur l'axe x_3 . C'est un champ vectoriel de dimension 3. Son expression analytique est donnée par la loi de Biot-Savart

$$B(x_1, x_2, x_3) = \frac{2I}{c(x_1^2 + x_2^2)} \begin{pmatrix} -x_2 \\ x_1 \\ 0 \end{pmatrix}, \quad (1.32)$$

où c est la vitesse de la lumière. Son domaine est l'extérieur du conducteur,

$$D(B) = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 > R^2\}.$$

En voici le portrait.



On vérifie facilement que les lignes de champ sont ici les cercles horizontaux centrés sur l'axe x_3 . En effet, un tel cercle est décrit par la courbe

$$x(t) = (x_1(t), x_2(t), x_3(t)) = (a \cos t, a \sin t, b),$$

où $a > R$ est le rayon du cercle et $(0, 0, b)$ son centre. Le vecteur vitesse de cette courbe, qui en dirige la tangente, est donc

$$x'(t) = \begin{pmatrix} x'_1(t) \\ x'_2(t) \\ x'_3(t) \end{pmatrix} = \begin{pmatrix} -a \sin t \\ a \cos t \\ 0 \end{pmatrix} = \begin{pmatrix} -x_2(t) \\ x_1(t) \\ 0 \end{pmatrix},$$

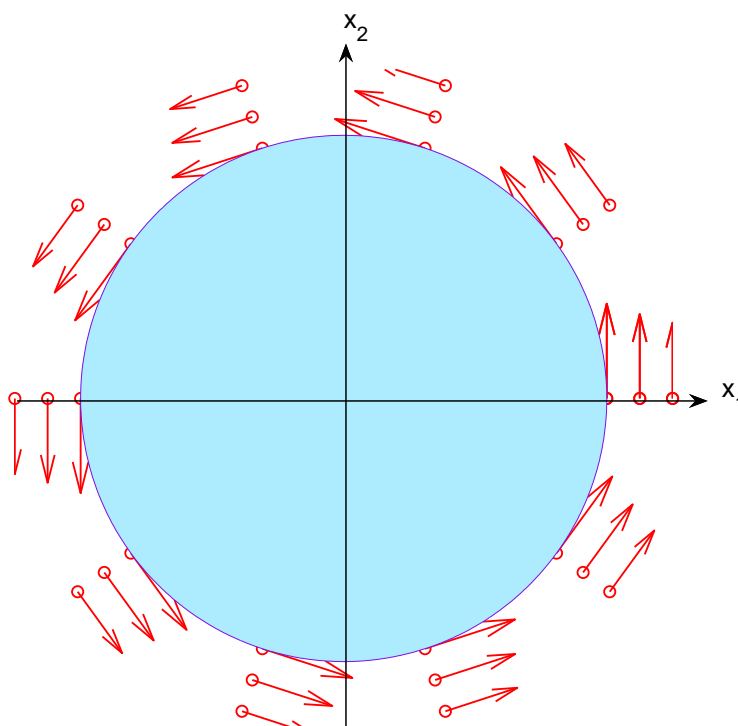
et on trouve bien qu'il est proportionnel à $B(x(t))$,

$$x'(t) = \frac{ca^2}{2I} B(x(t)).$$

Une autre caractéristique remarquable du champ B est que d'une part il ne dépend pas de la coordonnée x_3 et d'autre part sa troisième composante est toujours nulle. Cette propriété est une conséquence de la symétrie de translation dans la direction verticale du système physique en question. Elle permet en outre de réduire la description du champ B à l'aide d'une coupe horizontale. En se restreignant à un plan horizontal (par exemple $x_3 = 0$), on peut décrire le champ B à l'aide d'un champ de dimension 2,

$$b(x_1, x_2) = \frac{2I}{c(x_1^2 + x_2^2)} \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix},$$

dont voici le portrait.



◁

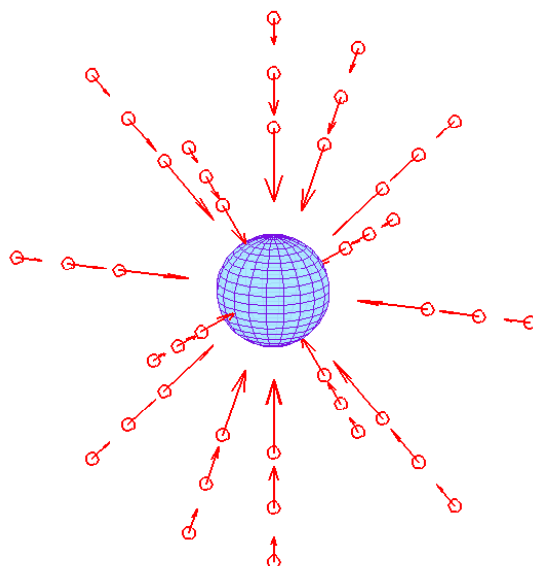
Exemple 1.39 Un autre exemple est le champ de gravitation terrestre. En plaçant la terre à l'origine du système de coordonnées, ce champ est donné par la formule de Newton

$$A(x_1, x_2, x_3) = -GM \frac{x}{\|x\|^3} = -GM \frac{1}{(x_1^2 + x_2^2 + x_3^2)^{3/2}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \quad (1.33)$$

où $G = 6.673 \cdot 10^{-11} \text{ m}^3\text{Kg}^{-1}\text{s}^{-2}$ est la constante de gravitation universelle et M la masse de la terre. Son domaine est

$$D(A) = \{x = (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1^2 + x_2^2 + x_3^2 > R\},$$

où R est le rayon terrestre (le champ de gravitation est différent à l'intérieur de la terre!). Tout corps de masse m plongé dans ce champ subit une force égale à mA . Le portrait du champ de gravitation est donné dans la figure suivante. On identifie aisément les lignes de champ : ce sont les demi-droites pointant vers l'origine.



Ici aussi on remarque la présence d’une symétrie. Si \mathcal{R} est une rotation d’angle quelconque autour d’un axe passant par l’origine, alors $A(\mathcal{R}x) = \mathcal{R}A(x)$. Comme dans l’exemple précédent, cette propriété permet de réduire la description du champ en effectuant une coupe par un plan quelconque passant par l’origine. En choisissant le plan $x_3 = 0$ on obtient une représentation de dimension 2 du champ gravitationnel

$$a(x_1, x_2) = -GM \frac{1}{(x_1^2 + x_2^2)^{3/2}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

◁

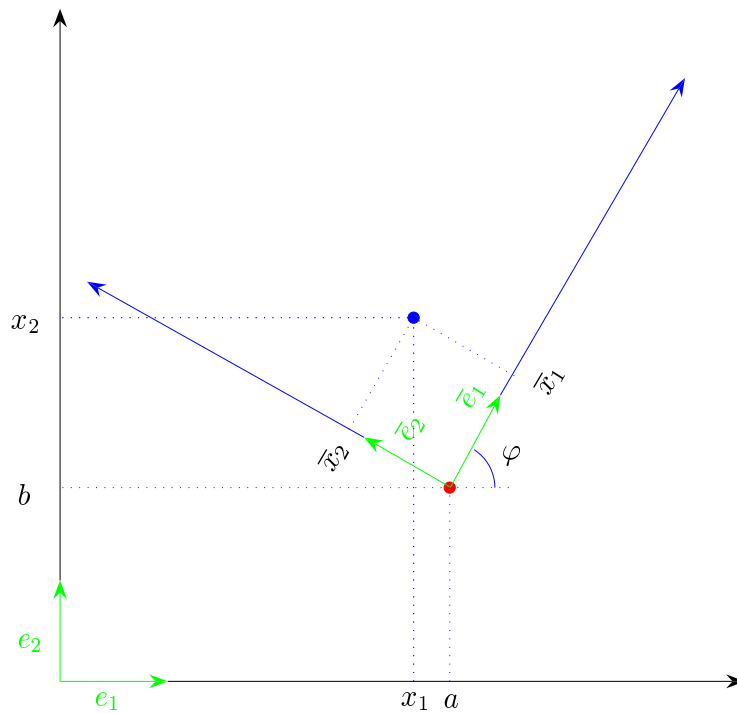
1.2.3.1 Changements de coordonnées

A première vue, un champ vectoriel semble être simplement une collection de champs scalaires. En effet les composantes $X_1(x_1, x_2)$ et $X_2(x_1, x_2)$ sont des fonctions à valeur réelle, tout comme des champs scalaires. Pour faire apparaître la distinction entre champ scalaire et champ vectoriel il faut considérer le comportement de ces objets lors d’un changement de coordonnées.

Rappelons tout d’abord qu’un système de coordonnées *cartésien* dans le *plan euclidien orienté* est spécifié par une origine \mathcal{O} et une base orthonormée $\{e_1, e_2\}$, formée de deux vecteurs tels que $\|e_1\| = 1$ et e_2 résulte d’une rotation d’angle $\pi/2$ dans le sens positif de e_1 . En particulier on a $e_1 \cdot e_1 = e_2 \cdot e_2 = 1$ et $e_1 \cdot e_2 = 0$. Les composantes d’un vecteur X dans ce système de coordonnées sont les deux nombres X_1 et X_2 tels que $X = X_1 e_1 + X_2 e_2$. On peut les obtenir en prenant le produit scalaire de X avec les vecteurs de base : $X_1 = X \cdot e_1$, $X_2 = X \cdot e_2$. Les coordonnées x_1, x_2 d’un point P sont les composantes du vecteur $\overrightarrow{\mathcal{O}P}$. Les axes du système de coordonnées sont les deux droites dirigées par e_1 et e_2 et passant par \mathcal{O} .

Introduisons un nouveau système de coordonnées d’origine $\overline{\mathcal{O}}$ et de base $\{\overline{e}_1, \overline{e}_2\}$. Soient (a, b) les coordonnées de $\overline{\mathcal{O}}$ dans l’ancien système de coordonnées et φ la mesure algébrique de l’angle entre e_1 et \overline{e}_1 : $\overrightarrow{\mathcal{O}\overline{\mathcal{O}}} = a e_1 + b e_2$, $\overline{e}_1 = e_1 \cos \varphi + e_2 \sin \varphi$, $\overline{e}_2 = -e_1 \sin \varphi + e_2 \cos \varphi$. Les nouvelles coordonnées $(\overline{x}_1, \overline{x}_2)$ d’un point P sont reliées à ses coordonnées originales (x_1, x_2) par la formule

$$\begin{aligned} x_1 &= a + \overline{x}_1 \cos \varphi - \overline{x}_2 \sin \varphi, \\ x_2 &= b + \overline{x}_1 \sin \varphi + \overline{x}_2 \cos \varphi. \end{aligned} \tag{1.34}$$

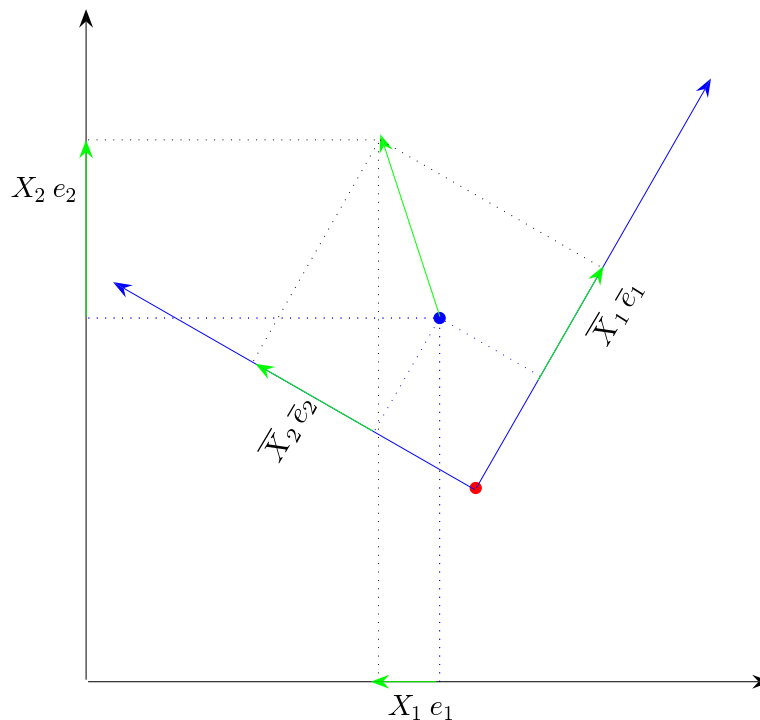


Si f est champ scalaire, $f(x_1, x_2)$ exprime une propriété (par exemple la température) liée au point P de coordonnées (x_1, x_2) . Dans le nouveau système de coordonnées la même propriété s'exprime donc par le champ scalaire \bar{f} défini par

$$\bar{f}(\bar{x}_1, \bar{x}_2) = f(x_1, x_2), \tag{1.35}$$

où x_1 et x_2 sont les fonctions de \bar{x}_1 et \bar{x}_2 définie par (1.34). De même, si X est un champ vectoriel ses composantes $\bar{X}_1(\bar{x}_1, \bar{x}_2)$ et $\bar{X}_2(\bar{x}_1, \bar{x}_2)$ dans le nouveau système de coordonnées sont déterminées par ses composantes $X_1(x_1, x_2)$ et $X_2(x_1, x_2)$ dans le système original par la relation

$$\bar{X}_1(\bar{x}_1, \bar{x}_2)\bar{e}_1 + \bar{X}_2(\bar{x}_1, \bar{x}_2)\bar{e}_2 = X_1(x_1, x_2)e_1 + X_2(x_1, x_2)e_2.$$



On en déduit aisément les formules de changement de coordonnées d'un champ vectoriel

$$\begin{aligned}\bar{X}_1(\bar{x}_1, \bar{x}_2) &= X_1(x_1, x_2) \cos \varphi + X_2(x_1, x_2) \sin \varphi, \\ \bar{X}_2(\bar{x}_1, \bar{x}_2) &= -X_1(x_1, x_2) \sin \varphi + X_2(x_1, x_2) \cos \varphi.\end{aligned}\tag{1.36}$$

Notez bien que les composantes d'un champ vectoriel *ne se transforment pas* comme un champ scalaire !

Déterminons encore comment se transforme le gradient d'un champ scalaire. Dans les nouvelles coordonnées, le gradient du champ f est donné par

$$\nabla \bar{f}(\bar{x}_1, \bar{x}_2) = \left(\frac{\partial \bar{f}}{\partial \bar{x}_1}(\bar{x}_1, \bar{x}_2), \frac{\partial \bar{f}}{\partial \bar{x}_2}(\bar{x}_1, \bar{x}_2) \right),$$

où $\bar{f}(\bar{x}_1, \bar{x}_2)$ est défini par la formule (1.35). Nous pouvons calculer ses composantes à l'aide du lemme 1.26 et de la formule (1.11). Pour cela il suffit de considérer \bar{x}_2 comme une constante dans (1.34) et $\xi(t) = (a + t \cos \varphi - \bar{x}_2 \sin \varphi, b + t \sin \varphi + \bar{x}_2 \cos \varphi)$ comme une courbe différentiable. On obtient

$$\frac{\partial \bar{f}}{\partial \bar{x}_1}(\bar{x}_1, \bar{x}_2) = \left. \frac{df(\xi(t))}{dt} \right|_{t=x_1} = \frac{\partial f}{\partial x_1}(x_1, x_2) \cos \varphi + \frac{\partial f}{\partial x_2}(x_1, x_2) \sin \varphi.$$

La seconde composante du gradient se calcule de manière analogue. On obtient ainsi la transformation

$$\begin{aligned}\frac{\partial \bar{f}}{\partial \bar{x}_1}(\bar{x}_1, \bar{x}_2) &= \frac{\partial f}{\partial x_1}(x_1, x_2) \cos \varphi + \frac{\partial f}{\partial x_2}(x_1, x_2) \sin \varphi, \\ \frac{\partial \bar{f}}{\partial \bar{x}_2}(\bar{x}_1, \bar{x}_2) &= -\frac{\partial f}{\partial x_1}(x_1, x_2) \sin \varphi + \frac{\partial f}{\partial x_2}(x_1, x_2) \cos \varphi.\end{aligned}$$

Une comparaison avec (1.36) montre que le gradient se transforme bien comme un vecteur dans un changement de coordonnées.

Remarque 1.42 Nous n'avons considéré que des systèmes de coordonnées cartésiens dans le sens précisé au début de ce paragraphe. Lorsqu'on admet des systèmes de coordonnées plus généraux (c'est-à-dire des transformations de coordonnées non orthogonales) on remarque que le gradient ne se transforme pas comme un vecteur, on dit que c'est un co-vecteur. Il en va de même de certains champs comme le champ magnétique (exemple 1.38) qui est un pseudo-vecteur (ou vecteur axial). Il change de signe par rapport aux vrais vecteurs (ou vecteurs polaires) lorsqu'on inverse l'orientation de l'espace. Nous n'approfondirons pas ces subtilités dans ce cours.

1.2.3.2 Dérivées partielles d'un champ vectoriel

Les composantes X_1, X_2 d'un champ vectoriel étant des fonctions de $D(X)$ dans \mathbb{R} on peut, comme pour un champ scalaire, définir les dérivées partielles d'un champ vectoriel

$$\frac{\partial X_1}{\partial x_1}(a, b), \quad \frac{\partial X_1}{\partial x_2}(a, b), \quad \frac{\partial X_2}{\partial x_1}(a, b), \quad \frac{\partial X_2}{\partial x_2}(a, b).$$

Tout comme nous avons regroupé les dérivées partielles d'un champ scalaire dans un vecteur, son gradient, il est commode de regrouper les dérivées d'un champ vectoriel sous la forme d'un tableau ou d'une matrice 2×2 (ou plus généralement $n \times n$ si le champ est de dimension n)

$$\frac{\partial(X_1, X_2)}{\partial(x_1, x_2)}(a, b) = \begin{pmatrix} \frac{\partial X_1}{\partial x_1}(a, b) & \frac{\partial X_1}{\partial x_2}(a, b) \\ \frac{\partial X_2}{\partial x_1}(a, b) & \frac{\partial X_2}{\partial x_2}(a, b) \end{pmatrix},\tag{1.37}$$

qu'on appelle **matrice jacobienne** du champ X . Le déterminant de cette matrice

$$JX_{(a,b)} = \det \left(\frac{\partial(X_1, X_2)}{\partial(x_1, x_2)}(a, b) \right) = \frac{\partial X_1}{\partial x_1}(a, b) \frac{\partial X_2}{\partial x_2}(a, b) - \frac{\partial X_1}{\partial x_2}(a, b) \frac{\partial X_2}{\partial x_1}(a, b),$$

est une quantité importante sur laquelle nous reviendrons au paragraphe 1.2.4.1, on l'appelle **jacobien** du champ X . Il en est de même de la trace de la matrice jacobienne qu'on appelle **divergence** du champ X

$$\operatorname{div} X_{(a,b)} = \operatorname{tr} \left(\frac{\partial(X_1, X_2)}{\partial(x_1, x_2)}(a, b) \right) = \frac{\partial X_1}{\partial x_1}(a, b) + \frac{\partial X_2}{\partial x_2}(a, b).$$

Exemple 1.40 Le champ (1.30) admet des dérivées partielles en chaque point de \mathbb{R}^2 . Sa matrice jacobienne est constante

$$\frac{\partial(X_1, X_2)}{\partial(x_1, x_2)}(a, b) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

ainsi que son jacobien $JX_{(a,b)} = 1$. Sa divergence est nulle $\operatorname{div} X_{(a,b)} = 0$. Le champ (1.31) admet quant à lui des dérivées partielles en tout point de $\mathbb{R}^2 \setminus \{0\}$. Sa matrice jacobienne est

$$\frac{\partial(Y_1, Y_2)}{\partial(x_1, x_2)}(a, b) = \begin{pmatrix} \frac{2ab}{(a^2 + b^2)^2} & \frac{-a^2 + b^2}{(a^2 + b^2)^2} \\ \frac{-a^2 + b^2}{(a^2 + b^2)^2} & \frac{-2ab}{(a^2 + b^2)^2} \end{pmatrix},$$

son jacobien $JY_{(a,b)} = -(a^2 + b^2)^{-2}$ et sa divergence est nulle $\operatorname{div} Y_{(a,b)} = 0$. ◁

1.2.3.3 Champs vectoriels différentiables

Définition 1.43 Un champs vectoriel X est **différentiable** en $(a, b) \in D(X)$ si ses composantes X_1 et X_2 sont des fonctions différentiables en (a, b) . Il est **différentiable sur l'ouvert** O si il est différentiable en chaque point de O . Il est de **classe C^n** sur O si ses composantes les sont.

Vérifiez que f est différentiable en (a, b) si et seulement si

$$X(a + h, b + k) = X(a, b) + dX_{(a,b)}(h, k) + \varepsilon(h, k)\|(h, k)\|,$$

où $\varepsilon(h, k) \rightarrow 0$ lorsque $(h, k) \rightarrow (0, 0)$ et $(h, k) \mapsto dX_{(a,b)}(h, k)$ dénote l'application linéaire

$$dX_{(a,b)}(h, k) = \left(\frac{\partial X_1}{\partial x_1}(a, b)h + \frac{\partial X_1}{\partial x_2}(a, b)k, \frac{\partial X_2}{\partial x_1}(a, b)h + \frac{\partial X_2}{\partial x_2}(a, b)k \right),$$

associée à la matrice jacobienne (1.37). Comme pour les champs scalaires, $dX_{(a,b)}$ est appelée **différentielle** ou **dérivée** du champ vectoriel X en (a, b) .

La plupart des résultats concernant les fonctions différentiables des paragraphes précédents s'étendent immédiatement aux champs vectoriels différentiables. Par exemple, si X est de classe C^1 sur O il est différentiable sur cet ouvert.

Exemple 1.41 Le champ de gravitation (1.33) est de classe C^∞ sur $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$. Sa matrice jacobienne est

$$dA_{(x_1, x_2, x_3)} = -GM \begin{pmatrix} \frac{-2x_1^2 + x_2^2 + x_3^2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} & \frac{-3x_1x_2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} & \frac{-3x_1x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \\ \frac{-3x_1x_2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} & \frac{x_1^2 - 2x_2^2 + x_3^2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} & \frac{-3x_2x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \\ \frac{-3x_1x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} & \frac{-3x_2x_3}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} & \frac{x_1^2 + x_2^2 - 2x_3^2}{(x_1^2 + x_2^2 + x_3^2)^{5/2}} \end{pmatrix}.$$

Malgré l'apparente complexité de cette formule, l'application linéaire correspondante est très simple. Pour s'en rendre compte il faut écrire la matrice jacobienne sous la forme

$$dA_x = -\frac{GM}{r^3} \left(\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 3 \begin{pmatrix} \hat{x}_1\hat{x}_1 & \hat{x}_1\hat{x}_2 & \hat{x}_1\hat{x}_3 \\ \hat{x}_2\hat{x}_1 & \hat{x}_2\hat{x}_2 & \hat{x}_2\hat{x}_3 \\ \hat{x}_3\hat{x}_1 & \hat{x}_3\hat{x}_2 & \hat{x}_3\hat{x}_3 \end{pmatrix} \right) = -\frac{GM}{r^3}(I - 3P),$$

où $x = (x_1, x_2, x_3)$, $r = \|x\|$ et $\hat{x} = (\hat{x}_1, \hat{x}_2, \hat{x}_3) = (x_1, x_2, x_3)/r = x/r$ est le vecteur unité dans la direction de x . La matrice identité I correspond à l'application identique $u \mapsto u$ de \mathbb{R}^3 alors qu'à la matrice P correspond la projection orthogonale sur la droite vectorielle engendrée par x , $u \mapsto (\hat{x} \cdot u)\hat{x}$. Relativement à une base orthonormée dont le premier vecteur est \hat{x} , la matrice jacobienne prend donc la forme simple

$$dA_x = -\frac{GM}{r^3} \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (1.38)$$

Son jacobien est le déterminant de cette matrice, $JA_x = 2G^3M^3r^{-9}$. On note finalement que $\text{div } A_x = 0$.

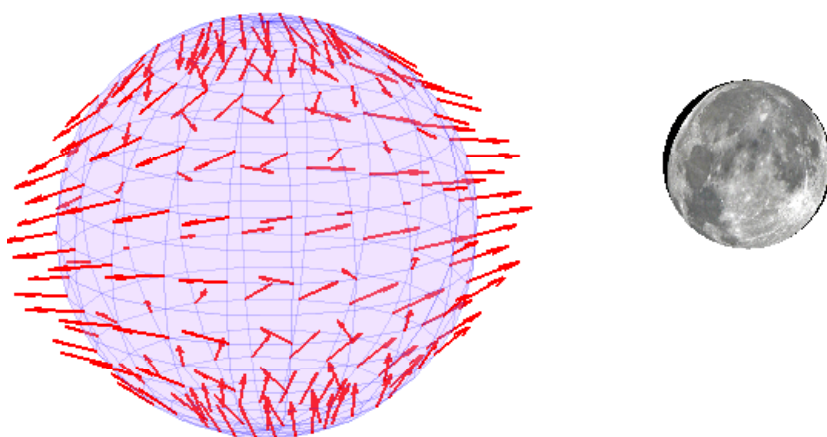
Une application de la matrice jacobienne que nous venons d'obtenir est le calcul des forces de marée. L'inhomogénéité du champ gravitationnel est à la source de ces forces. Pour le comprendre considérons l'effet du champ gravitationnel de lune sur la terre. Plaçons la lune à l'origine des coordonnées et désignons par x le centre de gravité de la terre. L'accélération de ce centre de gravité due à la lune est alors donnée par $A(x)$ (formule (1.33), où M désigne la masse de la lune). Un point $x + u$ situé à la surface de la terre subit quant à lui une accélération

$$A(x + u) = A(x) + \delta A(u).$$

Comme le champ A n'est pas constant, cette accélération n'est pas la même que celle du centre de gravité : $\delta A(u) \neq 0$. Un point de la surface de la terre est donc accéléré relativement à son centre de gravité. En vertu de la loi de Newton, cette différence d'accélération induit une force. Cette force de marée est responsable du phénomène bien connu portant ce nom. En première approximation, elle est donnée par le développement de Taylor

$$\delta A(u) = dA_x u. \quad (1.39)$$

La figure suivante en donne une illustration (pas à l'échelle!) : l'accélération différentielle $dA_x u$ à la surface de la terre est représentée par les flèches rouges. Le profil qui en résulte donne une idée qualitative de la déformation du globe terrestre et des masses océaniques due aux forces de marée.



Il est intéressant de comparer les accélérations du centre de masse de la terre dues à la lune et au soleil et les accélérations différentielles à la surface de la terre dues aux forces de marée correspondantes. L'accélération du centre de masse due à un corps de masse m placé à une distance d de la terre résulte de (1.33),

$$\|A\| = \frac{Gm}{d^2},$$

alors que l'accélération due aux forces de marée est approximativement donnée par (1.39) et (1.38),

$$\|\delta A(\mathbf{u})\| \simeq \frac{GmR}{d^3},$$

où $R = 6.4 \cdot 10^6$ m est le rayon moyen de la terre.

	masse	distance (moyenne) à la terre	$\ A\ $	$\ \delta A\ $
Lune	$7.3 \cdot 10^{22}$ Kg	$3.8 \cdot 10^8$ m	$3.3 \cdot 10^{-5}$ m/s ²	$4.5 \cdot 10^{-7}$ m/s ²
Soleil	$2.0 \cdot 10^{30}$ Kg	$1.5 \cdot 10^{11}$ m	$5.9 \cdot 10^{-3}$ m/s ²	$2.5 \cdot 10^{-7}$ m/s ²

On remarque que si l'accélération du centre de masse due à la lune est 100 fois plus faible que celle due au soleil, l'effet de marée de la lune est presque 2 fois plus important! \triangleleft

1.2.3.4 Laplacien et fonctions harmoniques

Si f est un champ scalaire de classe C^2 , son gradient est différentiable. La divergence de ce gradient est appelée **laplacien** de f et noté

$$\Delta f(x_1, x_2) = \operatorname{div} \nabla f(x_1, x_2) = \frac{\partial^2 f}{\partial x_1^2}(x_1, x_2) + \frac{\partial^2 f}{\partial x_2^2}(x_1, x_2).$$

Une fonction f est dite **harmonique** sur Ω si $\Delta f(x) = 0$ pour tout $x \in \Omega$.

Exemple 1.42 Le potentiel $\phi(x_1, x_2) = \log(x_1^2 + x_2^2)$ est harmonique sur $\mathbb{R}^2 \setminus \{0\}$. \triangleleft

1.2.4 Transformations

Dans les paragraphes 1.2.3 nous avons interprété les valeurs prises par une fonction $X : \Omega \rightarrow \mathbb{R}^2$ comme des vecteurs. Nous allons maintenant les considérer comme des points du plan. Dans cette perspective, une fonction

$$\begin{aligned} x : \Omega &\rightarrow \mathbb{R}^2 \\ (\mathbf{u}_1, \mathbf{u}_2) &\mapsto (x_1(\mathbf{u}_1, \mathbf{u}_2), x_2(\mathbf{u}_1, \mathbf{u}_2)), \end{aligned}$$

est une transformation du domaine Ω dans son image $x(\Omega) = \{x(\mathbf{u}_1, \mathbf{u}_2) \mid (\mathbf{u}_1, \mathbf{u}_2) \in \Omega\}$.

Exemple 1.43 Soient $\mathbf{a} = (a_1, a_2), \mathbf{b} = (b_1, b_2) \in \mathbb{R}^2$ deux vecteurs linéairement indépendants et $\mathbf{c} = (c_1, c_2) \in \mathbb{R}^2$. L'application affine x définie par $x(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{u}_1 \mathbf{a} + \mathbf{u}_2 \mathbf{b} + \mathbf{c}$ est une bijection du plan dans lui-même. A l'aide de la matrice

$$A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix},$$

cette transformation s'écrit $x(\mathbf{u}) = A\mathbf{u} + \mathbf{c}$. Comme $x(\mathbf{u} + \mathbf{h}) - x(\mathbf{u}) = A\mathbf{h}$, x est partout différentiable et sa matrice jacobienne est $dx_{\mathbf{u}} = A$ pour tout $\mathbf{u} \in \mathbb{R}^2$. Notez que la matrice A est non-singulière car ses colonnes sont linéairement indépendantes. Le jacobien de x est lui aussi constant $J_{\mathbf{u}}x = \det A = a_1 b_2 - a_2 b_1$.

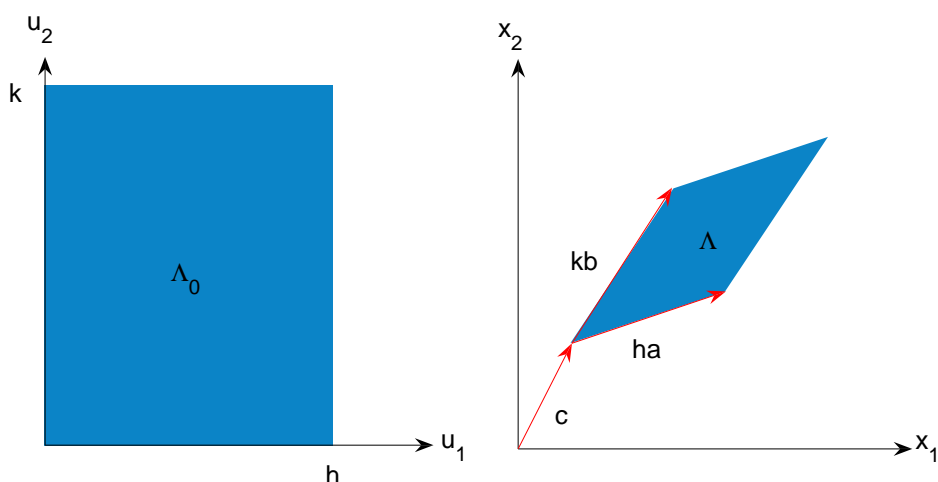
La matrice inverse

$$A^{-1} = \frac{1}{\det A} \begin{pmatrix} b_2 & -b_1 \\ -a_2 & a_1 \end{pmatrix},$$

permet d'écrire l'application réciproque de x comme $x^{-1}(\mathbf{u}) = A^{-1}\mathbf{u} + \mathbf{c}'$ où $\mathbf{c}' = -A^{-1}\mathbf{c}$. On remarque donc que $dx_{\mathbf{u}}^{-1} = A^{-1}$. Dans le cas particulier $\mathbf{c} = (0, 0), \mathbf{a} = (1, 0), \mathbf{b} = (0, 1)$ l'application x est l'identité $x(\mathbf{u}) = \mathbf{u}$ tout comme sa différentielle $dx_{\mathbf{u}}(\mathbf{h}) = \mathbf{h}$.

x transforme le rectangle $\Lambda_0 = [0, h] \times [0, k]$ en un parallélogramme

$$\Lambda = \{s\mathbf{a} + t\mathbf{b} + \mathbf{c} \mid s, t \in [0, 1]\}.$$



La valeur absolue du déterminant d'une matrice 2×2 est égale à l'aire du parallélogramme engendré par les deux vecteurs colonnes de cette matrice (démontrez-le !). On en déduit que l'aire du parallélogramme Λ est donné par

$$\left| \det \begin{pmatrix} ha_1 & kb_1 \\ ha_2 & kb_2 \end{pmatrix} \right| = |ha_1 kb_2 - ha_2 kb_1| = |J_u x| hk.$$

Comme hk est l'aire du rectangle Λ_0 , on peut résumer cette remarque en disant que la transformation x multiplie les aires par $|J_u x|$. \triangleleft

1.2.4.1 Jacobien d'une transformation

Une transformation x différentiable en (u_0, v_0) se comporte localement comme une transformation affine

$$x : (u_0 + h, v_0 + k) \mapsto x(u_0, v_0) + hx_u(u_0, v_0) + kx_v(u_0, v_0) + \text{erreur}.$$

L'exemple précédent nous amène naturellement à penser que l'aire de l'image $x(R)$ d'un petit rectangle $R = [u_0, u_0 + h] \times [v_0, v_0 + k]$ est bien approximée par

$$\text{Aire de } x(R) \simeq |J_{x(u_0, v_0)}| \text{ Aire de } R.$$

Dans ce paragraphe nous démontrons une formulation précise de ce fait qui nous sera très utile pour traiter du changement de variables dans les intégrales doubles au chapitre suivant.

Théorème 1.44 Soient $x : O \rightarrow \mathbb{R}^2$ une transformation de classe C^1 sur l'ouvert O , $\Lambda \subset O$ un sous-ensemble fermé et $\delta > 0$ tel que $(u + h, v + k) \in O$ pour tous $(u, v) \in \Lambda$ et $(h, k) \in U_\delta(0, 0)$. Il existe une fonction Δ telle que

$$\lim_{(h, k) \rightarrow (0, 0)} \Delta(h, k) = 0,$$

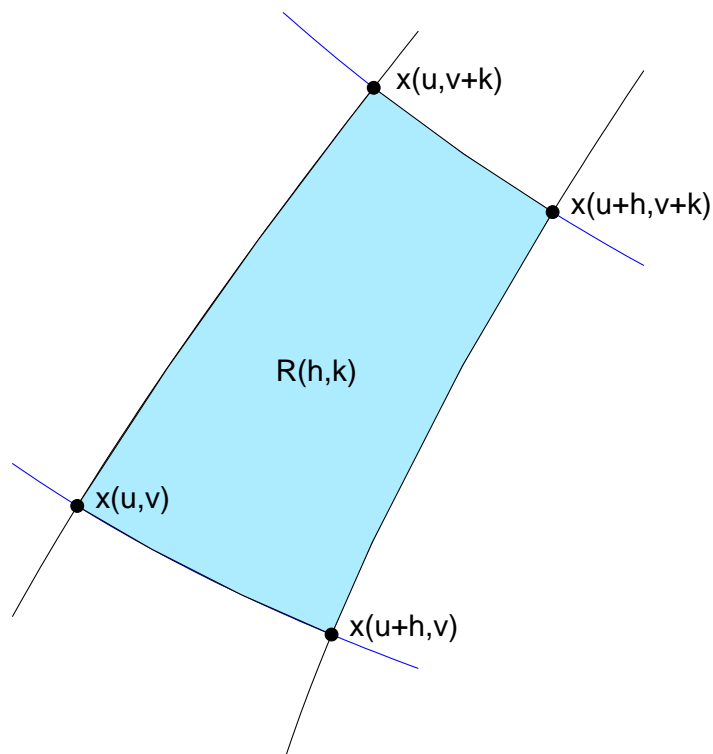
et

$$\text{Aire de } x([u, u + h] \times [v, v + k]) = |J_{x(u, v)}| |hk| + \|(h, k)\|^2 \Delta(h, k),$$

pour tous $(u, v) \in \Lambda$, $(h, k) \in U_\delta(0, 0)$.

Démonstration. Fixons $(u, v) \in \Lambda$ et pour tout $(h, k) \in U_\delta(0, 0)$ posons

$$R(h, k) = x([u, u + h] \times [v, v + k]).$$



La fonction x étant de classe C^1 , le corollaire 1.35 et la remarque 1.36 montrent que pour tous $s, t \in [0, 1]$,

$$x(u + sh, v + tk) - (sha + tkb + c) = \|(sh, tk)\| \varepsilon(u, v; sh, tk),$$

où les vecteurs a, b, c sont donnés par

$$a = x_u(u, v), \quad b = x_v(u, v), \quad c = x(u, v),$$

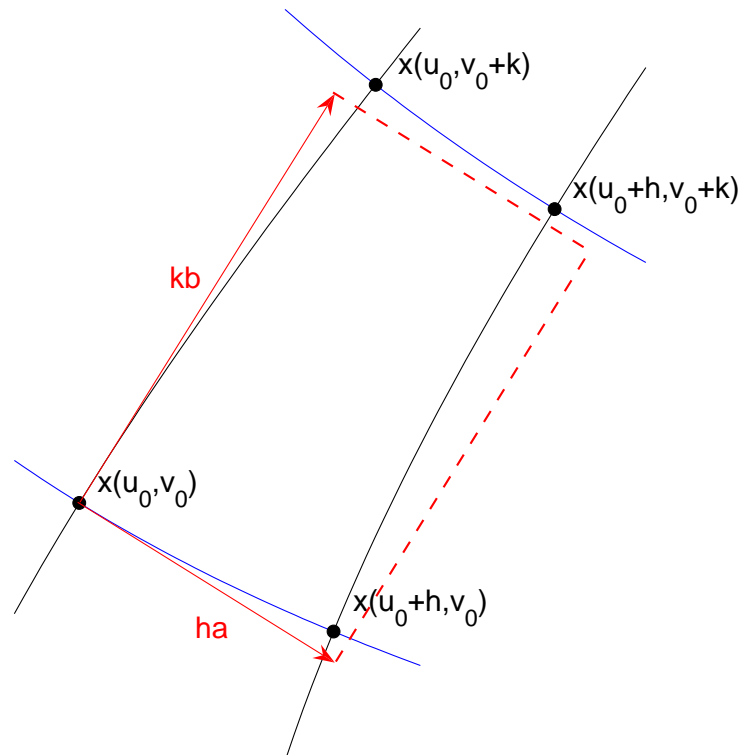
et ε est une fonction continue sur $\Lambda \times U_\delta(0, 0)$ telle que $\varepsilon(u, v; 0, 0) = 0$. En posant

$$\Delta(h, k) = \max_{s, t \in [0, 1], (u, v) \in \Lambda} |\varepsilon(u, v; sh, tk)|,$$

on remarque que $\Delta(h, k) \rightarrow 0$ lorsque $(h, k) \rightarrow (0, 0)$, qu'on peut écrire

$$\|x(u + sh, v + tk) - (sha + tkb + c)\| \leq \|(h, k)\| \Delta(h, k), \tag{1.40}$$

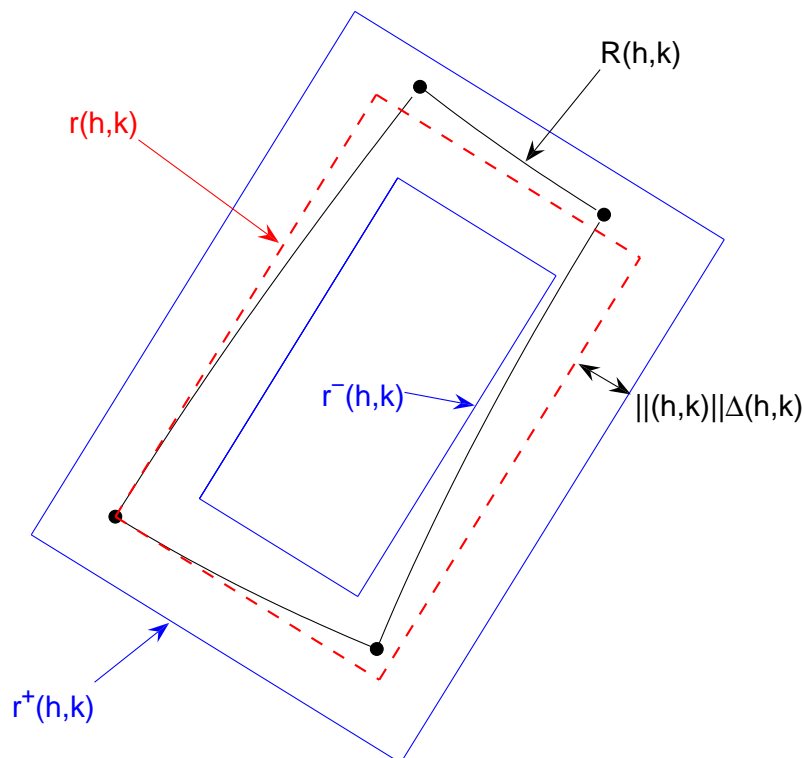
et que $r(h, k) = \{sha + tkb + c \mid s, t \in [0, 1]\}$ est un vrai parallélogramme (voir l'exemple 1.43).



L'estimation (1.40) montre en particulier que tout point

$$x(u + sh, v + tk), \quad (s, t) \in ([0, 1] \times \{0, 1\}) \cup (\{0, 1\} \times [0, 1]),$$

du bord de $R(h, k)$ se trouve à une distance inférieure ou égale à $\|(h, k)\| \Delta(h, k)$ du point correspondant $shu + tkb + c$ du bord du parallélogramme $r(h, k)$. On en déduit que $R(h, k)$ est inclus dans un parallélogramme $r^+(h, k)$ et contient un parallélogramme $r^-(h, k)$ comme illustré dans la figure suivante, la distance entre les côtés de $r(h, k)$ et ceux de $r^\pm(h, k)$ étant $\|(h, k)\| \Delta(h, k)$.



On en déduit d'une part que $A^-(h, k) \leq A(h, k) \leq A^+(h, k)$, $A^\pm(h, k)$ désignant l'aire de $r^\pm(h, k)$. D'autre part, on se convainc aisément qu'il existe une constante C telle que l'aire $A^0(h, k)$ de $r(h, k)$ satisfasse

$$|A^0(h, k) - A^\pm(h, k)| \leq C\|(h, k)\|^2 \Delta(h, k).$$

Il en résulte que

$$A^-(h, k) - A^0(h, k) \leq A(h, k) - A^0(h, k) \leq A^+(h, k) - A^0(h, k),$$

et donc

$$|A(h, k) - A^0(h, k)| \leq C\|(h, k)\|^2 \Delta(h, k).$$

Comme nous l'avons remarqué dans l'exemple 1.43, $A^0(h, k)$ est la valeur absolue du déterminant de la matrice dont les colonnes sont $h\mathbf{a}$ et $k\mathbf{b}$, c'est-à-dire

$$A^0(h, k) = \left| \det \begin{pmatrix} h \frac{\partial x_1}{\partial u}(u, v) & k \frac{\partial x_1}{\partial v}(u, v) \\ h \frac{\partial x_2}{\partial u}(u, v) & k \frac{\partial x_2}{\partial v}(u, v) \end{pmatrix} \right| = |J_{\mathbf{x}(u, v)}| |hk|,$$

d'où l'on déduit

$$|A(h, k) - |J_{\mathbf{x}(u, v)}| |hk|| \leq C\|(h, k)\|^2 \Delta(h, k).$$

□

1.2.4.2 Fonctions composées, changement de variables

Théorème 1.45 Soit F une transformation différentiable en $\mathbf{a} \in \mathbb{R}^2$. Si le champ scalaire f est différentiable en $F(\mathbf{a})$ alors $f \circ F$ est différentiable en \mathbf{a} et sa différentielle est donnée par

$$d(f \circ F)_\mathbf{a} = df_{F(\mathbf{a})} \circ dF_\mathbf{a}. \tag{1.41}$$

La formule (1.41) peut s'écrire de manière plus explicite en termes de dérivées partielles. Si $F(y_1, y_2) = (x_1(y_1, y_2), x_2(y_1, y_2))$ alors

$$\frac{\partial (f \circ F)}{\partial y_j}(\mathbf{a}) = \frac{\partial f}{\partial x_1}(F(\mathbf{a})) \frac{\partial x_1}{\partial y_j}(\mathbf{a}) + \frac{\partial f}{\partial x_2}(F(\mathbf{a})) \frac{\partial x_2}{\partial y_j}(\mathbf{a}), \tag{1.42}$$

pour $j = 1, 2$.

De même, si X est un champ vectoriel ou une transformation différentiable en $F(\mathbf{a})$, on peut appliquer le théorème 1.45 à chacune de ses composantes. On obtient ainsi la formule

$$d(X \circ F)_\mathbf{a} = dX_{F(\mathbf{a})} \circ dF_\mathbf{a},$$

ou encore

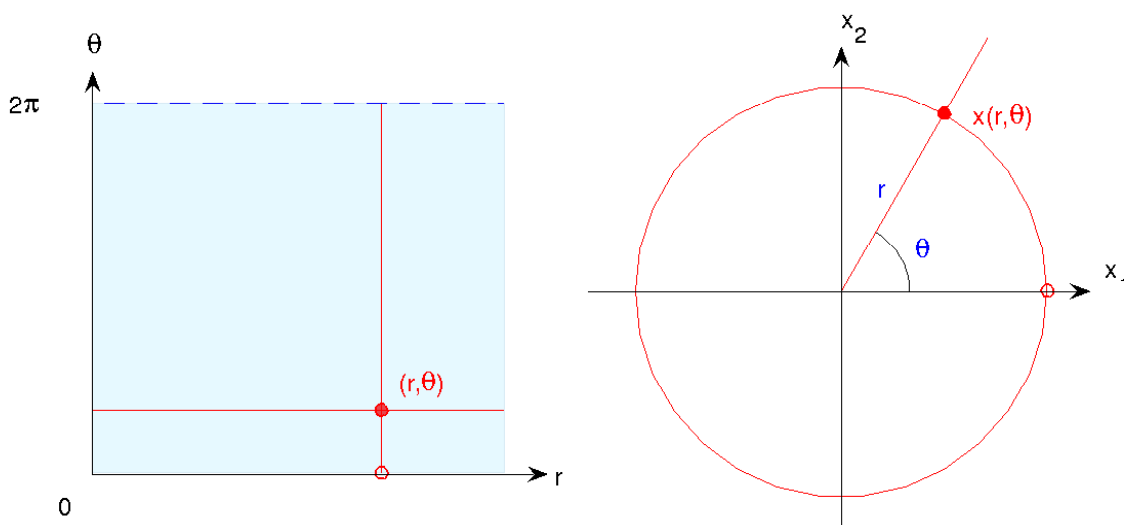
$$\frac{\partial (X_i \circ F)}{\partial y_j}(\mathbf{a}) = \frac{\partial X_i}{\partial x_1}(F(\mathbf{a})) \frac{\partial x_1}{\partial y_j}(\mathbf{a}) + \frac{\partial X_i}{\partial x_2}(F(\mathbf{a})) \frac{\partial x_2}{\partial y_j}(\mathbf{a}),$$

pour $i, j = 1, 2$.

Exemple 1.44 On dit que $(r, \theta) \in]0, \infty[\times]0, 2\pi[$ sont les **coordonnées polaires** du point

$$\mathbf{x}(r, \theta) = (x_1(r, \theta), x_2(r, \theta)) = (r \cos \theta, r \sin \theta) \in \mathbb{R}^2 \setminus \{(0, 0)\}.$$

Notez que les coordonnées polaires du point $(0, 0)$ ne sont pas bien définies puisque $\mathbf{x}(0, \theta) = (0, 0)$ pour tout θ . Les coordonnées polaires sont singulières à l'origine.



Soit $f(x_1, x_2)$ un champ scalaire sur le plan \mathbb{R}^2 . Dans certaines circonstances, en particulier en présence d'une symétrie de rotation autour de l'origine, il est plus commode de travailler avec des coordonnées polaires (r, θ) qu'avec les coordonnées cartésiennes (x_1, x_2) . En coordonnées polaires, le champ f est décrit par la fonction $\bar{f}(r, \theta) = f(r \cos \theta, r \sin \theta)$, c'est-à-dire $\bar{f} = f \circ \chi$. Par abus de notation, on omet souvent de distinguer les fonctions \bar{f} et f et on écrit $f(r, \theta)$ au lieu de $\bar{f}(r, \theta)$ pour désigner l'expression du champ f dans les coordonnées (r, θ) . Par exemple, le champ f défini en coordonnées cartésiennes sur $\mathbb{R}^2 \setminus \{0\}$ par la formule

$$f(x_1, x_2) = \frac{2x_1x_2}{x_1^2 + x_2^2}, \tag{1.43}$$

prend une forme beaucoup plus simple en coordonnées polaires

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

Etant donnée l'expression $f(r, \theta)$ du champ en coordonnées polaires, comment calcule-t-on son gradient

$$(\nabla f)(r, \theta) = \left(\frac{\partial f}{\partial x_1}(r \cos \theta, r \sin \theta), \frac{\partial f}{\partial x_2}(r \cos \theta, r \sin \theta) \right).$$

Pour répondre à cette question, on applique la formule (1.42) pour obtenir

$$f_r(r, \theta) = f_{x_1}(x(r, \theta)) \frac{\partial x_1}{\partial r}(r, \theta) + f_{x_2}(x(r, \theta)) \frac{\partial x_2}{\partial r}(r, \theta) = f_{x_1}(x(r, \theta)) \cos \theta + f_{x_2}(x(r, \theta)) \sin \theta,$$

et

$$f_\theta(r, \theta) = f_{x_1}(x(r, \theta)) \frac{\partial x_1}{\partial \theta}(r, \theta) + f_{x_2}(x(r, \theta)) \frac{\partial x_2}{\partial \theta}(r, \theta) = -f_{x_1}(x(r, \theta)) r \sin \theta + f_{x_2}(x(r, \theta)) r \cos \theta.$$

Ces relations peuvent s'écrire sous forme matricielle

$$\begin{pmatrix} f_r(r, \theta) \\ f_\theta(r, \theta) \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{pmatrix} \begin{pmatrix} f_{x_1}(x(r, \theta)) \\ f_{x_2}(x(r, \theta)) \end{pmatrix}$$

On obtient une expression du gradient du champ f en coordonnées polaires en résolvant le système linéaire ainsi obtenu pour les composantes f_{x_1} et f_{x_2} de ce gradient,

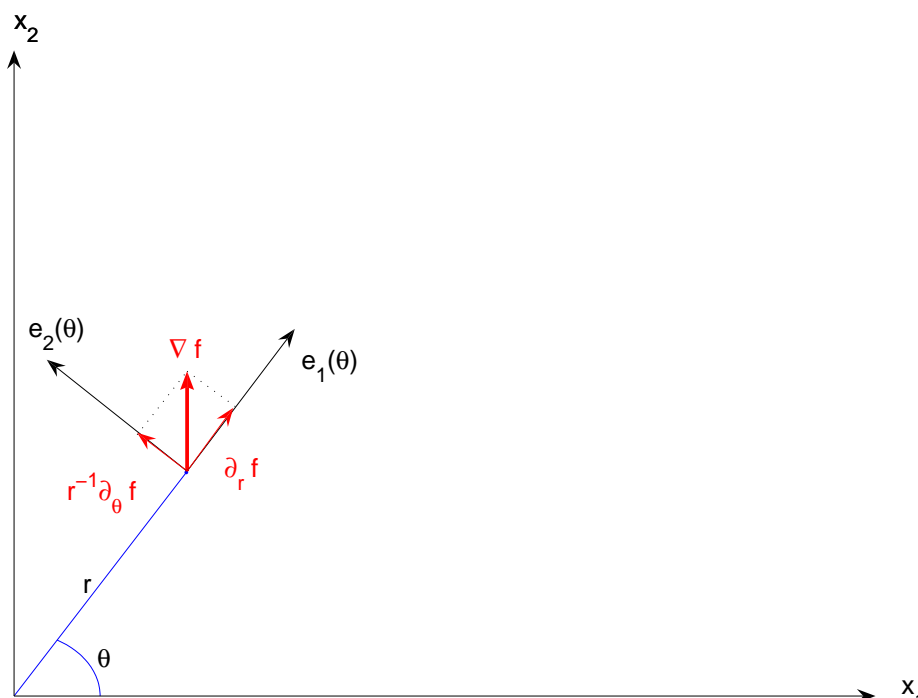
$$\begin{pmatrix} f_{x_1}(x(r, \theta)) \\ f_{x_2}(x(r, \theta)) \end{pmatrix} = \begin{pmatrix} \cos \theta & -r^{-1} \sin \theta \\ \sin \theta & r^{-1} \cos \theta \end{pmatrix} \begin{pmatrix} f_r(r, \theta) \\ f_\theta(r, \theta) \end{pmatrix}, \tag{1.44}$$

c'est-à-dire

$$(\nabla f)(r, \theta) = \frac{\partial f}{\partial r}(r, \theta) e_1(\theta) + \frac{1}{r} \frac{\partial f}{\partial \theta}(r, \theta) e_2(\theta). \tag{1.45}$$

où nous avons introduit les vecteurs orthonormés

$$e_1(\theta) = (\cos \theta, \sin \theta) \quad e_2(\theta) = (-\sin \theta, \cos \theta).$$



Remarquez l'apparition du facteur $1/r$ dans ces formules, une conséquence de la singularité des coordonnées polaires à l'origine.

Avec le champ (1.43) on obtient par exemple

$$(\nabla f)(r, \theta) = \frac{\partial \sin(2\theta)}{\partial r} e_1(\theta) + \frac{1}{r} \frac{\partial \sin(2\theta)}{\partial \theta} e_2(\theta) = \frac{2 \cos(2\theta)}{r} e_2(\theta).$$

◁

Intermède : le calcul des opérateurs différentiels. Voici une autre façon très pratique d'exprimer l'identité (1.44)

$$\begin{aligned} \frac{\partial}{\partial x_1} &= \cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta}, \\ \frac{\partial}{\partial x_2} &= \sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta}. \end{aligned} \tag{1.46}$$

Pour interpréter ces relations, il faut en appliquer le membre de gauche à l'expression d'un champ scalaire f en coordonnées cartésiennes (x_1, x_2) et le membre de droite à l'expression du même champ en coordonnées polaires (r, θ) . Chaque membre de ces relations est ce qu'on appelle un **opérateur différentiel**. Le laplacien en deux variables

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2},$$

ou en trois

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} + \frac{\partial^2}{\partial x_3^2},$$

sont des exemples d'opérateurs différentiels. En voici un autre

$$L = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1}.$$

De manière générale, un opérateur différentiel est une expression du type

$$D = \sum_{n_1, n_2} a_{n_1 n_2}(x_1, x_2) \frac{\partial^{n_1+n_2}}{\partial x_1^{n_1} \partial x_2^{n_2}}.$$

qui agit naturellement sur des fonctions suffisamment différentiables, $D : f \mapsto Df$, selon la formule

$$(Df)(x_1, x_2) = \sum_{n_1, n_2} a_{n_1, n_2}(x_1, x_2) \frac{\partial^{n_1+n_2} f}{\partial x_1^{n_1} \partial x_2^{n_2}}(x_1, x_2).$$

Le gradient est un opérateur différentiel à valeur vectorielle

$$\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2} \right).$$

Un opérateur différentiel est clairement **linéaire**, pour tous $a, b \in \mathbb{R}$ et toutes fonctions f, g suffisamment différentiables,

$$D(af + bg) = aDf + bDg.$$

Un opérateur différentiel est donc une application linéaire sur tout espace vectoriel de fonctions suffisamment différentiables. Par exemple le laplacien est une application linéaire de l'espace vectoriel des fonctions de classe C^2 dans l'espace vectoriel des fonctions continues ou de l'espace vectoriel des fonctions de classe C^∞ dans lui-même.

On peut faire des calculs algébriques avec des opérateurs différentiels, par exemple

$$\begin{aligned} L^2 &= \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \right) \left(x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} \right) \\ &= \left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_1 \frac{\partial}{\partial x_2} \right) - \left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_2 \frac{\partial}{\partial x_1} \right) \\ &\quad - \left(x_2 \frac{\partial}{\partial x_1} \right) \left(x_1 \frac{\partial}{\partial x_2} \right) + \left(x_2 \frac{\partial}{\partial x_1} \right) \left(x_2 \frac{\partial}{\partial x_1} \right). \end{aligned}$$

Pour calculer un produit comme

$$\left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_2 \frac{\partial}{\partial x_1} \right)$$

il faut tenir compte du fait qu'une dérivée agit sur tout ce qui se trouve à sa droite, en effet

$$\begin{aligned} \left(x_1 \frac{\partial}{\partial x_2} x_2 \frac{\partial}{\partial x_1} f \right) (x_1, x_2) &= x_1 \frac{\partial}{\partial x_2} \left(x_2 \frac{\partial f}{\partial x_1} \right) (x_1, x_2) \\ &= x_1 \frac{\partial x_2}{\partial x_2} (x_1, x_2) \frac{\partial f}{\partial x_1} (x_1, x_2) + x_1 x_2 \frac{\partial^2 f}{\partial x_2 \partial x_1} (x_1, x_2) \\ &= x_1 \frac{\partial f}{\partial x_1} (x_1, x_2) + x_1 x_2 \frac{\partial^2 f}{\partial x_2 \partial x_1} (x_1, x_2), \end{aligned}$$

c'est-à-dire que

$$\left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_2 \frac{\partial}{\partial x_1} \right) = x_1 \frac{\partial}{\partial x_1} + x_1 x_2 \frac{\partial^2}{\partial x_2 \partial x_1}.$$

Notez en particulier que la multiplication n'est pas commutative

$$\left(x_2 \frac{\partial}{\partial x_1} \right) \left(x_1 \frac{\partial}{\partial x_2} \right) = x_2 \frac{\partial}{\partial x_2} + x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2} \neq \left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_2 \frac{\partial}{\partial x_1} \right).$$

On peut donc terminer le calcul de L^2 comme suit

$$\begin{aligned} L^2 &= \left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_1 \frac{\partial}{\partial x_2} \right) - \left(x_1 \frac{\partial}{\partial x_2} \right) \left(x_2 \frac{\partial}{\partial x_1} \right) \\ &\quad - \left(x_2 \frac{\partial}{\partial x_1} \right) \left(x_1 \frac{\partial}{\partial x_2} \right) + \left(x_2 \frac{\partial}{\partial x_1} \right) \left(x_2 \frac{\partial}{\partial x_1} \right) \\ &= x_1^2 \frac{\partial^2}{\partial x_2^2} - x_1 \frac{\partial}{\partial x_1} - x_1 x_2 \frac{\partial^2}{\partial x_2 \partial x_1} \\ &\quad - x_2 \frac{\partial}{\partial x_2} - x_1 x_2 \frac{\partial^2}{\partial x_1 \partial x_2} + x_2^2 \frac{\partial^2}{\partial x_1^2}. \end{aligned}$$

Lorsqu'on fait agir cet opérateur sur des fonctions de classe C^2 , les deux dérivées mixtes sont égales (par le théorème 1.32). On peut alors écrire

$$L^2 = x_1^2 \frac{\partial^2}{\partial x_2^2} - 2x_1 x_2 \frac{\partial^2}{\partial x_2 \partial x_1} + x_2^2 \frac{\partial^2}{\partial x_1^2} - x_1 \frac{\partial}{\partial x_1} - x_2 \frac{\partial}{\partial x_2}.$$

A l'aide des formules (1.46) on peut exprimer l'opérateur L en coordonnées polaires

$$L = x_1 \frac{\partial}{\partial x_2} - x_2 \frac{\partial}{\partial x_1} = r \cos \theta \left(\sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta} \right) - r \sin \theta \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) = \frac{\partial}{\partial \theta}.$$

Le calcul de L^2 en coordonnées polaires est donc beaucoup plus simple qu'en coordonnées cartésiennes,

$$L^2 = \frac{\partial^2}{\partial \theta^2}.$$

Exemple 1.45 Pour calculer le laplacien d'un champ scalaire en coordonnées polaires

$$(\Delta f)(r, \theta) = \frac{\partial^2 f}{\partial x_1^2}(r \cos \theta, r \sin \theta) + \frac{\partial^2 f}{\partial x_2^2}(r \cos \theta, r \sin \theta).$$

on applique les formules (1.46). On a d'une part

$$\begin{aligned} \frac{\partial^2}{\partial x_1^2} &= \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) \\ &= \cos^2 \theta \frac{\partial^2}{\partial r^2} + \frac{\sin \theta \cos \theta}{r^2} \frac{\partial}{\partial \theta} - \frac{\sin \theta \cos \theta}{r} \frac{\partial^2}{\partial r \partial \theta} \\ &+ \frac{\sin^2 \theta}{r} \frac{\partial}{\partial r} + \frac{\sin \theta \cos \theta}{r^2} \frac{\partial}{\partial \theta} + \frac{\sin^2 \theta}{r^2} \frac{\partial^2}{\partial \theta^2}, \end{aligned}$$

et d'autre part

$$\begin{aligned} \frac{\partial^2}{\partial x_2^2} &= \left(\sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta} \right) \left(\sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta} \right) \\ &= \sin^2 \theta \frac{\partial^2}{\partial r^2} - \frac{\sin \theta \cos \theta}{r^2} \frac{\partial}{\partial \theta} + \frac{\sin \theta \cos \theta}{r} \frac{\partial^2}{\partial r \partial \theta} \\ &+ \frac{\cos^2 \theta}{r} \frac{\partial}{\partial r} - \frac{\sin \theta \cos \theta}{r^2} \frac{\partial}{\partial \theta} + \frac{\cos^2 \theta}{r^2} \frac{\partial^2}{\partial \theta^2}. \end{aligned}$$

En additionnant ces deux formules on obtient l'expression recherchée

$$\Delta = \frac{\partial^2}{\partial x_1^2} + \frac{\partial^2}{\partial x_2^2} = \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2} = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \theta^2}. \quad (1.47)$$

◁

Exemple 1.46 Le champ de vitesse $V(x)$ d'un flot stationnaire incompressible satisfait l'équation de continuité

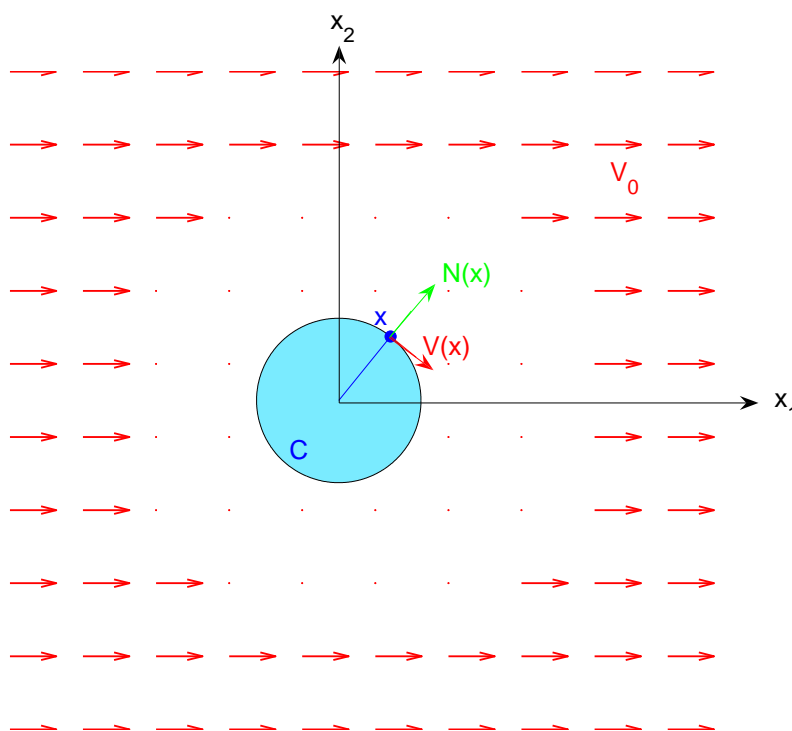
$$\operatorname{div} V(x) = \frac{\partial V_1}{\partial x_1}(x_1, x_2, x_3) + \frac{\partial V_2}{\partial x_2}(x_1, x_2, x_3) + \frac{\partial V_3}{\partial x_3}(x_1, x_2, x_3) = 0.$$

Si le flot est irrotationnel, ce champ dérive d'un potentiel $V(x) = \nabla \varphi(x)$. Le potentiel φ satisfait alors l'équation de Laplace

$$\operatorname{div} V(x) = \operatorname{div} \nabla \varphi(x) = \Delta \varphi(x) = 0.$$

Le potentiel d'un flot constant $V_0(x) = (v_0, 0, 0)$ est clairement donné par $\varphi_0(x_1, x_2, x_3) = v_0 x_1$. On cherche à décrire comment ce flot est perturbé par la présence d'un obstacle cylindrique infini placé le long de l'axe x_3

$$C = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 \leq R^2\}.$$



En supposant que le flot perturbé $V(x)$ reste irrotationnel, son potentiel $\varphi(x)$ satisfait l'équation de Laplace à l'extérieur du cylindre C ,

$$\Delta\varphi(x_1, x_2, x_3) = 0 \text{ pour } x_1^2 + x_2^2 > R^2. \quad (1.48)$$

La vitesse $V(x)$ doit être tangente au cylindre sur sa surface. Si $x = (x_1, x_2, x_3)$ est un point sur la surface du cylindre alors $x_1^2 + x_2^2 = R^2$ et le vecteur normal à la surface en ce point est donné par

$$N(x) = \frac{1}{R} \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix}.$$

La vitesse $V(x) = \nabla\varphi(x)$ est tangente à la surface si et seulement si sa composante normale s'annule, c'est-à-dire

$$N(x) \cdot V(x) = \frac{x_1}{R} \frac{\partial\varphi}{\partial x_1}(x_1, x_2, x_3) + \frac{x_2}{R} \frac{\partial\varphi}{\partial x_2}(x_1, x_2, x_3) = 0 \text{ pour } x_1^2 + x_2^2 = R^2. \quad (1.49)$$

Finalement, lorsqu'on s'éloigne du cylindre C la vitesse $V(x)$ doit s'approcher de la vitesse non perturbée $V_0(x)$,

$$\lim_{x_1^2 + x_2^2 \rightarrow \infty} V(x) = \lim_{x_1^2 + x_2^2 \rightarrow \infty} \nabla\varphi(x) = \begin{pmatrix} v_0 \\ 0 \\ 0 \end{pmatrix}. \quad (1.50)$$

Notre problème consiste à résoudre l'équation aux dérivées partielles (1.48) avec les conditions aux limites (1.49) et (1.50).

La symétrie de translation du problème dans la direction x_3 nous permet de conclure que la vitesse $V(x)$ ne dépend pas de la coordonnée x_3 et que sa troisième composante $V_3(x)$ est toujours nulle. Nous pouvons donc réduire le problème au plan x_1, x_2 . Afin de prendre en compte la nature perturbative du problème, il est judicieux de poser

$$\varphi(x) = \varphi_0(x) + \psi(x) = v_0 x_1 + \psi(x_1, x_2),$$

soit, en coordonnées polaires,

$$\varphi(r, \theta) = \varphi_0(r, \theta) + \psi(r, \theta) = v_0 r \cos \theta + \psi(r, \theta).$$

Comme $\Delta\varphi_0 = 0$, l'équation de Laplace (1.48) peut s'écrire $\Delta\psi(x) = 0$ en dehors de C . En utilisant la formule (1.47), cette équation devient

$$\Delta\psi(r, \theta) = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial \psi}{\partial r} + \frac{1}{r^2} \frac{\partial^2 \psi}{\partial \theta^2} = 0 \text{ pour } r > R. \quad (1.51)$$

Pour formuler la condition sur la surface de cylindre C , on note que les formules (1.46) impliquent

$$x_1 \frac{\partial}{\partial x_1} + x_2 \frac{\partial}{\partial x_2} = r \cos \theta \left(\cos \theta \frac{\partial}{\partial r} - \frac{\sin \theta}{r} \frac{\partial}{\partial \theta} \right) + r \sin \theta \left(\sin \theta \frac{\partial}{\partial r} + \frac{\cos \theta}{r} \frac{\partial}{\partial \theta} \right) = r \frac{\partial}{\partial r},$$

si bien qu'on peut écrire la condition (1.49) comme

$$\frac{\partial \varphi}{\partial r}(R, \theta) = \frac{\partial \varphi_0}{\partial r}(R, \theta) + \frac{\partial \psi}{\partial r}(R, \theta) = v_0 \cos \theta + \frac{\partial \psi}{\partial r}(R, \theta) = 0,$$

soit

$$\frac{\partial \psi}{\partial r}(R, \theta) = -v_0 \cos \theta. \quad (1.52)$$

Finalement, la condition (1.50) devient

$$\lim_{r \rightarrow \infty} \nabla \psi(r, \theta) = 0,$$

ce qu'on peut écrire, à l'aide de la relation (1.45),

$$\lim_{r \rightarrow \infty} \frac{\partial \psi}{\partial r}(r, \theta) = \lim_{r \rightarrow \infty} \frac{1}{r} \frac{\partial \psi}{\partial \theta}(r, \theta) = 0. \quad (1.53)$$

Pour résoudre le problème (1.51), (1.52), (1.53), on fait maintenant un Ansatz de séparation des variables

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

La condition (1.52) s'écrit alors $f'(R)g(\theta) = -v_0 \cos \theta$ d'où on conclut que

$$g(\theta) = -\frac{v_0}{f'(R)} \cos \theta.$$

L'équation de Laplace (1.51) devient

$$\frac{1}{r} (rf')' \left(-\frac{v_0}{f'(R)} \cos \theta \right) + \frac{1}{r^2} f \left(\frac{v_0}{f'(R)} \cos \theta \right) = 0,$$

soit

$$r(rf')' = f. \quad (1.54)$$

En faisant l'Ansatz $f(r) = r^\ell$ on a $rf' = \ell f$ et le membre de gauche de (1.54) devient

$$r(rf')' = r(\ell f)' = \ell^2 f.$$

On tire donc de (1.54) que $\ell^2 = 1$ et on obtient ainsi deux solutions $f_1(r) = r^{-1}$ et $f_2(r) = r$. Nous étudierons les équations différentielles au chapitre 3. En particulier nous montrerons dans l'exemple 3.6 que toute solution de (1.54) est de la forme

$$f(r) = \frac{A}{r} + Br,$$

où A et B sont des constantes qui restent à déterminer. Comme

$$\begin{aligned} \frac{\partial \psi}{\partial r} &= \left(-\frac{A}{r^2} + B \right) \left(-\frac{v_0}{-AR^{-2} + B} \cos \theta \right), \\ \frac{1}{r} \frac{\partial \psi}{\partial \theta} &= \left(\frac{A}{r^2} + B \right) \left(\frac{v_0}{-AR^{-2} + B} \sin \theta \right), \end{aligned}$$

la condition (1.53) nous permet de conclure que $B = 0$. Dans ce cas on remarque que la constante A disparaît lorsqu'on multiplie $f(r)$ par $g(\theta)$. On peut donc, sans restriction de la généralité, poser $A = 1$. Nous avons ainsi obtenu la solution de notre problème

$$\psi(r, \theta) = \frac{R^2 v_0}{r} \cos \theta,$$

d'où l'on déduit

$$\varphi(r, \theta) = \varphi_0(r, \theta) + \psi(r, \theta) = v_0 \left(r + \frac{R^2}{r} \right) \cos \theta.$$

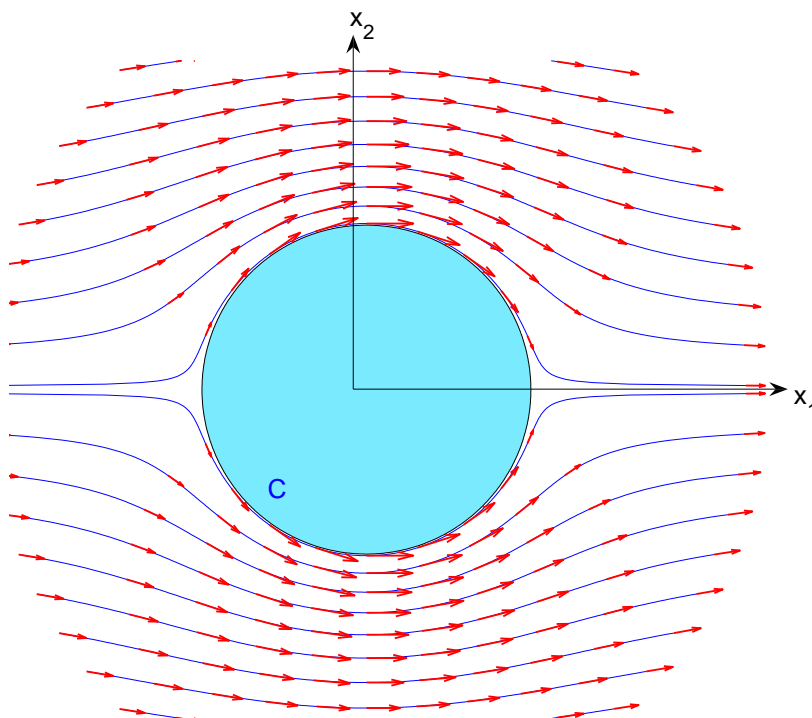
En utilisant encore la relation (1.45),

$$V(r, \theta) = \nabla \varphi(r, \theta) = v_0 \left(1 - \frac{R^2}{r^2} \right) \cos(\theta) \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} - v_0 \left(1 + \frac{R^2}{r^2} \right) \sin \theta \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix},$$

et après quelques manipulations élémentaires

$$V(r, \theta) = v_0 \begin{pmatrix} 1 - \frac{R^2}{r^2} \cos 2\theta \\ -\frac{R^2}{r^2} \sin 2\theta \end{pmatrix}.$$

Ce champ ainsi que ses lignes de champ sont illustrés dans la figure suivante.



◁

Nous terminons cette section avec la démonstration du théorème 1.45 pour laquelle nous avons besoin d'une propriété élémentaire des applications linéaires.

Lemme 1.46 Si $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ est linéaire, il existe une constante C telle que $\|A(x)\| \leq C \|x\|$ pour tout $x \in \mathbb{R}^n$.

Démonstration. Soit e_1, \dots, e_n la base canonique de \mathbb{R}^n . L'inégalité triangulaire appliquée à la formule

$$A(x) = \sum_{j=1}^n x_j A(e_j),$$

donne

$$\|A(x)\| \leq \sum_{j=1}^n |x_j| \|A(e_j)\|.$$

En interprétant cette somme comme un produit scalaire, nous pouvons invoquer l'inégalité de Cauchy-Schwarz pour obtenir

$$\|A(x)\| \leq \sqrt{\sum_{j=1}^n \|A(e_j)\|^2} \|x\|.$$

□

Preuve du théorème 1.45 Posons $b = F(a)$ et remarquons que $F(a + h) = b + k$ où $k = dF_a(h) + \varepsilon(h) \|h\|$ avec $\varepsilon(h) \rightarrow (0, 0)$ et $k \rightarrow (0, 0)$ lorsque $h \rightarrow (0, 0)$. On a donc

$f \circ F(a + h) - f \circ F(a) = f(b + k) - f(b) = df_b(k) + \delta(k) \|k\| = df_b \circ dF_a(h) + df_b(\varepsilon(h)) \|h\| + \delta(k) \|k\|$,
où $\delta(k) \rightarrow 0$ lorsque $k \rightarrow (0, 0)$. Il en résulte

$$\frac{f \circ F(a + h) - [f \circ F(a) + df_b \circ dF_a(h)]}{\|h\|} = df_b(\varepsilon(h)) + \delta(k) \frac{\|k\|}{\|h\|}. \quad (1.55)$$

En appliquant le lemme précédent à la différentielle dF_a , on montre qu'il existe une constante C telle que $\|k\| \leq C \|h\|$. On en déduit que le membre de droite de inégalité (1.55) tend vers 0 lorsque $h \rightarrow (0, 0)$. □

1.2.4.3 Le théorème de la fonction inverse

Considérons une transformation injective $F : \Omega \rightarrow \mathbb{R}^2$ et désignons par Ω' son image. Si l'application réciproque $F^{-1} : \Omega' \rightarrow \Omega$ est différentiable en a et si F est différentiable en $F^{-1}(a)$ on déduit de $F \circ F^{-1} = I$ (l'identité) et du fait que $dI_a = I$ (c.f. exemple 1.43, I est linéaire) que

$$dF_{F^{-1}(a)} \circ dF_a^{-1} = I.$$

Il en résulte que la différentielle $dF_{F^{-1}(a)}$ est injective et que

$$dF_a^{-1} = (dF_{F^{-1}(a)})^{-1}.$$

La différentielle de F a donc "hérité" de l'injectivité de F !

Il est beaucoup plus remarquable que la réciproque est vraie sous une hypothèse un peu plus forte : Si la différentielle de F est injective en un point, F est injective dans un voisinage de ce point. Plus précisément, on a le théorème suivant dont la démonstration dépasse le cadre de ce cours.

Théorème 1.47 (Théorème de la fonction inverse) *Si la transformation F est de classe C^1 sur l'ouvert $\Omega \subset \mathbb{R}^2$ et si sa différentielle dF_a en $a \in \Omega$ est injective (c'est-à-dire si son jacobien JF_a n'est pas nul), il existe un voisinage U de a et un voisinage V de $F(a)$ tel que F soit une bijection de U dans V . De plus la fonction réciproque F^{-1} est de classe C^1 sur U et sa différentielle est donnée par*

$$dF_y^{-1} = (dF_{F^{-1}(y)})^{-1},$$

pour tout $y \in U$.

Exemple 1.47 La fonction $F(r, \theta) = (r \cos \theta, r \sin \theta)$ transforme le domaine rectangulaire $\Omega =]0, R[\times]-\pi/3, \pi/3[$ dans le secteur circulaire

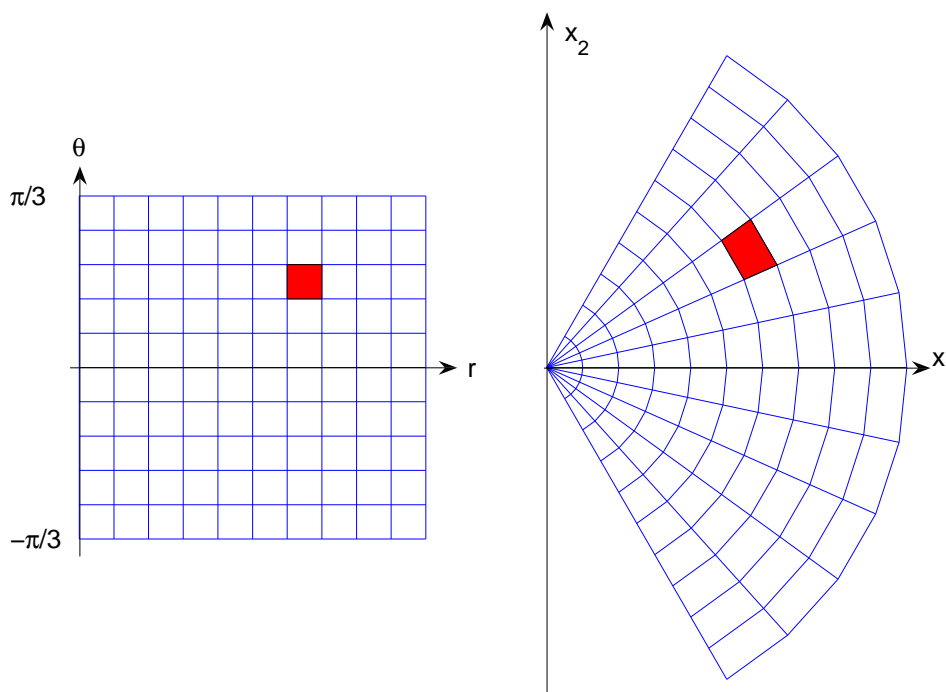
$$\Omega' = \{x = (r \cos \theta, r \sin \theta) \mid 0 < r < R, -\pi/3 < \theta < \pi/3\}.$$

Sa matrice jacobienne est

$$dF_{(r,\theta)} = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}, \quad (1.56)$$

et son jacobien $JF_{(r,\theta)} = r > 0$ sur Ω . Le théorème de la fonction inverse nous permet donc de conclure que tout point de Ω possède un voisinage sur lequel F est injective. On dit dans ce cas que F est **localement injective** sur Ω . La différentielle de l'application réciproque est

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



En fait F est globalement injective sur Ω et on vérifie facilement que

$$G(x_1, x_2) = \left(\sqrt{x_1^2 + x_2^2}, \arctan \left(\frac{x_2}{x_1} \right) \right), \tag{1.57}$$

en est l'application réciproque. En effet on a d'une part, pour $r > 0$ et $|\theta| < \pi/2$,

$$G \circ F(r, \theta) = \left(\sqrt{r^2 \cos^2 \theta + r^2 \sin^2 \theta}, \arctan \left(\frac{r \sin \theta}{r \cos \theta} \right) \right) = (\sqrt{r^2}, \arctan(\tan \theta)) = (r, \theta).$$

D'autre part, à l'aide des formules

$$\cos(\arctan u) = \frac{1}{\sqrt{1+u^2}}, \quad \sin(\arctan u) = \frac{u}{\sqrt{1+u^2}}, \tag{1.58}$$

on obtient pour $x_1 > 0$,

$$\begin{aligned} F \circ G(x_1, x_2) &= \left(\sqrt{x_1^2 + x_2^2} \cos \left(\arctan \left(\frac{x_2}{x_1} \right) \right), \sqrt{x_1^2 + x_2^2} \sin \left(\arctan \left(\frac{x_2}{x_1} \right) \right) \right) \\ &= \left(\frac{\sqrt{x_1^2 + x_2^2}}{\sqrt{1 + (x_2/x_1)^2}}, \frac{x_2 \sqrt{x_1^2 + x_2^2}}{x_1 \sqrt{1 + (x_2/x_1)^2}} \right) \\ &= \left(|x_1|, x_2 \frac{|x_1|}{x_1} \right) = (x_1, x_2). \end{aligned}$$

◁

Une transformation $F : \Omega \rightarrow \mathbb{R}^2$ peut être localement injective sur Ω sans pour autant être injective sur Ω .

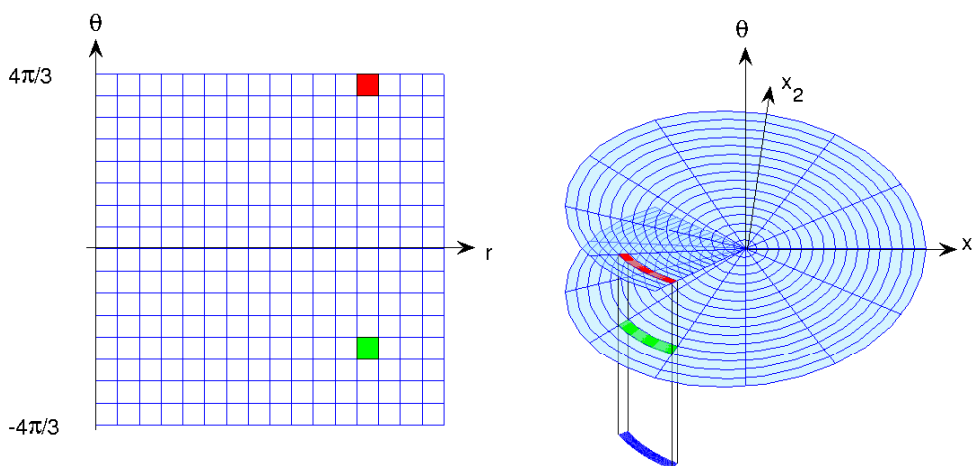
Exemple 1.48 On reprend la fonction $F(r, \theta) = (r \cos \theta, r \sin \theta)$ de l'exemple précédent mais cette fois sur le domaine rectangulaire $\Omega =]0, \mathbb{R}[\times] - 4\pi/3, 4\pi/3[$. Sa matrice jacobienne est toujours donnée par la formule (1.56). On peut donc conclure, comme dans l'exemple précédent, que F est localement injective sur Ω . La restriction de F au domaine $]0, \mathbb{R}[\times] - \pi/2, \pi/2[$ admet la fonction G , définie en (1.57), comme réciproque (la démonstration est identique à celle de l'exemple 1.47!). On note cependant que cette fonction n'est pas définie si $x_1 = 0$. La fonction

$$G_+(x_1, x_2) = \left(\sqrt{x_1^2 + x_2^2}, \arctan \left(\frac{x_1 + x_2}{x_1 - x_2} + \frac{3\pi}{4} \right) \right),$$

est bien définie si $x_2 > x_1$. Vérifiez qu'elle est réciproque à F restreinte au rectangle $]0, \mathbb{R}[\times]\pi/4, 4\pi/3[$. De même

$$G_-(x_1, x_2) = \left(\sqrt{x_1^2 + x_2^2}, -\arctan \left(\frac{x_1 - x_2}{x_1 + x_2} + \frac{3\pi}{4} \right) \right),$$

est réciproque à la restriction de F à la région $]0, \mathbb{R}[\times] - 4\pi/3, -\pi/4[$.



◁

Cependant F n'est pas globalement injective! En effet, si $(r, \theta) \in]0, \mathbb{R}[\times]2\pi/3, 4\pi/3[\subset \Omega$ on a $(r, \theta - 2\pi) \in]0, \mathbb{R}[\times] - 4\pi/3, -2\pi/3[\subset \Omega$, $(r, \theta) \neq (r, \theta - 2\pi)$ mais $F(r, \theta) = F(r, \theta - 2\pi)$.

1.2.4.4 Le théorème de la fonction implicite

Si $b \neq 0$, l'équation $ax + by + c = 0$ définit une fonction affine $y = -(ax + c)/b$. Dans ce paragraphe nous allons généraliser ce fait aux équations du type $F(x, y) = 0$ où F est une fonction différentiable.

Une fonction $f(x)$ est **définie implicitement** près de $x = a$ par l'équation $F(x, y) = 0$ si toutes les solutions de cette équation dans un voisinage de $(a, f(a))$ sont sur le graphe $\{(x, y) \mid y = f(x)\}$ de f .

Exemple 1.49 Si $F(x, y) = x^2 + y^2 - 1$, l'équation $F(x, y) = 0$ est celle d'un cercle de rayon 1 centré en $(0, 0)$. Ce cercle n'est pas globalement le graphe d'une fonction, cependant l'équation $F(x, y) = 0$ peut être résolue explicitement pour y . On trouve les deux solutions $y = \pm\sqrt{1 - x^2}$. Les fonctions $f(x) = \sqrt{1 - x^2}$ et $g(x) = -\sqrt{1 - x^2}$ sont définies implicitement par l'équation $F(x, y) = 0$ près de $x = 1$.

◁

De nombreux problèmes d'analyse se réduisent à résoudre une équation du type $F(x, y) = 0$. Le corollaire du théorème de la fonction inverse que nous démontrons maintenant est donc particulièrement important.

Théorème 1.48 (Théorème de la fonction implicite) Soit $F : \Omega \rightarrow \mathbb{R}$ de classe C^1 dans un voisinage de $(a, b) \in \Omega \subset \mathbb{R}^2$. Si

$$F(a, b) = 0, \quad F_{x_2}(a, b) \neq 0,$$

il existe un intervalle ouvert $I \ni a$ et une fonction de classe C^1 $f : I \rightarrow \mathbb{R}$ telle que

$$f(a) = b, \quad F(x, f(x)) = 0, \quad f'(x) = -\frac{F_{x_1}(x, f(x))}{F_{x_2}(x, f(x))},$$

pour tout $x \in I$.

Démonstration. L'application

$$\Phi : (x_1, x_2) \mapsto (x_1, F(x_1, x_2)),$$

est de classe C^1 dans un voisinage de (a, b) . Sa différentielle en (a, b) est

$$d\Phi_{(a,b)} = \begin{pmatrix} 1 & 0 \\ F_{x_1}(a, b) & F_{x_2}(a, b) \end{pmatrix},$$

et son jacobien $J\Phi_{(a,b)} = F_{x_2}(a, b) \neq 0$. Φ est donc une bijection d'un voisinage U de (a, b) dans un voisinage V de $\Phi(a, b) = (a, 0)$. Pour $(\xi, \eta) \in V$ on pose $(x(\xi, \eta), \varphi(\xi, \eta)) = \Phi^{-1}(\xi, \eta)$. On a donc $(\xi, \eta) = \Phi \circ \Phi^{-1}(\xi, \eta) = \Phi(x(\xi, \eta), \varphi(\xi, \eta))$, c'est-à-dire

$$\xi = x(\xi, \eta), \quad \eta = F(\xi, \varphi(\xi, \eta)),$$

pour $(\xi, \eta) \in V$. De plus Φ^{-1} est de classe C^1 et la relation $d\Phi_{(\xi,\eta)}^{-1} = (d\Phi_{\Phi^{-1}(\xi,\eta)})^{-1}$ s'écrit

$$\begin{pmatrix} x_\xi(\xi, \eta) & x_\eta(\xi, \eta) \\ \varphi_\xi(\xi, \eta) & \varphi_\eta(\xi, \eta) \end{pmatrix} = \frac{1}{F_{x_2}(x(\xi, \eta), \varphi(\xi, \eta))} \begin{pmatrix} F_{x_2}(x(\xi, \eta), \varphi(\xi, \eta)) & 0 \\ -F_{x_1}(x(\xi, \eta), \varphi(\xi, \eta)) & 1 \end{pmatrix}.$$

Soit I un intervalle contenant a et tel que $I \times \{0\} \subset U$. Pour tout $\xi \in I$ on a alors $F(\xi, \varphi(\xi, 0)) = 0$. La fonction recherchée est donc $f(\xi) = \varphi(\xi, 0)$. Elle est de classe C^1 et sa dérivée est donnée par

$$f'(\xi) = \varphi_\xi(\xi, 0) = -\frac{F_{x_1}(\xi, f(\xi))}{F_{x_2}(\xi, f(\xi))}.$$

□

Voici une application simple mais typique du théorème de la fonction implicite.

Exemple 1.50 Soient f et g deux fonctions d'une variable de classe C^1 , f ayant un zéro simple en $x = 0$: $f(0) = 0$ et $f'(0) \neq 0$. Soit $f_\lambda(x) = f(x) + \lambda g(x)$. Si $\lambda \in \mathbb{R}$ est assez petit la fonction f_λ admet un zéro simple $x(\lambda)$ près de 0 et la fonction $\lambda \mapsto x(\lambda)$ est de classe C^1 .

Pour appliquer le théorème 1.48 on considère la fonction $F(\lambda, x) = f(x) + \lambda g(x)$ qui est de classe C^1 . Il suffit de remarquer que $F(0, 0) = 0$ et $F_x(0, 0) = f'(0) \neq 0$. On en déduit l'existence de la fonction $x(\lambda)$. Sa dérivée est

$$x'(\lambda) = -\frac{g(x(\lambda))}{f'(x(\lambda)) + \lambda g'(x(\lambda))}.$$

◁

Chapitre 2

Intégrales multiples

2.1 Sommes et intégrales de Riemann

Dans le chapitre ?? nous avons défini l'intégrale de Riemann d'une fonction continue sur un intervalle fermé $[a, b]$. La relation de Chasles,

$$\int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx,$$

qui exprime l'additivité de l'intégrale, permet d'étendre cette définition de manière consistante à toute réunion finie d'intervalles qui, pris deux à deux, n'ont pas plus d'un point commun. Bien que de tels sous-ensembles forment une classe suffisante pour traiter beaucoup de problèmes, tous les sous-ensembles de \mathbb{R} , même bornés, ne sont pas de ce type. En fait il est impossible de définir l'intégrale d'une fonction aussi simple que la fonction constante $f(x) = 1$ sur n'importe quel sous-ensemble borné de \mathbb{R} : c'est le **paradoxe de Banach-Tarski**. En particulier, comme

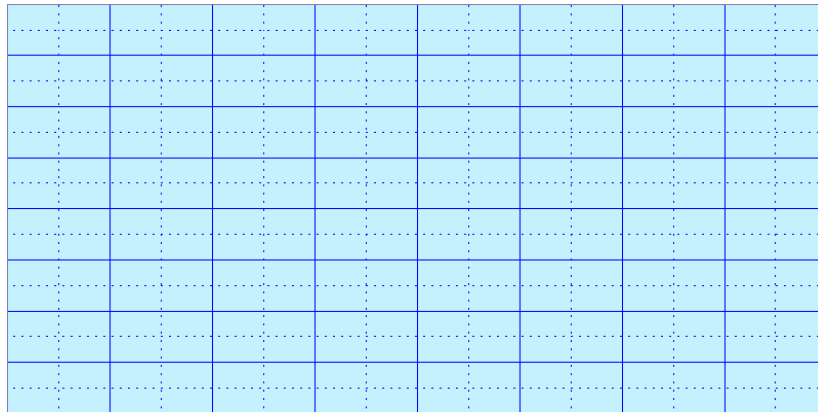
$$\int_a^b 1 dx = b - a$$

est la longueur de l'intervalle $[a, b]$, il est impossible d'assigner une longueur de manière consistante à tous les sous-ensembles de \mathbb{R} . De même on ne peut attribuer une aire (ou un volume) à tous les sous-ensembles bornés du plan (ou de l'espace). Les mathématiciens ont contourné ce problème en développant la théorie de la mesure qui généralise la définition de l'intégrale à une classe beaucoup plus étendue de sous-ensembles ainsi qu'à de nombreuses fonctions non continues : **l'intégrale de Lebesgue**. Ce sujet avancé n'est traité qu'en troisième année de la Licence de Mathématiques. Dans ce cours nous nous contenterons d'une approche pragmatique de l'intégration qui, bien que présentant de nombreuses déficiences, nous permettra au moins de calculer les intégrales apparaissant dans de nombreux problèmes élémentaires. En particulier nous ne chercherons pas à caractériser les fonctions "intégrables" et nous nous contenterons d'intégrer des fonctions continues. Nous traiterons en détail l'intégration des fonctions de deux variables dans la section 2.3 et très sommairement l'extension aux fonctions de trois variables dans la section 2.4.

Dans le plan, les intervalles fermés sont remplacés par des rectangles fermés $R = [a, b] \times [c, d]$ et les partitions d'un intervalle que nous avons utilisées au chapitre ?? par des partitions d'un rectangle en sous-rectangles. On notera $\mathcal{P}_n(R)$ la partition du rectangle R constituée des micro-rectangles

$$R_{ij}^{(n)} = \left[a + (i-1) \frac{(b-a)}{2^n}, a + i \frac{(b-a)}{2^n} \right] \times \left[c + (j-1) \frac{(d-c)}{2^n}, c + j \frac{(d-c)}{2^n} \right], \quad (2.1)$$

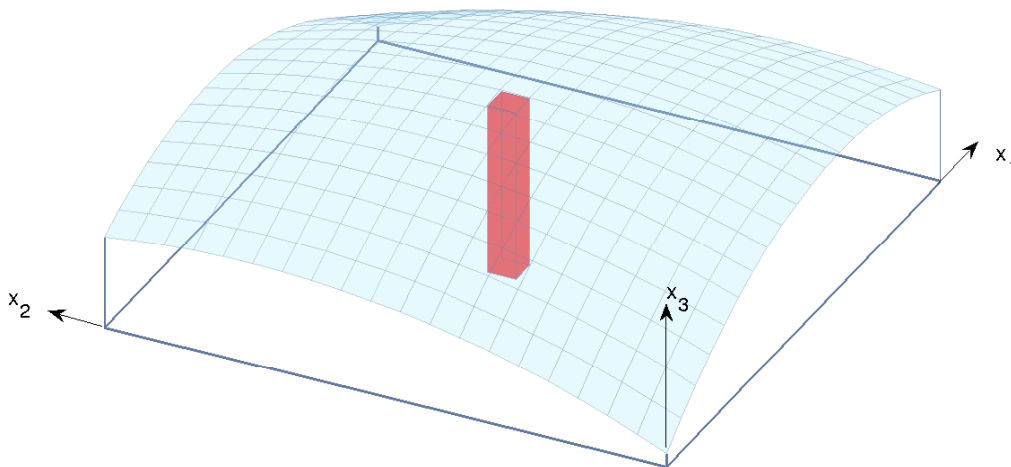
où $i, j = 1, \dots, 2^n$. Notez que la partition $\mathcal{P}_{n+1}(R)$ est obtenue à partir de $\mathcal{P}_n(R)$ en subdivisant chaque rectangle $R_{ij}^{(n)}$ en quatre rectangles congruents, comme illustré dans la figure suivante.



La somme de Riemann d'une fonction f , continue sur \mathbb{R} , associée à la partition $\mathcal{P}_n(\mathbb{R})$ est

$$\Sigma_{\mathcal{P}_n(\mathbb{R})}(f) = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} f \left(a + (i-1) \frac{(b-a)}{2^n}, c + (j-1) \frac{(d-c)}{2^n} \right) \frac{b-a}{2^n} \frac{d-c}{2^n}. \quad (2.2)$$

Dans cette expression, le sommand n'est autre que le volume du parallélépipède de base $\mathbb{R}_{ij}^{(n)}$ dont la hauteur est la valeur prise par f en un sommet de $\mathbb{R}_{ij}^{(n)}$.



La somme de Riemann est donc une approximation du volume de la région comprise entre le plan horizontal $x_3 = 0$ et le graphe de f , compté algébriquement, c'est-à-dire positivement pour la partie comprise en dessus du plan et négativement pour celle située en dessous (par la suite, nous parlerons simplement du **volume sous le graphe de f**). Une façon équivalente d'interpréter les sommes de Riemann (2.2) consiste à introduire la suite de fonctions f_n , constante sur chaque micro-rectangle de la partition $\mathcal{P}_n(\mathbb{R})$, et définie par

$$f_n(x_1, x_2) = f \left(a + (i-1) \frac{(b-a)}{2^n}, c + (j-1) \frac{(d-c)}{2^n} \right) \quad \text{si } (x_1, x_2) \in \mathbb{R}_{ij}^{(n)}.$$

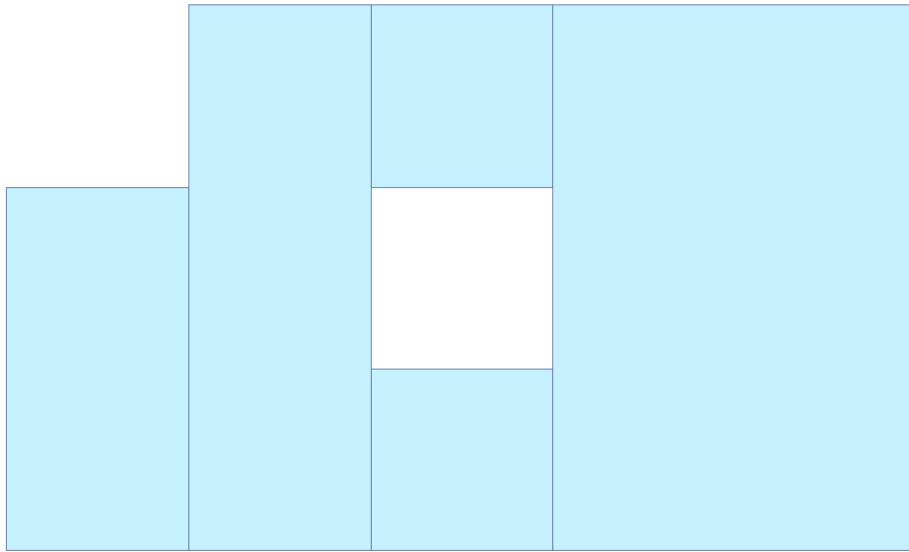
Pour chaque n , la fonction f_n est l'approximation de f obtenue en fixant la valeur de f_n sur chaque micro-rectangle $\mathbb{R}_{ij}^{(n)}$ à la valeur prise par f en un sommet de ce dernier. La somme $\Sigma_{\mathcal{P}_n(\mathbb{R})}(f)$ est alors exactement le volume sous le graphe de f_n .

On peut montrer que la somme de Riemann converge lorsque $n \rightarrow \infty$, c'est-à-dire lorsque la partition $\mathcal{P}_n(\mathbb{R})$ devient de plus en plus fine. Par définition,

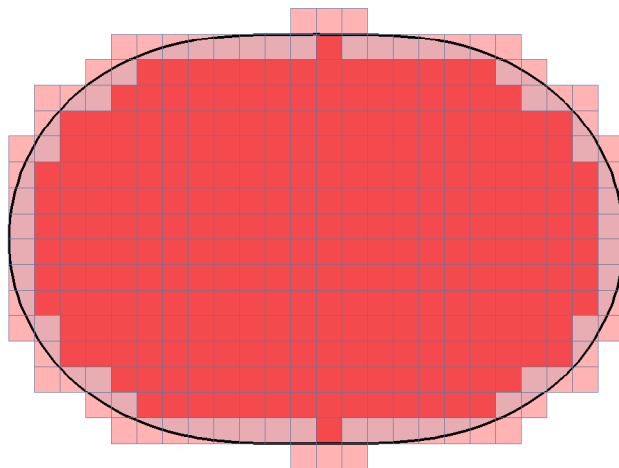
$$\iint_{\mathbb{R}} f(x_1, x_2) dx_1 dx_2 = \lim_{n \rightarrow \infty} \Sigma_{\mathcal{P}_n(\mathbb{R})}(f),$$

est l'**intégrale de Riemann** de f sur le rectangle \mathbb{R} . C'est aussi le volume sous le graphe de f . Pour distinguer une telle intégrale de l'intégrale d'une fonction d'une variable, on parle également d'**intégrale double** par opposition à une **intégrale simple**. Nous verrons au paragraphe 2.3.1 comment réduire le calcul d'une intégrale double à l'évaluation de deux intégrales simples.

En invoquant l'additivité on étend sans difficulté la définition de l'intégrale de f aux domaines qui sont des réunions finies de rectangles essentiellement disjoints, comme celui de la figure suivante.



Cependant la géométrie du plan est très riche et la plupart des domaines apparaissant dans les applications, par exemple un disque, ne sont pas de ce type. Pour appliquer la "philosophie riemannienne" il faut non seulement, comme précédemment, approximer la fonction f par une suite f_n de fonctions constantes sur des micro-rectangles, mais aussi approximer le domaine Ω sur lequel on intègre par une suite Ω_n de réunions de tels micro-rectangles (voir la figure ci-dessous). On se heurte alors à une nouvelle difficulté. Pour assurer que la suite des sommes de Riemann correspondant à des suites d'approximations f_n, Ω_n converge vers une limite indépendante du choix de ces suites il est nécessaire de faire des hypothèses sur la "régularité" du bord du domaine Ω .



Nous n'entrerons pas dans une discussion plus détaillée de ces finesses techniques et nous contenterons d'une approche heuristique dans la section 2.3.2. Nous aborderons aussi l'approche consistant à transformer un domaine Ω en rectangle par changement de coordonnées dans la section 2.3.3.

2.2 Intégrales itérées

Comme nous l'avons déjà évoqué dans la section précédente, il est possible de réduire le calcul d'intégrales multiples à l'évaluation d'intégrales simples. Nous élaborons dans cette section la notion d'intégrales itérées qui nous servira à cet usage.

Théorème 2.1 Soit f une fonction continue sur le rectangle $R = [a, b] \times [c, d]$ et α, β deux fonctions continues sur $[a, b]$ telles que $\alpha(x) \in [c, d]$ et $\beta(x) \in [c, d]$ pour tout $x \in [a, b]$. La fonction

$$g(x) = \int_{\alpha(x)}^{\beta(x)} f(x, y) dy,$$

est continue sur $[a, b]$. De plus, si f, α et β sont de classe C^1 , g est de classe C^1 et

$$g'(x) = \int_{\alpha(x)}^{\beta(x)} \frac{\partial f}{\partial x}(x, y) dy + f(x, \beta(x))\beta'(x) - f(x, \alpha(x))\alpha'(x). \quad (2.3)$$

En conséquence, il est possible de définir l'intégrale itérée d'une fonction continue de deux variables

$$\int_a^b \left[\int_{\alpha(x_1)}^{\beta(x_1)} f(x_1, x_2) dx_2 \right] dx_1.$$

Exemple 2.1 Pour calculer l'intégrale itérée

$$I = \int_0^1 \left[\int_{-x_1}^{x_1} \frac{1}{(1+x_1+x_2)^2} dx_2 \right] dx_1,$$

on évalue tout d'abord

$$g(x_1) = \int_{-x_1}^{x_1} \frac{1}{(1+x_1+x_2)^2} dx_2 = - \frac{1}{1+x_1+x_2} \Big|_{x_2=-x_1}^{x_2=x_1} = 1 - \frac{1}{1+2x_1}.$$

On a donc

$$I = \int_0^1 \left(1 - \frac{1}{1+2x_1} \right) dx_1 = x_1 - \frac{1}{2} \log(1+2x_1) \Big|_{x_1=0}^{x_1=1} = 1 - \frac{1}{2} \log 3.$$

◁

Il est clair qu'on peut définir de même l'intégrale itérée d'une fonction de trois variables

$$\int_a^b \left[\int_{\alpha(x_1)}^{\beta(x_1)} \left[\int_{\gamma(x_1, x_2)}^{\delta(x_1, x_2)} f(x_1, x_2, x_3) dx_3 \right] dx_2 \right] dx_1.$$

Démonstration du théorème 2.1. Soit $x \in [a, b]$ fixé et $I(x)$ l'intervalle fermé d'extrémités $\alpha(x)$ et $\beta(x)$. En utilisant la relation de Chasles on peut écrire

$$g(x+h) - g(x) = \int_{\alpha(x)}^{\beta(x)} (f(x+h, y) - f(x, y)) dy - \int_{\alpha(x)}^{\alpha(x+h)} f(x+h, y) dy + \int_{\beta(x)}^{\beta(x+h)} f(x+h, y) dy.$$

Nous considérons tout d'abord le cas où f, α et β sont des fonctions continues en invoquant le théorème de la moyenne sur chaque terme du membre de droite de l'identité précédente.

Le premier terme donne lieu à

$$\int_{\alpha(x)}^{\beta(x)} (f(x+h, y) - f(x, y)) dy = (\beta(x) - \alpha(x))(f(x+h, \theta) - f(x, \theta)),$$

où $\theta \in I(x)$. Par le théorème 1.18, f est uniformément continue sur R . Par définition, pour tout $\varepsilon > 0$ il existe donc $\delta > 0$ tel que $|h| < \delta$ implique $|f(x+h, \theta) - f(x, \theta)| < \varepsilon$. Il en résulte que

$$\lim_{h \rightarrow 0} \int_{\alpha(x)}^{\beta(x)} (f(x+h, y) - f(x, y)) dy = 0.$$

On peut écrire le second terme comme

$$\int_{\alpha(x)}^{\alpha(x+h)} f(x+h, y) dy = (\alpha(x+h) - \alpha(x))f(x+h, \alpha(x) + \theta(\alpha(x+h) - \alpha(x))),$$

avec $\theta \in [0, 1]$. Puisque α est continue et f bornée sur R on obtient

$$\lim_{h \rightarrow 0} \int_{\alpha(x)}^{\alpha(x+h)} f(x+h, y) dy = 0.$$

Le troisième terme se traite de façon similaire et nous pouvons donc conclure que

$$\lim_{h \rightarrow 0} g(x+h) = g(x),$$

c'est-à-dire que g est continue.

Supposons maintenant que f , α et β sont de classe C^1 . Le premier terme s'estime à l'aide du théorème des accroissements finis

$$\frac{1}{h} \int_{\alpha(x)}^{\beta(x)} (f(x+h, y) - f(x, y)) dy = \int_{\alpha(x)}^{\beta(x)} \frac{\partial f}{\partial x}(x + \theta h, y) dy,$$

pour un $\theta \in [0, 1]$. La continuité uniforme de la dérivée partielle sur R permet de conclure, comme précédemment, que

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_{\alpha(x)}^{\beta(x)} (f(x+h, y) - f(x, y)) dy = \int_{\alpha(x)}^{\beta(x)} \frac{\partial f}{\partial x}(x, y) dy.$$

Pour le second terme on a

$$\frac{1}{h} \int_{\alpha(x)}^{\alpha(x+h)} f(x+h, y) dy = \frac{\alpha(x+h) - \alpha(x)}{h} f(x+h, \alpha(x) + h(\alpha(x+h) - \alpha(x))),$$

et la différentiabilité de α et la continuité de f nous permettent de conclure que

$$\lim_{h \rightarrow 0} \frac{1}{h} \int_{\alpha(x)}^{\alpha(x+h)} f(x+h, y) dy = \alpha'(x) f(x, \alpha(x)).$$

Le troisième terme se traite de façon analogue, ce qui achève la démonstration de la formule (2.3). \square

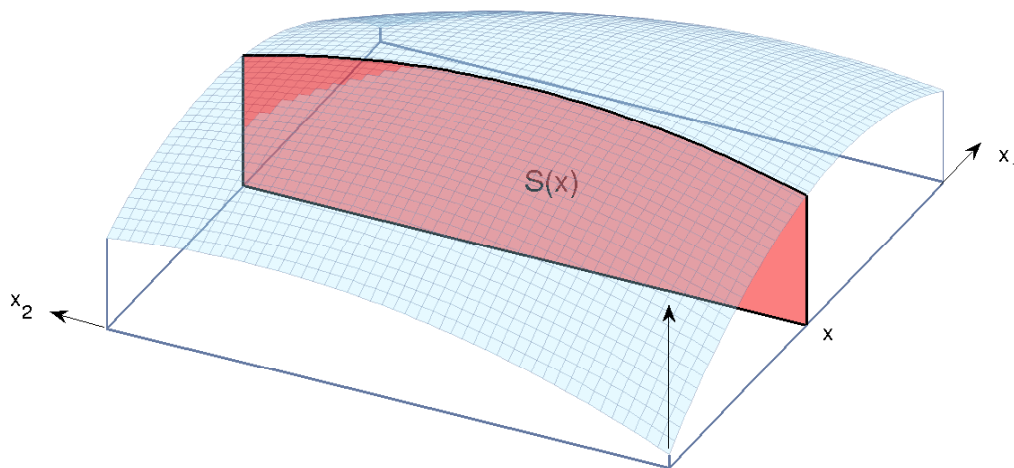
2.3 Intégrales doubles

2.3.1 Intégrales sur un rectangle

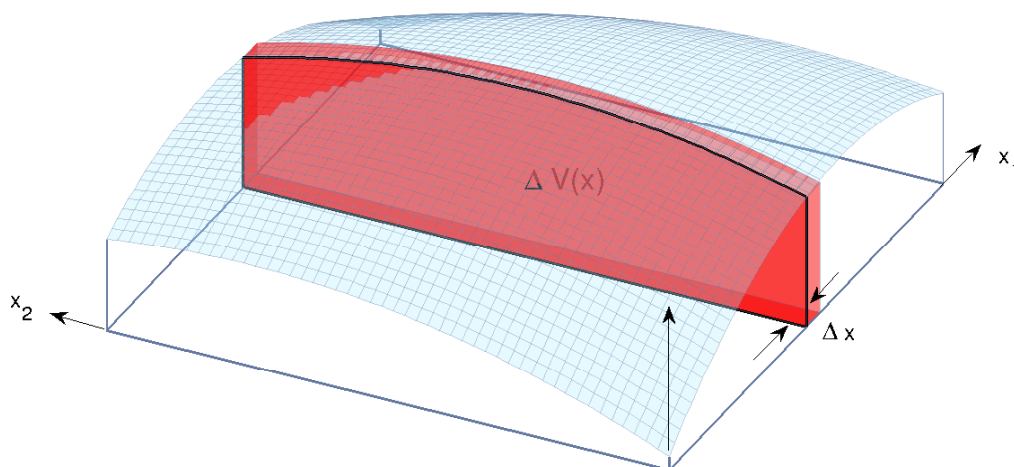
Etant donnée une fonction continue f sur le rectangle $R = [a, b] \times [c, d]$, nous cherchons à calculer l'intégrale de f sur R , c'est-à-dire le volume (compté algébriquement) de la région V limitée par R , le plan horizontal $x_3 = 0$ et le graphe de f . Comme

$$S(x) = \int_c^d f(x, x_2) dx_2$$

est l'aire (comptée algébriquement) de l'intersection de cette région avec plan $x_1 = x$.



Le volume de la tranche de \$V\$ comprise entre les plans \$x_1 = x\$ et \$x_1 = x + \Delta x\$ est approximativement égal à \$\Delta V(x) = S(x)\Delta x\$.



A une partition \$\mathcal{P} = \{a = x_0 < x_1 < \dots < x_n < x_{n+1} = b\}\$ de l'intervalle \$[a, b]\$ on peut donc associer une approximation du volume cherché en additionnant les volumes de chaque tranche

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

On reconnaît ici la somme de Riemann \$\Sigma_{\mathcal{P}}(S)\$ de la fonction \$S(x)\$ associée à la partition \$\mathcal{P}\$. Lorsque la finesse \$\delta(\mathcal{P})\$ de cette partition tend vers 0, cette somme converge vers l'intégrale de Riemann

$$\int_a^b S(x) dx.$$

Le volume recherché est donc donné par l'intégrale itérée

$$V = \int_a^b \left[\int_c^d f(x_1, x_2) dx_2 \right] dx_1.$$

Bien entendu, au lieu de découper ce volume en tranches parallèles au plan \$x_2, x_3\$ nous pouvons tout aussi bien le trancher parallèlement au plan \$x_1, x_3\$. Dans ce cas nous obtenons l'expression

$$V = \int_c^d \left[\int_a^b f(x_1, x_2) dx_1 \right] dx_2.$$

On en déduit l'identité de Fubini

$$\int_c^d \left[\int_a^b f(x_1, x_2) dx_1 \right] dx_2 = \int_a^b \left[\int_c^d f(x_1, x_2) dx_2 \right] dx_1. \quad (2.4)$$

Bien que l'argument qui nous a mené à cette conclusion ne soit pas complètement rigoureux, nous nous en satisferons et formulons ce résultat de la manière suivante.

Théorème 2.2 (Théorème de Fubini) *Si la fonction f est continue sur le rectangle $[a, b] \times [c, d]$ alors son intégrale est donnée par*

$$\iint_{\mathbb{R}^2} f(x_1, x_2) dx_1 dx_2 = \int_c^d \left[\int_a^b f(x_1, x_2) dx_1 \right] dx_2 = \int_a^b \left[\int_c^d f(x_1, x_2) dx_2 \right] dx_1.$$

Exemple 2.2 Les composantes non triviales du tenseur d'inertie (relativement au centre de gravité) d'une plaque rectangulaire homogène de masse m et de côtés a et b sont données par les intégrales

$$\begin{aligned} \Theta_{11} &= \frac{m}{ab} \iint_{\mathbb{R}^2} x_2^2 dx_1 dx_2, \\ \Theta_{22} &= \frac{m}{ab} \iint_{\mathbb{R}^2} x_1^2 dx_1 dx_2, \\ \Theta_{12} &= \frac{m}{ab} \iint_{\mathbb{R}^2} x_1 x_2 dx_1 dx_2, \end{aligned}$$

où $\mathbb{R} = [-a/2, a/2] \times [-b/2, b/2]$. On a donc

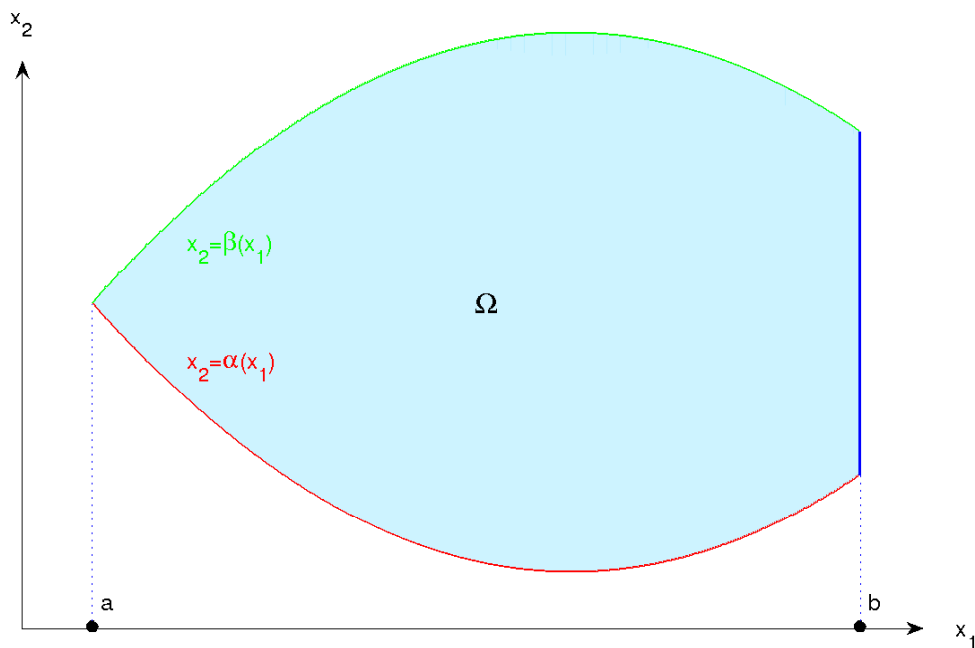
$$\begin{aligned} \Theta_{11} &= \frac{m}{ab} \int_{-a/2}^{a/2} \left[\int_{-b/2}^{b/2} x_2^2 dx_2 \right] dx_1 = \frac{m}{ab} \int_{-a/2}^{a/2} \left[\frac{x_2^3}{3} \right]_{x_2=-b/2}^{x_2=b/2} dx_1 = \frac{m}{ab} \int_{-a/2}^{a/2} \left[\frac{b^3}{12} \right] dx_1 = \frac{mb^2}{12}, \\ \Theta_{22} &= \frac{m}{ab} \int_{-b/2}^{b/2} \left[\int_{-a/2}^{a/2} x_1^2 dx_1 \right] dx_2 = \frac{m}{ab} \int_{-b/2}^{b/2} \left[\frac{x_1^3}{3} \right]_{x_1=-a/2}^{x_1=a/2} dx_2 = \frac{m}{ab} \int_{-b/2}^{b/2} \left[\frac{a^3}{12} \right] dx_2 = \frac{ma^2}{12}, \\ \Theta_{12} &= \frac{m}{ab} \int_{-a/2}^{a/2} \left[\int_{-b/2}^{b/2} x_1 x_2 dx_2 \right] dx_1 = \frac{m}{ab} \int_{-a/2}^{a/2} \left[\frac{x_1 x_2^2}{2} \right]_{x_2=-b/2}^{x_2=b/2} dx_1 = \frac{m}{ab} \int_{-a/2}^{a/2} 0 dx_1 = 0. \end{aligned}$$

◁

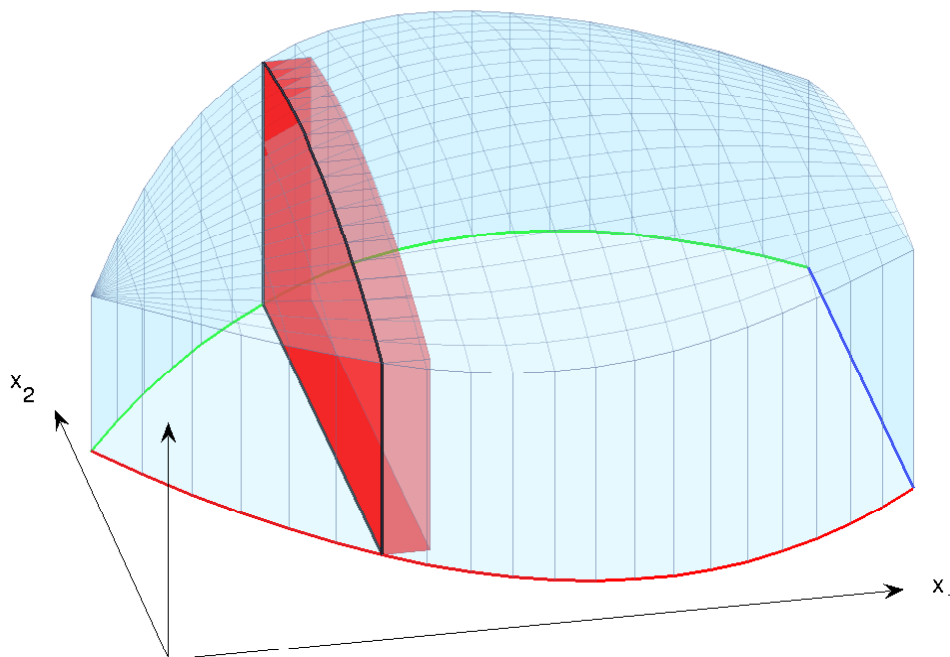
2.3.2 Intégrales sur un domaine simple

Nous dirons que $\Omega \subset \mathbb{R}^2$ est un **domaine simple** s'il est limité par les graphes des deux fonctions continues, plus précisément s'il existe un intervalle $[a, b]$ ainsi que deux fonctions α et β continues sur cet intervalle et telles que

$$\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \mid x_1 \in [a, b], \alpha(x_1) \leq x_2 \leq \beta(x_1)\},$$



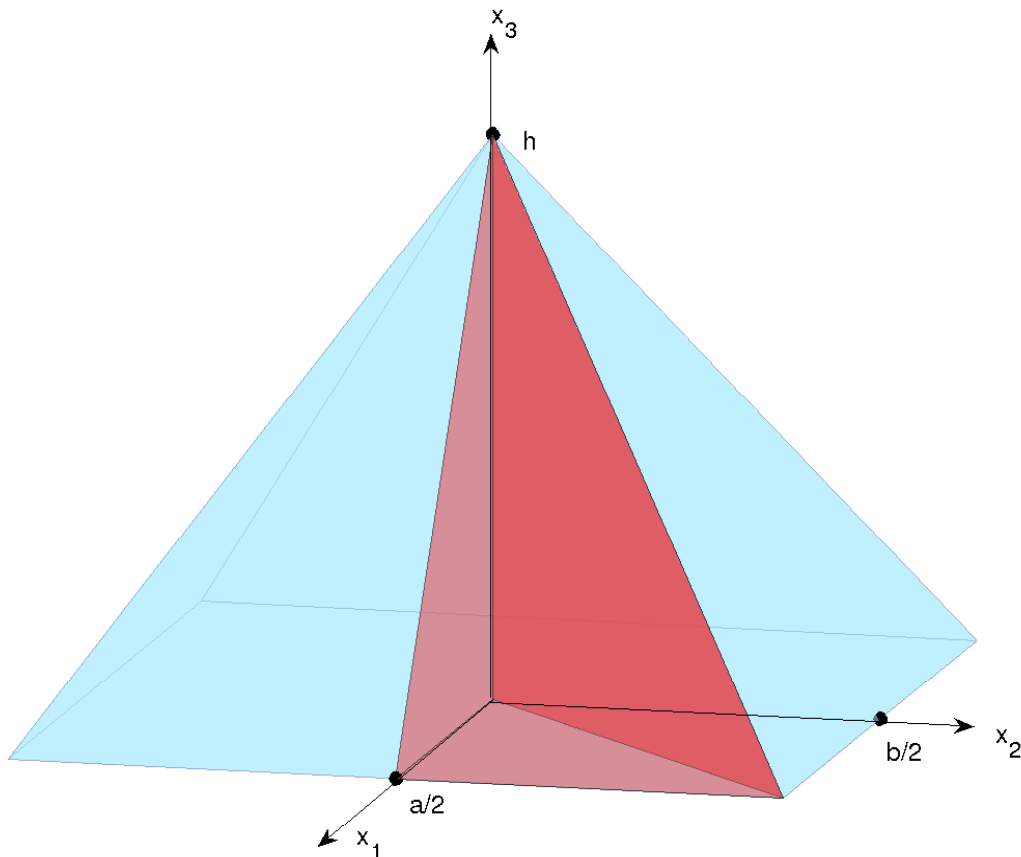
L'intégrale d'une fonction continue f sur un domaine simple Ω se calcule selon le même principe que l'intégrale sur un rectangle, la seule modification consistant à limiter latéralement les tranches par les fonctions α et β , comme illustré dans le graphique ci-dessous.



En suivant le même argument que dans la section précédente, on obtient la formule suivante pour l'intégrale de f sur le domaine Ω ,

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2 = \int_a^b \left[\int_{\alpha(x_1)}^{\beta(x_1)} f(x_1, x_2) dx_2 \right] dx_1.$$

Exemple 2.3 Calculons le volume d'une pyramide de base rectangulaire de côtés a et b et de hauteur h .



Par symétrie, ce volume vaut 8 fois le volume du tétraèdre rouge dans la figure ci-dessus. La face oblique de ce tétraèdre, de sommets $(a/2, 0, 0)$, $(a/2, b/2, 0)$ et $(0, 0, h)$, se situe sur le plan d'équation $x_3 = f(x_1, x_2) = (a - 2x_1)h/a$. Si T désigne la base du tétraèdre, le volume de ce dernier est donné par

$$\int_T f(x_1, x_2) dx_1 dx_2.$$

L'hypoténuse du triangle T se trouvant sur la droite d'équation $x_2 = bx_1/a$, on a

$$\int_T f(x_1, x_2) dx_1 dx_2 = \int_0^{a/2} \left[\int_0^{bx_1/a} \frac{h}{a} (a - 2x_1) dx_2 \right] dx_1 = \int_0^{a/2} \frac{hb}{a^2} (ax_1 - 2x_1^2) dx_1 = \frac{1}{8} \frac{abh}{3},$$

ce qui montre que le volume de la pyramide est $V = abh/3$. ◁

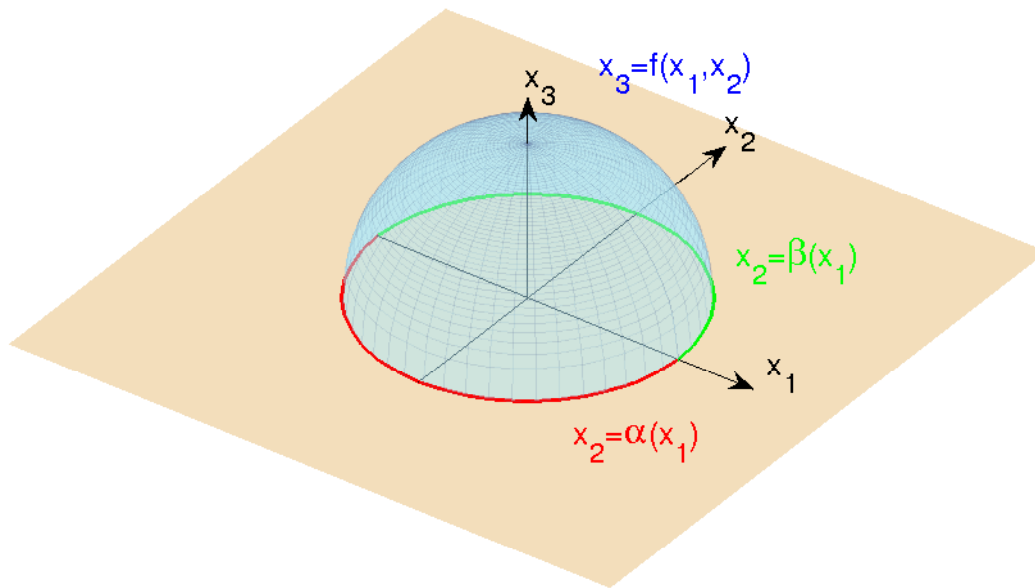
Exemple 2.4 Le volume d'une boule de rayon R (sans restreindre la généralité nous la supposons centrée à l'origine) est le double du volume contenu entre le plan $x_3 = 0$ et l'hémisphère nord de la sphère de rayon R centrée en $(0, 0, 0)$. En résolvant l'équation de cette sphère

$$x_1^2 + x_2^2 + x_3^2 = R^2,$$

pour x_3 , nous pouvons décrire son hémisphère nord comme le graphe de la fonction

$$f(x_1, x_2) = \sqrt{R^2 - x_1^2 - x_2^2},$$

dont le domaine est le disque D_R de rayon R , intersection de la boule avec le plan $x_3 = 0$. Ce disque est un domaine simple limité par les graphes des fonctions $\alpha(x_1) = -\sqrt{R^2 - x_1^2}$ et $\beta(x_1) = \sqrt{R^2 - x_1^2}$ sur l'intervalle $[-R, R]$.



Le volume de la boule est donc

$$V = 2 \int_{D_R} f(x_1, x_2) dx_1 dx_2 = 2 \int_{-R}^R \left[\int_{\alpha(x_1)}^{\beta(x_1)} f(x_1, x_2) dx_2 \right] dx_1.$$

La première intégrale, sur x_2 , s'écrit comme

$$\int_{\alpha(x_1)}^{\beta(x_1)} f(x_1, x_2) dx_2 = \int_{\alpha(x_1)}^{\beta(x_1)} \sqrt{a^2 - x_2^2} dx_2 = \frac{1}{2} \left(x_2 \sqrt{a^2 - x_2^2} + a^2 \arcsin \left(\frac{x_2}{a} \right) \right) \Big|_{x_2=\alpha(x_1)}^{x_2=\beta(x_1)},$$

où nous avons posé $a^2 = R^2 - x_1^2$. En notant que $\alpha(x_1)^2 = \beta(x_1)^2 = a^2$ et $\arcsin(\pm 1) = \pm \pi/2$ on obtient

$$\int_{\alpha(x_1)}^{\beta(x_1)} f(x_1, x_2) dx_2 = \frac{\pi}{2} (R^2 - x_1^2).$$

Nous pouvons maintenant évaluer la seconde intégrale, sur x_1 , et terminer ainsi le calcul

$$V = \pi \int_{-R}^R (R^2 - x_1^2) dx_1 = \pi \left(R^2 x_1 - \frac{x_1^3}{3} \right) \Big|_{x_1=-R}^{x_1=R} = \frac{4}{3} \pi R^3.$$

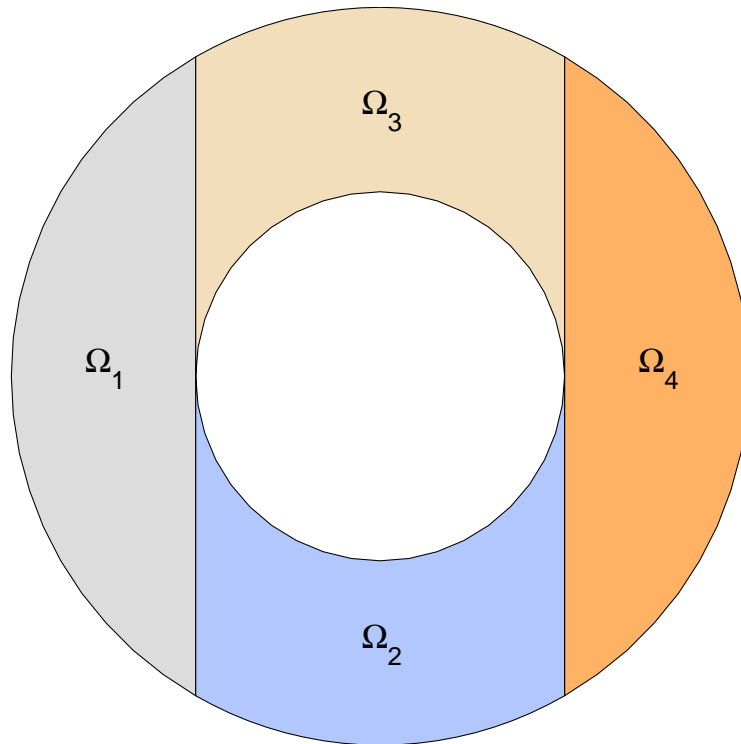
◁

Pour calculer l'intégrale d'une fonction sur des domaines plus généraux on cherche à décomposer ce domaine en sous domaines simples et on utilise l'additivité de l'intégrale. Nous ne donnerons ici qu'un exemple de cette technique.

Exemple 2.5 Pour intégrer la fonction f sur le domaine annulaire

$$\Omega = \{(x_1, x_2) \mid 1 \leq x_1^2 + x_2^2 \leq 4\},$$

on peut décomposer celui-ci en 4 domaines simples, $\Omega_1, \dots, \Omega_4$,



On obtient ainsi la décomposition suivante

$$\begin{aligned}
 \iint_{\Omega} f(x_1, x_2) \, dx_1 \, dx_2 &= \int_{-2}^{-1} \left[\int_{-\sqrt{4-x_1^2}}^{\sqrt{4-x_1^2}} f(x_1, x_2) \, dx_2 \right] dx_1 \\
 &+ \int_{-1}^1 \left[\int_{-\sqrt{4-x_1^2}}^{-\sqrt{1-x_1^2}} f(x_1, x_2) \, dx_2 \right] dx_1 \\
 &+ \int_{-1}^1 \left[\int_{\sqrt{1-x_1^2}}^{\sqrt{4-x_1^2}} f(x_1, x_2) \, dx_2 \right] dx_1 \\
 &+ \int_1^2 \left[\int_{-\sqrt{4-x_1^2}}^{\sqrt{4-x_1^2}} f(x_1, x_2) \, dx_2 \right] dx_1.
 \end{aligned}$$

◁

2.3.3 Changement de variables

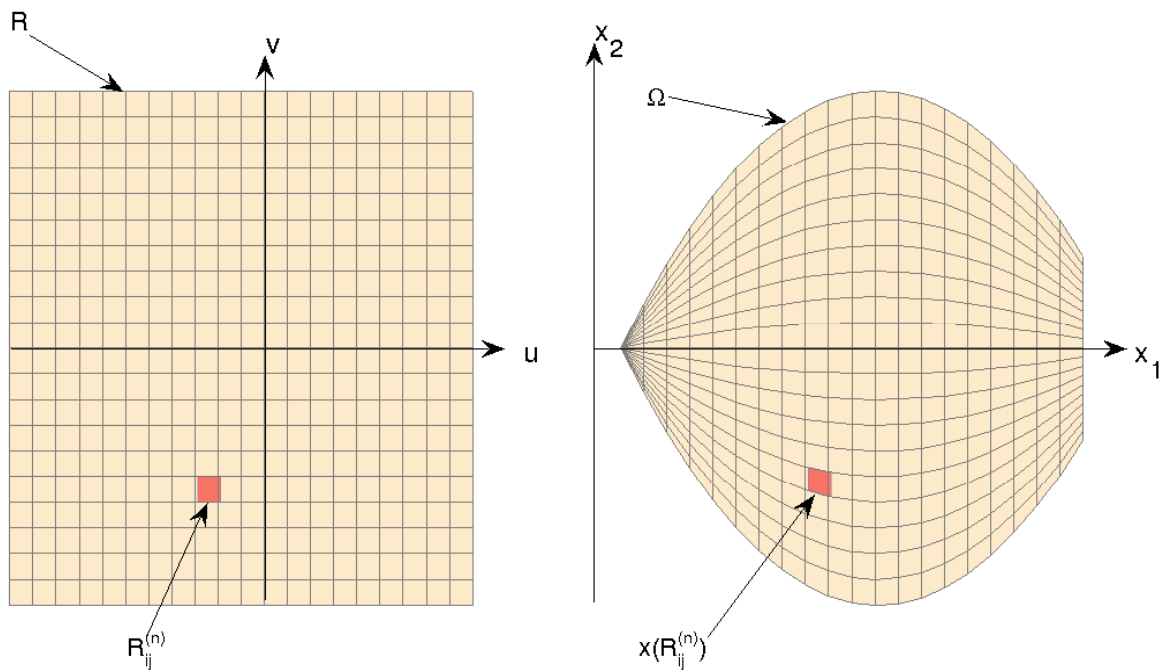
Le changement de variable d'intégration est une technique très utile pour calculer des intégrales simples. Notre but dans ce paragraphe est d'étendre cette technique aux intégrales multiples.

Soit $R = [a, b] \times [c, d] \ni (u, v) \mapsto x(u, v) \in \mathbb{R}^2$ une transformation injective que nous supposons de classe C^1 et désignons par Ω son image. (Pour être précis, nous supposons x de classe C^1 sur un rectangle ouvert $]a - \delta, b + \delta[\times]c - \delta, d - \delta[$ contenant R , voir les hypothèses du théorème 1.44 et son utilisation ci-dessous). Nous allons exprimer

$$\iint_{\Omega} f(x_1, x_2) \, dx_1 \, dx_2,$$

comme une intégrale sur le rectangle R dans les variables u et v .

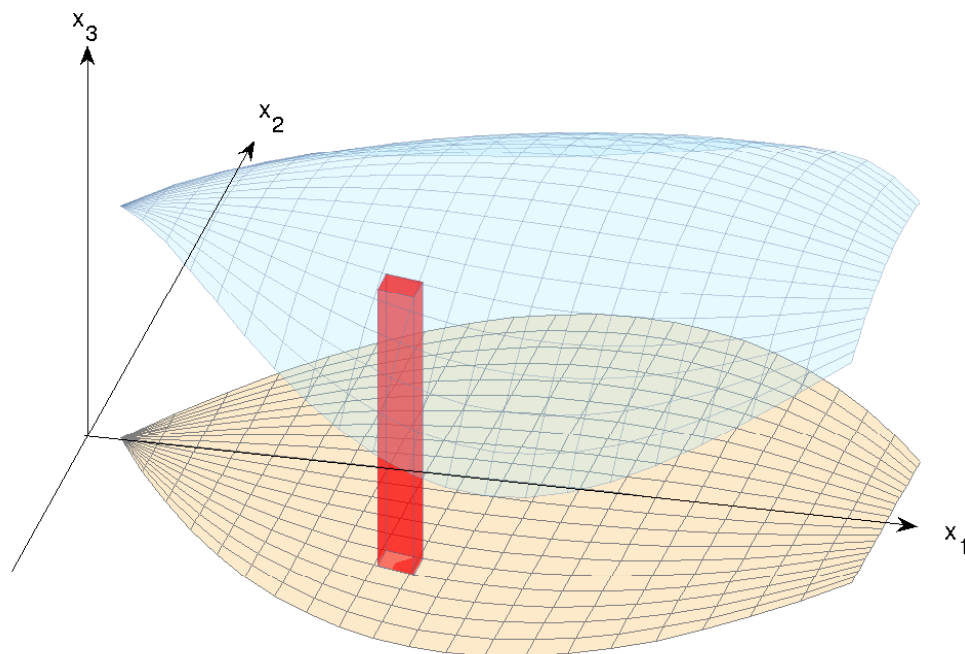
Soit $\mathcal{P}_n(R)$ la partition du rectangle R constituée des micro-rectangles $R_{ij}^{(n)}$ définis par (2.1). L'image de chacun de ces micro-rectangles est un parallélogramme curviligne $x(R_{ij}^{(n)})$ d'aire $A_{ij}^{(n)}$. L'ensemble des $x(R_{ij}^{(n)})$ forme une partition $\tilde{\mathcal{P}}_n(\Omega)$ du domaine Ω , illustrée dans la figure suivante.



A chaque partition $\tilde{\mathcal{P}}_n(R)$, nous pouvons associer une somme de Riemann de la fonction f

$$\Sigma_{\tilde{\mathcal{P}}_n(R)}(f) = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} f \left(x \left(a + (i-1) \frac{(b-a)}{2^n}, c + (j-1) \frac{(d-c)}{2^n} \right) \right) A_{ij}^{(n)}, \quad (2.5)$$

dont l'interprétation devrait maintenant vous apparaître clairement : chaque sommand de cette somme est le volume d'un cylindre dont la base $x(R_{ij}^{(n)})$ est d'aire $A_{ij}^{(n)}$ et dont la hauteur est la valeur prise par la fonction f en un des sommets du parallélogramme curviligne $x(R_{ij}^{(n)})$.



On arrive donc naturellement à conjecturer que la suite de ces sommes de Riemann converge vers

l'intégrale de f sur Ω ,

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2 = \lim_{n \rightarrow \infty} \Sigma_{\tilde{\mathcal{P}}_n(R)}(f),$$

ce qu'on peut montrer facilement.

Considérons maintenant l'aire d'un parallélogramme curviligne $\chi(R_{ij}^{(n)})$. Le théorème 1.44 nous permet d'écrire

$$A_{ij}^{(n)} = \left| J\chi_{(a+(i-1)(b-a)2^{-n}, c+(j-1)(d-c)2^{-n})} \right| \frac{(b-a)(d-c)}{2^n} + \frac{(b-a)^2 + (d-c)^2}{2^{2n}} \Delta_n,$$

où $\Delta_n \rightarrow 0$ lorsque $n \rightarrow \infty$. En introduisant cette expression dans la somme de Riemann (2.5), on obtient

$$\Sigma_{\tilde{\mathcal{P}}_n(R)}(f) = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} \tilde{f} \left(a + (i-1) \frac{(b-a)}{2^n}, c + (j-1) \frac{(d-c)}{2^n} \right) \frac{(b-a)(d-c)}{2^n} + \text{erreur}_n,$$

où nous avons posé $\tilde{f}(u, v) = f(\chi(u, v)) |J_{\chi(u, v)}|$ et où l'erreur est donnée par

$$\text{erreur}_n = \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} f \left(\chi \left(a + (i-1) \frac{(b-a)}{2^n}, c + (j-1) \frac{(d-c)}{2^n} \right) \right) \frac{(b-a)^2 + (d-c)^2}{2^{2n}} \Delta_n,$$

Nous estimons cette dernière par

$$|\text{erreur}_n| \leq \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} \max_{x \in \Omega} |f(x)| \frac{(b-a)^2 + (d-c)^2}{2^{2n}} |\Delta_n| = \max_{x \in \Omega} |f(x)| ((b-a)^2 + (d-c)^2) |\Delta_n|,$$

d'où nous obtenons

$$\lim_{n \rightarrow \infty} \text{erreur}_n = 0.$$

Nous avons donc

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2 = \lim_{n \rightarrow \infty} \sum_{i=1}^{2^n} \sum_{j=1}^{2^n} \tilde{f} \left(a + (i-1) \frac{(b-a)}{2^n}, c + (j-1) \frac{(d-c)}{2^n} \right) \frac{(b-a)(d-c)}{2^n},$$

et, en reconnaissant dans cette expression la somme de Riemann de la fonction \tilde{f} , on conclut par

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2 = \iint_R f(\chi(u, v)) |J_{\chi(u, v)}| du dv.$$

En remplaçant le rectangle R par un domaine plus général, on obtient

Théorème 2.3 Soit χ une transformation de classe C^1 sur l'ouvert O et $\Lambda \subset O$ un domaine simple (ou une réunion de tels domaines). Si f est injective sur Λ , alors pour toute fonction continue sur $\chi(\Lambda)$ on a l'identité

$$\iint_{\chi(\Lambda)} f(x_1, x_2) dx_1 dx_2 = \iint_{\Lambda} f(\chi(u, v)) |J_{\chi(u, v)}| du dv, \quad (2.6)$$

ce que l'on exprime parfois en écrivant $dx_1 dx_2 = |J_{\chi(u, v)}| du dv$.



L'hypothèse d'injectivité est essentielle dans le théorème précédent ! C'est souvent le point le plus délicat à vérifier en pratique.

Exemple 2.6 La fonction $F(r, \theta)$ de l'exemple 1.48 transforme $\Lambda = [0, R] \times [-4\pi/3, 4\pi/3]$ dans le disque D_R de rayon R centré à l'origine. On a donc, en réduisant comme dans l'exemple 2.4 l'intégrale double sur le disque D_R en une intégrale itérée,

$$\iint_{F(\Lambda)} dx_1 dx_2 = \int_{-R}^R \left[\int_{-\sqrt{R^2-x_1^2}}^{\sqrt{R^2-x_1^2}} dx_2 \right] dx_1 = \int_{-R}^R 2\sqrt{R^2-x_1^2} dx_1 = \pi R^2.$$

F est de classe C^∞ , on peut lire son jacobien sur la formule (1.56) : $J_{(r,\theta)}F = r$. On en déduit

$$\iint_{\Lambda} |J_{(r,\theta)}F| dr d\theta = \int_{-4\pi/3}^{4\pi/3} \left[\int_0^R r dr \right] d\theta = \int_{-4\pi/3}^{4\pi/3} \frac{R^2}{2} d\theta = \frac{4}{3}\pi R^2.$$

Puisque

$$\iint_{F(\Lambda)} dx_1 dx_2 \neq \iint_{\Lambda} |J_{(r,\theta)}F| dr d\theta,$$

la conclusion du théorème 2.3 est fautive. La raison en est que l'hypothèse d'injectivité n'est pas satisfaite. La différence

$$\iint_{\Lambda} |J_{(r,\theta)}F| dr d\theta - \iint_{F(\Lambda)} dx_1 dx_2 = \frac{1}{3}\pi R^2,$$

est l'aire du secteur circulaire $\{F(r, \theta) \mid r \in [0, R], \theta \in [2\pi/3, 4\pi/3]\}$ qui est "couvert" 2 fois par F (revoir la figure de l'exemple 1.48). \triangleleft

Remarque 2.4 Pour pouvoir appliquer la formule 2.6 au calcul de l'intégrale

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2,$$

on doit déterminer le domaine Λ tel que $\Omega = x(\Lambda)$. Dans beaucoup d'applications on est forcé, pour assurer l'injectivité de la transformation $(u, v) \mapsto x(u, v)$, de restreindre son domaine de telle façon à ce que son image ne recouvre pas Ω tout entier. Dans ce cas tout n'est pas perdu. En effet, si x est injective sur Λ et si $x(\Lambda) \subset \Omega$, l'additivité de l'intégrale permet d'écrire

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2 = \iint_{x(\Lambda)} f(x_1, x_2) dx_1 dx_2 + \iint_{\Omega \setminus x(\Lambda)} f(x_1, x_2) dx_1 dx_2.$$

On remarque alors que si l'aire de $\Omega \setminus x(\Lambda)$ est nulle la seconde intégrale dans le membre de droite de cette identité est nulle et on en conclut que

$$\iint_{\Omega} f(x_1, x_2) dx_1 dx_2 = \iint_{\Lambda} f(x(u, v)) |J_{x(u,v)}| du dv.$$

Exemple 2.7 Coordonnées polaire (c.f. exemple 1.44). La transformation

$$F: \mathbb{R}^2 \rightarrow \mathbb{R}^2 \\ (r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

est de classe C^∞ . Elle est surjective. En raison de la périodicité des fonctions trigonométriques, $F(r, \theta + 2\pi) = F(r, \theta)$, elle n'est pas injective. Par contre, sa restriction

$$x:]0, \infty[\times]0, 2\pi[\rightarrow \mathbb{R}^2 \\ (r, \theta) \mapsto (r \cos \theta, r \sin \theta).$$

est injective, mais n'est pas surjective. Son image $\mathbb{R}^2 \setminus \{(0, 0)\}$ ne recouvre pas totalement le plan. Cependant, ce "défaut" ne porte pas à conséquence lorsqu'il s'agit d'appliquer le théorème 2.3 car l'aire de l'ensemble manquant, $\{(0, 0)\}$, est nulle.

Le jacobien de x est $J_{(r,\theta)}x = r$. On passe donc des coordonnées cartésiennes (x_1, x_2) aux coordonnées polaires (r, θ) à l'aide de la formule

$$\iint_{x(\Lambda)} f(x_1, x_2) dx_1 dx_2 = \iint_{\Lambda} f(r \cos \theta, r \sin \theta) dr d\theta.$$

Pour illustrer l'usage de cette formule, calculons l'intégrale I de la fonction $f(x_1, x_2) = e^{a\sqrt{x_1^2+x_2^2}}$ sur le disque

$$D_R = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq R^2\}.$$

On note que D_R est l'image par la transformation x du rectangle $Q_R = [0, R] \times [0, 2\pi]$ et que $f(x(r, \theta)) = e^{ar}$. On a donc

$$I = \iint_{D_R} e^{a\sqrt{x_1^2+x_2^2}} dx_1 dx_2 = \iint_{Q_R} e^{ar} r dr d\theta.$$

Cette dernière intégrale se transforme en intégrale itérée

$$I = \int_0^{2\pi} \left[\int_0^R r e^{ar} dr \right] d\theta.$$

La première intégrale se calcule par partie

$$\int_0^R r e^{ar} dr = r \frac{e^{ar}}{a} \Big|_{r=0}^{r=R} - \int_0^R \frac{e^{ar}}{a} dr = R \frac{e^{aR}}{a} - \frac{e^{ar}}{a^2} \Big|_{r=0}^{r=R} = \frac{e^{aR}}{a^2} (e^{-aR} - (1 - aR)).$$

Comme cette fonction ne dépend pas de θ , la seconde intégrale se calcule immédiatement et livre le résultat final

$$I = 2\pi \frac{e^{aR}}{a^2} (e^{-aR} - (1 - aR)).$$

◁

Exemple 2.8 L'intégrale gaussienne

$$I(a, b) = \int_{-\infty}^{\infty} e^{-ax^2 - bx} dx,$$

où $a > 0$ et $b \in \mathbb{R}$, intervient dans beaucoup d'applications et dans des domaines variés des mathématiques et de la physique. En complétant le carré dans l'exponentielle, on la réduit à

$$I(a, b) = e^{b^2/4a} \int_{-\infty}^{\infty} e^{-a(x+b/2a)^2} dx,$$

et en changeant de variable d'intégration $y = \sqrt{a}(x + b/2a)$ on arrive à

$$I(a, b) = \frac{e^{b^2/4a}}{\sqrt{a}} \int_{-\infty}^{\infty} e^{-y^2} dy = \frac{e^{b^2/4a}}{\sqrt{a}} I(1, 0).$$

Pour évaluer l'intégrale résiduelle considérons l'intégrale double

$$I_R = \iint_{C_R} e^{-x_1^2 - x_2^2} dx_1 dx_2,$$

sur le carré $C_R = [-R, R] \times [-R, R]$. En l'écrivant comme une intégrale itérée, on obtient

$$I_R = \int_{-R}^R \left[\int_{-R}^R e^{-x_1^2 - x_2^2} dx_1 \right] dx_2 = \int_{-R}^R e^{-x_2^2} \left[\int_{-R}^R e^{-x_1^2} dx_1 \right] dx_2 = \left[\int_{-R}^R e^{-x^2} dx \right]^2,$$

et donc

$$\lim_{R \rightarrow \infty} I_R = I(1, 0)^2.$$

Soit $D_r = \{(x_1, x_2) \mid x_1^2 + x_2^2 \leq r^2\}$ le disque de rayon r et

$$J_r = \iint_{D_r} e^{-x_1^2 - x_2^2} dx_1 dx_2.$$

Comme $D_R \subset C_R \subset D_{\sqrt{2}R}$ et $e^{-x_1^2 - x_2^2} > 0$ on a

$$J_R \leq I_R \leq J_{\sqrt{2}R},$$

et par conséquent

$$\lim_{R \rightarrow \infty} I_R = \lim_{R \rightarrow \infty} J_R,$$

a condition que cette dernière limite existe. Nous calculons cette limite en passant en coordonnées polaires $\chi(r, \theta) = (r \cos \theta, r \sin \theta)$. En procédant comme dans l'exemple précédent on obtient

$$J_R = \int_0^{2\pi} \left[\int_0^R e^{-r^2} r dr \right] d\theta.$$

Après changement de variable $u = r^2$, la première intégrale devient

$$\int_0^R e^{-r^2} r dr = \frac{1}{2} \int_0^R e^{-u} du = \frac{1 - e^{-R}}{2},$$

et donc

$$J_R = \pi(1 - e^{-R}).$$

On en déduit que

$$I(1, 0) = \sqrt{\lim_{R \rightarrow \infty} J_R} = \sqrt{\pi},$$

et l'expression finale de l'intégrale gaussienne

$$I(a, b) = \sqrt{\frac{\pi}{a}} e^{b^2/4a}.$$

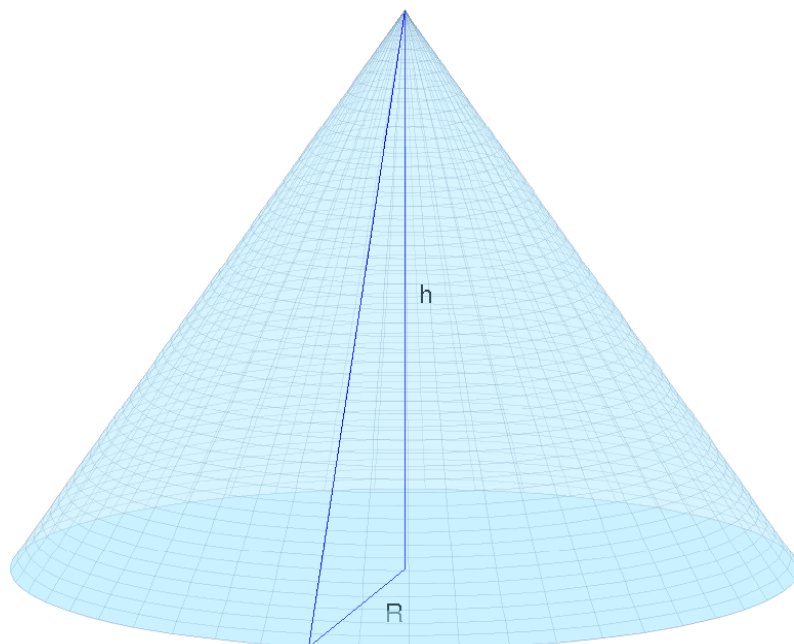
◁

Exemple 2.9 Le cône de hauteur h dont la base est un disque D_R de rayon R est le graphe de la fonction

$$f(x_1, x_2) = h \left(1 - \frac{\sqrt{x_1^2 + x_2^2}}{R} \right).$$

Le volume de ce cône est donc

$$V = \iint_{D_R} f(x_1, x_2) dx_1 dx_2 = \int_0^{2\pi} \left[\int_0^R h \left(1 - \frac{r}{R} \right) r dr \right] d\theta = 2\pi h \left(\frac{R^2}{2} - \frac{R^2}{3} \right) = \frac{1}{3} \pi R^2 h.$$



◁

2.3.4 Calcul d'aires

Le volume sous le graphe de la fonction $f: \Omega \rightarrow \mathbb{R}$ définie par $f(x_1, x_2) = 1$ est l'aire du domaine $\Omega \subset \mathbb{R}^2$. On a donc la formule générale pour l'aire d'un domaine plan

$$\text{Aire}(\Omega) = \iint_{\Omega} dx_1 dx_2.$$

Exemple 2.10 Comme premier exemple, calculons l'aire d'un disque D_R de rayon R

$$\text{Aire}(D_R) = \iint_{D_R} dx_1 dx_2.$$

Comme dans l'exemple 2.7, nous transformons le disque D_R en rectangle $Q_R = [0, R] \times [0, 2\pi]$ en passant en coordonnées polaires. On a donc

$$\text{Aire}(D_R) = \iint_{Q_R} r dr d\theta = \int_0^{2\pi} \left[\int_0^R r dr \right] d\theta = \int_0^{2\pi} \left[\frac{R^2}{2} \right] d\theta = \pi R^2.$$

◁

Exemple 2.11 Calculons l'aire du domaine limité par une ellipse de demi-axes a et b ,

$$E(a, b) = \left\{ (x_1, x_2) \mid \frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} \leq 1 \right\}.$$

Ce domaine est l'image par la transformation $x(u_1, u_2) = (au_1, bu_2)$ du disque de rayon 1. En effet,

$$\frac{x_1(u_1, u_2)^2}{a^2} + \frac{x_2(u_1, u_2)^2}{b^2} = u_1^2 + u_2^2.$$

Le jacobien de x étant $J_{(u_1, u_2)}x = ab$, on a $dx_1 dx_2 = ab du_1 du_2$. On en déduit

$$\text{Aire}(E(a, b)) = \iint_{E(a, b)} dx_1 dx_2 = \iint_{D_1} ab du_1 du_2 = ab \text{Aire}(D_1) = \pi ab.$$

Lorsque $a = b$ on retrouve l'aire du cercle de rayon a , πa^2 .

◁

2.4 Intégrales triples

Le calcul d'intégrales triples se fait de façon similaire. Nous nous contenterons de l'illustrer par quelques exemples.

2.4.1 Intégrales sur des quadrilatères

L'intégrale de la fonction $f(x_1, x_2, x_3)$ sur le quadrilatère $Q = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2]$ se réduit à l'intégrale itérée

$$\iiint_Q f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \int_{a_1}^{a_2} \left[\int_{b_1}^{b_2} \left[\int_{c_1}^{c_2} f(x_1, x_2, x_3) dx_3 \right] dx_2 \right] dx_1.$$

Le théorème de Fubini permet de permuter à volonté l'ordre d'intégration. Il y a donc 6 façons différentes de calculer une telle intégrale. Par exemple

$$\iiint_Q f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \int_{b_1}^{b_2} \left[\int_{c_1}^{c_2} \left[\int_{a_1}^{a_2} f(x_1, x_2, x_3) dx_1 \right] dx_3 \right] dx_2.$$



Alors que le résultat, c'est-à-dire l'intégrale de f sur Q , ne dépend pas de cet ordre, la difficulté que représente chacune des 3 intégrales peut, elle, en dépendre.

Exemple 2.12 Soit $Q = [-1, 1] \times [-1, 1] \times [-1, 1]$. L'intégrale

$$I = \iiint_Q x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} dx_1 dx_2 dx_3,$$

se calcule aisément si on l'écrit

$$I = \int_{-1}^1 \left[\int_{-1}^1 \left[\int_{-1}^1 x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} dx_1 \right] dx_3 \right] dx_2.$$

En effet la première intégrale s'annule en raison de la symétrie $x_1 \mapsto -x_1$,

$$\int_{-1}^1 x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} dx_1 = \int_0^1 \left(x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} - x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} \right) dx_1 = 0.$$

On en déduit donc que $I = 0$. Par contre, si on écrit

$$I = \int_{-1}^1 \left[\int_{-1}^1 \left[\int_{-1}^1 x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} dx_3 \right] dx_2 \right] dx_1,$$

la première intégrale

$$\int_{-1}^1 x_1 e^{-\sqrt{x_1^2 + x_2^2 + x_3^2}} dx_3,$$

ne se calcule pas analytiquement. ◁

Exemple 2.13 L'intégrale de la fonction $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ sur le quadrilatère $Q = [0, a] \times [0, b] \times [0, c]$ se réduit à l'intégrale itérée

$$\iiint_Q f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \int_0^a \left[\int_0^b \left[\int_0^c (x_1^2 + x_2^2 + x_3^2) dx_3 \right] dx_2 \right] dx_1.$$

La première intégrale est

$$\int_0^c (x_1^2 + x_2^2 + x_3^2) dx_3 = cx_1^2 + cx_2^2 + \frac{c^3}{3}.$$

La deuxième intégrale se calcule tout aussi aisément

$$\int_0^b \left(cx_1^2 + cx_2^2 + \frac{c^3}{3} \right) dx_2 = bcx_1^2 + \frac{b^3c + bc^3}{3}.$$

Finalement, la troisième intégrale donne

$$\int_0^a \left(bcx_1^2 + \frac{b^3c + bc^3}{3} \right) dx_1 = \frac{a^3bc + ab^3c + abc^3}{3} = \frac{1}{3}abc(a^2 + b^2 + c^2). ◁$$

2.4.2 Intégrales sur des domaines simples

Des domaines plus généraux peuvent souvent être découpés en tranches qui permettent d'écrire l'intégrale comme une intégrale itérée.

Exemple 2.14 Soit $B_R = \{x \in \mathbb{R}^3 \mid \|x\| \leq R\}$ la boule de rayon R centrée à l'origine. Pour tout $h \in [-R, R]$, soit $D_R(h)$ l'intersection de B_R avec le plan $x_3 = h$. Elle est déterminée par le système

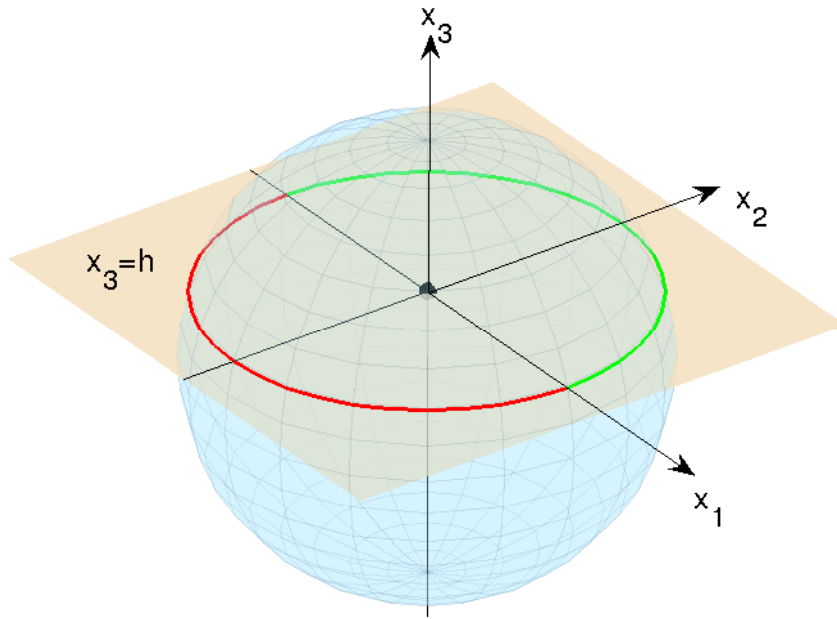
$$x_1^2 + x_2^2 + x_3^2 \leq R^2, \quad x_3 = h,$$

qu'on peut résoudre pour obtenir

$$D_R(h) = \{(x_1, x_2, x_3) \mid x_1^2 + x_2^2 \leq R^2 - h^2 \text{ et } x_3 = h\}.$$

Ce découpage de la boule B_R permet d'écrire

$$\iiint_{B_R} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \int_{-R}^R \left[\iint_{D_R(x_3)} f(x_1, x_2, x_3) dx_1 dx_2 \right] dx_3.$$



L'intégrale double sur $D_R(x_3)$, un disque de rayon $a(x_3) = \sqrt{R^2 - x_3^2}$, peut s'effectuer comme dans l'exemple 2.4,

$$\iint_{D_R(x_3)} f(x_1, x_2, x_3) dx_1 dx_2 = \int_{-a(x_3)}^{a(x_3)} \left[\int_{-\sqrt{a(x_3)^2 - x_1^2}}^{\sqrt{a(x_3)^2 - x_1^2}} f(x_1, x_2, x_3) dx_2 \right] dx_1.$$

Nous avons ainsi obtenu la réduction à une intégrale itérée

$$\iiint_{B_R} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \int_{-R}^R \left[\int_{-\sqrt{R^2 - x_3^2}}^{\sqrt{R^2 - x_3^2}} \left[\int_{-\sqrt{R^2 - x_3^2 - x_1^2}}^{\sqrt{R^2 - x_3^2 - x_1^2}} f(x_1, x_2, x_3) dx_2 \right] dx_1 \right] dx_3.$$

Il est aussi possible d'effectuer l'intégrale sur le disque $D_R(x_3)$ en coordonnées polaires, comme dans l'exemple 2.7. Dans ce cas on obtient

$$\iint_{D_R(x_3)} f(x_1, x_2, x_3) dx_1 dx_2 = \int_0^{2\pi} \left[\int_0^{\sqrt{R^2 - x_3^2}} f(r \cos \theta, r \sin \theta, x_3) r dr \right] d\theta,$$

et par conséquence

$$\iiint_{B_R} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \int_{-R}^R \left[\int_0^{2\pi} \left[\int_0^{\sqrt{R^2 - x_3^2}} f(r \cos \theta, r \sin \theta, x_3) r dr \right] d\theta \right] dx_3. \quad (2.7)$$

Pour illustrer cette réduction, calculons le potentiel de gravitation d'une boule homogène de rayon R et de densité ρ à une distance $a > R$ de son centre. Il est donné par l'intégrale

$$\Phi(a) = -G\rho \iiint_{B_R} \frac{1}{\sqrt{x_1^2 + x_2^2 + (x_3 - a)^2}} dx_1 dx_2 dx_3.$$

La symétrie de l'intégrand par rapport aux rotations autour de l'axe x_3 invite à l'usage de coordonnées polaires. Nous appliquons donc la formule (2.7) pour obtenir

$$\Phi(a) = -G\rho \int_{-R}^R \left[\int_0^{2\pi} \left[\int_0^{\sqrt{R^2 - x_3^2}} \frac{r}{\sqrt{r^2 + (x_3 - a)^2}} dr \right] d\theta \right] dx_3.$$

La première intégrale donne, compte tenu du fait que $x_3 \leq R < a$,

$$\int_0^{\sqrt{R^2 - x_3^2}} \frac{r}{\sqrt{r^2 + (x_3 - a)^2}} dr = \sqrt{r^2 + (x_3 - a)^2} \Big|_{r=0}^{r=\sqrt{R^2 - x_3^2}} = \sqrt{R^2 + a^2 - 2ax_3} - (a - x_3).$$

La seconde intégrale est très simple puisque son intégrand est indépendant de la variable d'intégration θ (c'est ici que la symétrie invoquée plus haut entre en jeu !)

$$\int_0^{2\pi} \left[\int_0^{\sqrt{R^2 - x_3^2}} \frac{r}{\sqrt{r^2 + (x_3 - a)^2}} dr \right] d\theta = 2\pi \left(\sqrt{R^2 + a^2 - 2ax_3} - (a - x_3) \right).$$

La dernière intégrale donne le résultat final

$$\begin{aligned} \Phi(a) &= -G\rho \int_{-R}^R 2\pi \left(\sqrt{R^2 + a^2 - 2ax_3} - (a - x_3) \right) dx_3 \\ &= -2\pi G\rho \left(-\frac{(R^2 + a^2 - 2ax_3)^{3/2}}{3a} - ax_3 + \frac{x_3^2}{2} \right) \Big|_{x_3=-R}^{x_3=R} \\ &= -G \left(\frac{4}{3}\pi R^3 \right) \rho \frac{1}{a}. \end{aligned}$$

En dénotant par

$$M = \left(\frac{4}{3}\pi R^3 \right) \rho,$$

la masse totale de la boule, on peut formuler le résultat comme

$$\Phi(a) = -\frac{GM}{a},$$

ce qui démontre une propriété bien connue du potentiel newtonien : le potentiel d'une boule homogène est, à l'extérieur de celle-ci, égal au potentiel d'un point matériel de même masse placé en son centre. <

Exemple 2.15 La fonction de structure d'un solide Ω

$$\mathbb{R}^3 \ni k \mapsto I_\Omega(k) = \iiint_\Omega e^{i(k_1 x_1 + k_2 x_2 + k_3 x_3)} dx_1 dx_2 dx_3,$$

joue un rôle important dans l'étude de la diffraction des ondes sur ce solide. Calculons la fonction de structure du tétraèdre limité par les plans d'équations

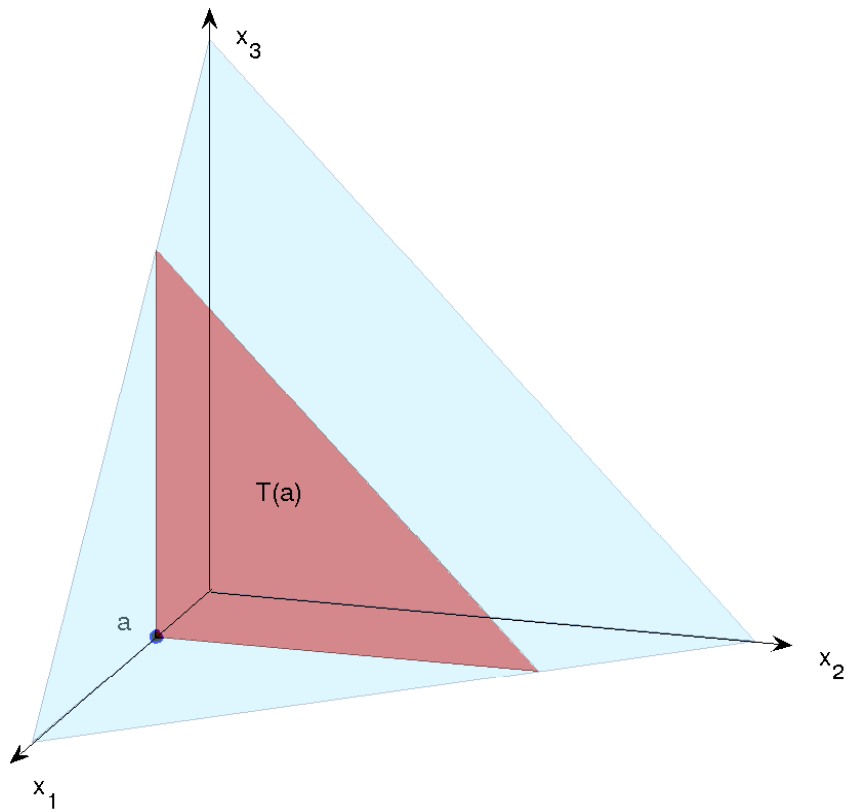
$$x_1 = 0, \quad x_2 = 0, \quad x_3 = 0, \quad x_1 + x_2 + x_3 = 1,$$

c'est-à-dire du solide

$$\Omega = \{(x_1, x_2, x_3) \mid x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_1 + x_2 + x_3 \leq 1\}.$$

Pour tout $a \in [0, 1]$, l'intersection de Ω avec le plan d'équation $x_1 = a$ est l'intérieur d'un triangle

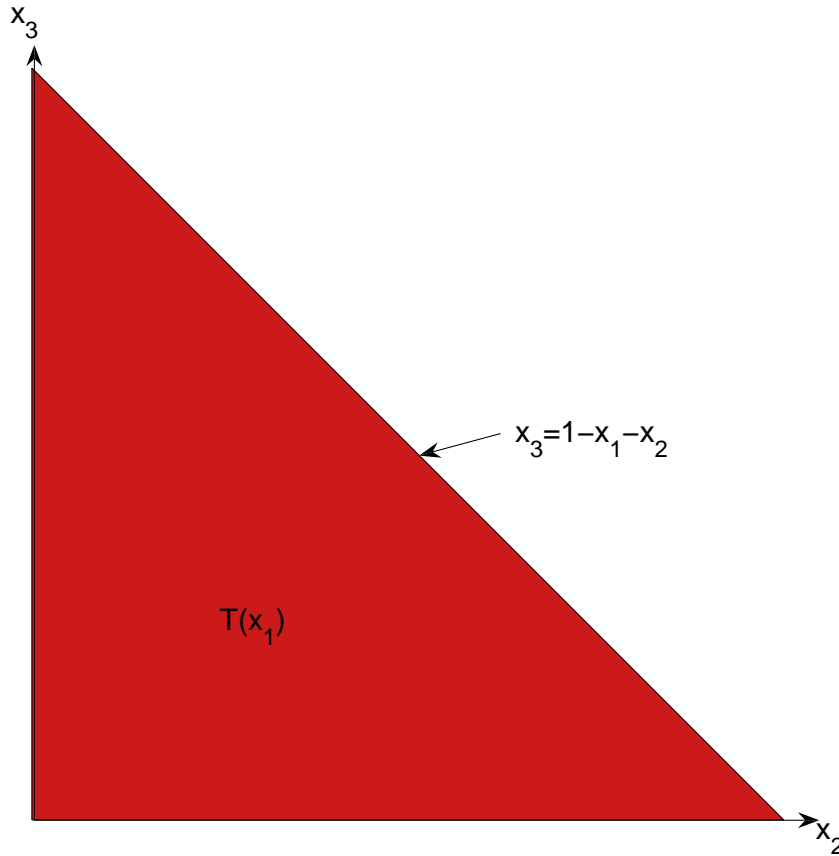
$$T(a) = \{(x_1, x_2, x_3) \mid x_1 = a, x_2 \geq 0, x_3 \geq 0, x_2 + x_3 \leq 1 - a\}.$$



On a donc

$$\begin{aligned} I_{\Omega}(\mathbf{k}) &= \int_0^1 \left[\iint_{T(x_1)} e^{i(k_1 x_1 + k_2 x_2 + k_3 x_3)} dx_2 dx_3 \right] dx_1 \\ &= \int_0^1 e^{ik_1 x_1} \left[\iint_{T(x_1)} e^{i(k_2 x_2 + k_3 x_3)} dx_2 dx_3 \right] dx_1. \end{aligned}$$

L'intégrale sur le triangle $T(x_1)$



se décompose simplement

$$\begin{aligned} \iint_{T(x_1)} e^{i(k_2 x_2 + k_3 x_3)} dx_2 dx_3 &= \int_0^{1-x_1} \left[\int_0^{1-x_1-x_2} e^{i(k_2 x_2 + k_3 x_3)} dx_3 \right] dx_2 \\ &= \int_0^{1-x_1} e^{ik_2 x_2} \left[\int_0^{1-x_1-x_2} e^{ik_3 x_3} dx_3 \right] dx_2. \end{aligned}$$

On a donc l'intégrale itérée

$$I_{\Omega}(k) = \int_0^1 e^{ik_1 x_1} \left[\int_0^{1-x_1} e^{ik_2 x_2} \left[\int_0^{1-x_1-x_2} e^{ik_3 x_3} dx_3 \right] dx_2 \right] dx_1.$$

La première intégrale est

$$\int_0^{1-x_1-x_2} e^{ik_3 x_3} dx_3 = \frac{e^{ik_3(1-x_1-x_2)} - 1}{ik_3}.$$

La seconde devient

$$\int_0^{1-x_1} \frac{e^{i(k_2-k_3)x_2} e^{ik_3(1-x_1)} - e^{ik_2 x_2}}{ik_3} dx_2 = \frac{e^{ik_3(1-x_1)} - e^{ik_2(1-x_1)}}{k_3(k_2-k_3)} + \frac{e^{ik_2(1-x_1)} - 1}{k_2 k_3}.$$

Finalement, la dernière intégrale nous donne

$$\begin{aligned} I_{\Omega}(k) &= \int_0^1 \left(\frac{e^{i(k_1-k_3)x_1} e^{ik_3} - e^{i(k_1-k_2)x_1} e^{ik_2}}{k_3(k_2-k_3)} + \frac{e^{i(k_1-k_2)x_1} e^{ik_2} - e^{ik_1 x_1}}{k_2 k_3} \right) dx_1 \\ &= \frac{(e^{i(k_1-k_3)} - 1)e^{ik_3}}{i(k_1-k_3)(k_2-k_3)k_3} - \frac{(e^{i(k_1-k_2)} - 1)e^{ik_2}}{i(k_1-k_2)(k_2-k_3)k_3} + \frac{(e^{i(k_1-k_2)} - 1)e^{ik_2}}{i(k_1-k_2)k_2 k_3} - \frac{e^{ik_1} - 1}{k_1 k_2 k_3}. \end{aligned}$$

Après un peu d'algèbre élémentaire, on peut exprimer ce résultat sous la forme plus symétrique

$$I_{\Omega}(k) = G(k_1, k_2, k_3) + G(k_2, k_3, k_1) + G(k_3, k_1, k_2),$$

où

$$G(u, v, w) = e^{iu/2} \frac{\sin(u/2)}{(u/2)} \frac{1}{(u-v)(w-u)}.$$

2.4.3 Changement de variables

Pour effectuer un changement de variables dans une intégrale triple on suit une procédure complètement analogue à celle décrite dans le théorème 2.3 pour les intégrales doubles. Si $\chi : \Omega \subset \mathbb{R}^3 \rightarrow \mathbb{R}^3$ est une transformation de classe C^1 injective sur $\Lambda \subset \Omega$

$$\iiint_{\chi(\Lambda)} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \iiint_{\Lambda} f(\chi(u_1, u_2, u_3)) |J_{(u_1, u_2, u_3)\chi}| du_1 du_2 du_3,$$

où

$$J_{(u_1, u_2, u_3)\chi} = \frac{\partial(x_1, x_2, x_3)}{\partial(u_1, u_2, u_3)},$$

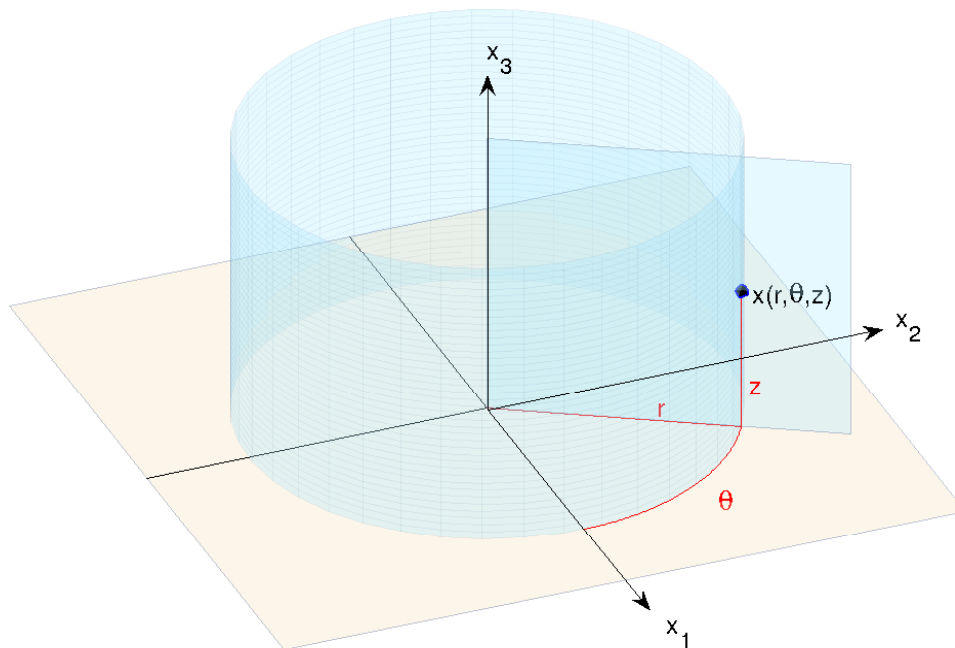
est le jacobien de la transformation χ . La remarque 2.4 s'applique également. Si Ω est tel que $\chi(\Lambda) \subset \Omega$ et si le volume de $\Omega \setminus \chi(\Lambda)$ est nul, alors

$$\iiint_{\Omega} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \iiint_{\Lambda} f(\chi(u_1, u_2, u_3)) |J_{(u_1, u_2, u_3)\chi}| du_1 du_2 du_3,$$

Exemple 2.16 Les coordonnées cylindriques (r, θ, z) d'un point $x \in \mathbb{R}^3$ sont obtenues en introduisant des coordonnées polaires dans le plan x_1, x_2 et en conservant la troisième coordonnée x_3 . On peut donc associer ces coordonnées à la transformation

$$\chi(r, \theta, z) = (r \cos \theta, r \sin \theta, z),$$

de classe C^∞ sur \mathbb{R}^3 et injective lorsqu'on la restreint à $]0, \infty[\times]0, 2\pi[\times]-\infty, \infty[$. L'image par χ de ce domaine est \mathbb{R}^3 privé de l'axe x_3 . Le volume de ce dernier étant nul, on a pas à se soucier de ce "défaut".



Son jacobien étant

$$J_{(r, \theta, z)\chi} = \det \begin{pmatrix} \cos \theta & -r \sin \theta & 0 \\ \sin \theta & r \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} = r,$$

on a la formule de changement de variables suivante

$$\iiint_{\chi(\Lambda)} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \iiint_{\Lambda} f(r \cos \theta, r \sin \theta, z) r dr d\theta dz.$$

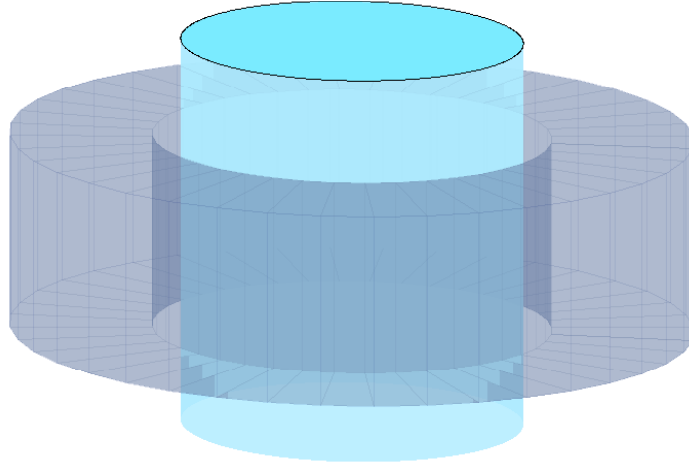
L'usage de coordonnées cylindriques est particulièrement adapté aux problèmes possédant une symétrie de rotation autour de l'axe x_3 ou/et une symétrie de translation dans la direction de cet axe.

Pour illustrer l'usage de ces coordonnées, nous calculons l'énergie du champ magnétique (1.32). L'électrodynamique nous apprend que l'énergie d'un champ magnétique $B(x)$ contenue dans une région $V \subset \mathbb{R}^3$ vide de matière est donnée par

$$E(V) = \frac{1}{8\pi} \iiint_V \|B(x)\|^2 dx_1 dx_2 dx_3.$$

Soit V le manchon cylindrique de hauteur L , de rayon intérieur $R_1 > R$ et de rayon extérieur $R_2 > R_1$,

$$V = \{(x_1, x_2, x_3) \mid R_1 \leq x_1^2 + x_2^2 \leq R_2, 0 \leq x_3 \leq L\}.$$



Son expression en coordonnées cylindriques est $V = x(\Lambda)$ où $\Lambda = [R_1, R_2] \times [0, 2\pi] \times [0, L]$. On a donc

$$E(V) = \frac{1}{8\pi} \int_0^L \left[\int_0^{2\pi} \left[\int_{R_1}^{R_2} \|B(r \cos \theta, r \sin \theta, z)\|^2 r dr \right] d\theta \right] dz.$$

Pour le champ (1.32),

$$B(r \cos \theta, r \sin \theta, z) = \frac{2I}{cr} \begin{pmatrix} -\sin \theta \\ \cos \theta \\ 0 \end{pmatrix},$$

et donc

$$E(V) = \frac{1}{8\pi} \int_0^L \left[\int_0^{2\pi} \left[\int_{R_1}^{R_2} \left(\frac{2I}{cr} \right)^2 r dr \right] d\theta \right] dz.$$

La première intégrale donne

$$\int_{R_1}^{R_2} \frac{4I^2}{c^2 r} dr = \frac{4I^2}{c^2} \log \left(\frac{R_2}{R_1} \right).$$

Les deux intégrales restantes sont donc triviales et livrent le résultat

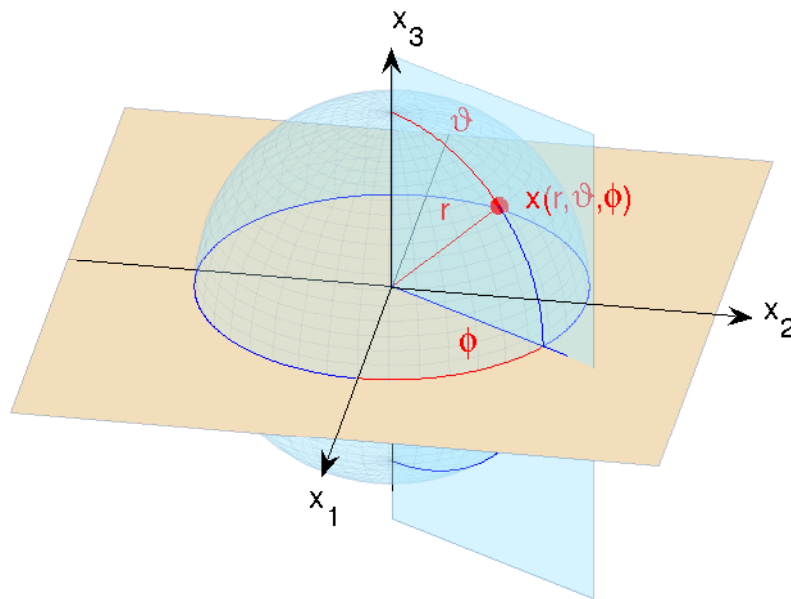
$$E(V) = \frac{I^2}{c^2} L \log \left(\frac{R_2}{R_1} \right).$$

◁

Exemple 2.17 Les coordonnées sphériques (r, ϑ, ϕ) dans \mathbb{R}^3 sont associées à la transformation

$$x(r, \vartheta, \phi) = (r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta).$$

De classe C^∞ sur \mathbb{R}^3 , cette transformation est injective sur $]0, \infty[\times]0, \pi[\times [0, 2\pi[$. L'image par x de ce domaine est \mathbb{R}^3 privé de l'axe x_3 . Tout comme dans le cas des coordonnées cylindriques, on peut donc négliger ce "défaut".



La signification géométrique des coordonnées sphériques est simple. La coordonnée r du point x est sa distance à l'origine, $r = \sqrt{x_1^2 + x_2^2 + x_3^2}$. Sa coordonnée ϑ est sa latitude mesurée depuis le "pôle nord" alors que ϕ est sa longitude mesurée depuis le méridien passant par le côté positif de l'axe x_1 .

Le jacobien de la transformation x est

$$J_{(r,\vartheta,\phi)x} = \det \begin{pmatrix} \cos \phi \sin \vartheta & r \cos \phi \cos \vartheta & -r \sin \phi \sin \vartheta \\ \sin \phi \sin \vartheta & r \sin \phi \cos \vartheta & r \cos \phi \sin \vartheta \\ \cos \vartheta & -r \sin \vartheta & 0 \end{pmatrix} = r^2 \sin \vartheta,$$

et la formule de changement de variables correspondante

$$\iiint_{x(\Lambda)} f(x_1, x_2, x_3) dx_1 dx_2 dx_3 = \iiint_{\Lambda} f(r \cos \phi \sin \vartheta, r \sin \phi \sin \vartheta, r \cos \vartheta) r^2 \sin \vartheta dr d\vartheta d\phi.$$

On utilise des coordonnées sphériques dans les problèmes possédant une symétrie de rotation complète autour de l'origine. Comme illustration, reprenons le problème de l'exemple 2.4 : le calcul du volume de la boule B_R de rayon R . Comme nous le verrons dans le paragraphe suivant, ce dernier est donné par

$$\iiint_{B_R} dx_1 dx_2 dx_3.$$

En coordonnées sphériques, la boule B_R est simplement décrite par $r \leq R$. On a donc

$$\iiint_{B_R} dx_1 dx_2 dx_3 = \int_0^R \left[\int_0^\pi \left[\int_0^{2\pi} r^2 \sin \vartheta d\phi \right] d\vartheta \right] dr.$$

La première intégrale est

$$\int_0^{2\pi} r^2 \sin \vartheta d\phi = 2\pi r^2 \sin \vartheta.$$

La seconde

$$\int_0^\pi 2\pi r^2 \sin \vartheta d\vartheta = -2\pi r^2 \cos \vartheta \Big|_{\vartheta=0}^{\vartheta=\pi} = 4\pi r^2,$$

et finalement

$$\int_0^R 4\pi r^2 dr = 4\pi \frac{r^3}{3} \Big|_{r=0}^{r=R} = \frac{4}{3} \pi R^3.$$

Exemple 2.18 Comme exemple, calculons le potentiel de gravitation d'une boule homogène de densité ρ et de rayon R .

$$\Phi(x) = -G \iiint_{B_R} \frac{\rho}{\|x - y\|} dy_1 dy_2 dy_3.$$

On remarque tout d'abord que ce potentiel est invariant par rotation. En effet, soit R une rotation de l'espace

$$\Phi(Rx) = -G \iiint_{B_R} \frac{\rho}{\|Rx - y\|} dy_1 dy_2 dy_3.$$

Comme une rotation est une isométrie, elle préserve la norme et on a

$$\|Rx - y\| = \|R(x - R^{-1}y)\| = \|x - R^{-1}y\|.$$

Dans l'intégrale

$$\Phi(Rx) = -G \iiint_{B_R} \frac{\rho}{\|x - R^{-1}y\|} dy_1 dy_2 dy_3,$$

effectuons le changement de variables $y(z) = Rz$. Le jacobien de cette transformation est $J_z y = \det R = 1$. ◁

2.4.4 Calcul de volumes

Tout comme l'intégrale double de la fonction constante $f(x) = 1$ sur un domaine du plan donne l'aire de ce domaine, l'intégrale triple de cette fonction sur un domaine $\Omega \subset \mathbb{R}^3$ donne son volume

$$\text{Volume}(\Omega) = \iiint_{\Omega} dx_1 dx_2 dx_3.$$

Exemple 2.19 Le tore T est le solide de révolution obtenu par la rotation d'un disque de rayon a autour d'un axe coplanaire à ce disque passant à distance $b > a$ de son centre. En coordonnées cylindriques, on le décrit par

$$(r - b)^2 + z^2 \leq a^2.$$

On peut donc paramétrer les coordonnées r et z par

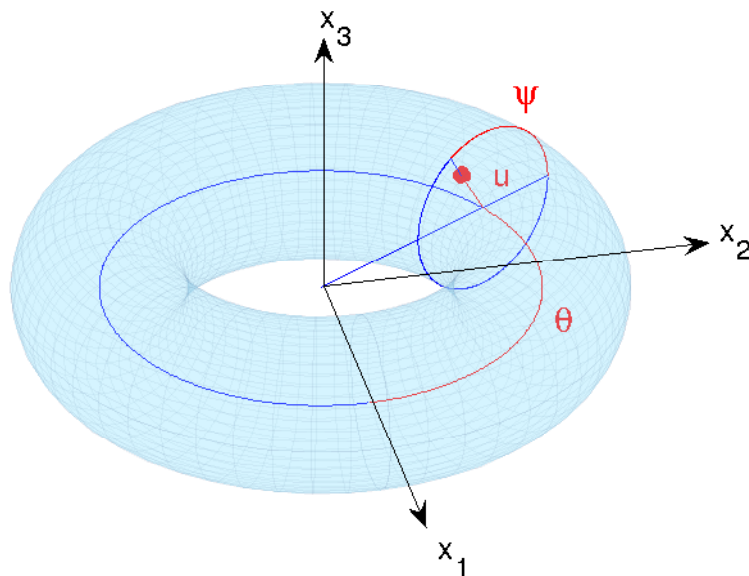
$$r - b = u \cos \psi, \quad z = u \sin \psi,$$

avec $u \in [0, a]$ et $\psi \in [0, 2\pi[$. Comme $r^2 = x_1^2 + x_2^2 = (b + u \cos \psi)^2$ on a

$$x_1 = (b + u \cos \psi) \cos \theta, \quad x_2 = (b + u \cos \psi) \sin \theta,$$

avec $\theta \in [0, 2\pi[$. On en conclut que $T = x(\Lambda)$ avec $\Lambda = [0, a] \times [0, 2\pi[\times [0, 2\pi[$ et

$$x(u, \theta, \psi) = ((b + u \cos \psi) \cos \theta, (b + u \cos \psi) \sin \theta, u \sin \psi).$$



Le jacobien de la transformation χ est

$$J_{(u,\theta,\psi)\chi} = \det \begin{pmatrix} \cos \psi \cos \theta & -(b + u \cos \psi) \sin \theta & -u \sin \psi \cos \theta \\ \cos \psi \sin \theta & (b + u \cos \psi) \cos \theta & -u \sin \psi \sin \theta \\ \sin \psi & 0 & u \cos \psi \end{pmatrix} = u(b + u \cos \psi),$$

et le volume du tore T est donné par

$$\text{Volume}(T) = \iiint_T dx_1 dx_2 dx_3 = \iiint_{\Lambda} u(b + u \cos \psi) du d\theta d\psi,$$

qu'on réduit à l'intégrale itérée

$$\text{Volume}(T) = \int_0^a \left[\int_0^{2\pi} \left[\int_0^{2\pi} u(b + u \cos \psi) d\psi \right] d\theta \right] du.$$

La première intégrale est

$$\int_0^{2\pi} u(b + u \cos \psi) d\psi = 2\pi bu.$$

La seconde intégrale donne

$$\int_0^{2\pi} 2\pi bu d\theta = (2\pi)^2 bu.$$

Finalement, la dernière intégrale nous livre le résultat

$$\text{Volume}(T) = \int_0^a (2\pi)^2 bu du = (2\pi)^2 b \frac{a^2}{2} = 2\pi^2 a^2 b.$$

◁

Chapitre 3

Equations différentielles

De nombreux problèmes mathématiques mais aussi de physique, de biologie, d'économie ... conduisent à déterminer une fonction $y(x)$ satisfaisant une équation dy type

$$F(x, y(x), y'(x), y''(x), \dots, y^{(n)}(x)) = 0, \quad (3.1)$$

où F une fonction suffisamment régulière de $n + 1$ variables. On dit que (3.1) est une **équation différentielle ordinaire (EDO) d'ordre n** . Une fonction $f :]a, b[\rightarrow \mathbb{R}$ est une solution de cette équation si elle est de classe C^n sur l'intervalle $]a, b[$ et si, en chaque point $x \in]a, b[$, les n premières dérivées de f satisfont $F(x, f(x), f'(x), \dots, f^{(n)}(x)) = 0$.

Exemple 3.1 Soit $F(x, y_0, y_1, y_2) = y_2^2 + x^2 y_1^2 + y_0^2 + 1$. L'équation différentielle

$$F(x, y(x), y'(x), y''(x)) = (y''(x))^2 + (xy'(x))^2 + (y(x))^4 + 1 = 0,$$

n'a pas de solution. En effet, $F(x, y_0, y_1, y_2) \geq 1$ pour tout $(x, y_0, y_1, y_2) \in \mathbb{R}^4$. Par contre l'équation

$$(y''(x))^2 + (xy'(x))^2 + (y(x))^4 = 0,$$

admet une unique solution, la fonction constante $f(x) = 0$ alors que l'équation

$$(y''(x))^2 + (y(x))^4 = 0,$$

admet une infinité de solutions, comme par exemple

$$f(x) = \pm \frac{6}{(x - a)^2},$$

où $a \in \mathbb{R}$ est une constante arbitraire. Elle admet d'autres solutions qu'on ne peut pas exprimer en termes de fonctions élémentaires. \triangleleft

Remarque 3.1 On parle d'équations différentielles *ordinaires* (EDO) pour les différencier des **équations aux dérivées partielles (EDP)** qui, comme leur nom l'indique, font intervenir les dérivées partielles d'une fonction de plusieurs variables. Un exemple fameux d'EDP est l'équation de Laplace

$$\Delta f(x_1, x_2) = \frac{\partial^2 f}{\partial x_1^2} f(x_1, x_2) + \frac{\partial^2 f}{\partial x_2^2} f(x_1, x_2) = 0,$$

que nous avons déjà rencontrée dans l'exemple 1.46. Comme nous l'avons remarqué dans ce dernier, il est parfois possible de ramener une telle EDP à une ou plusieurs EDO.

Comme le montre l'exemple précédent, une EDO peut admettre un infinité de solutions. En fait c'est le cas de la plupart des équations qu'on rencontre dans la pratique. Il est donc généralement nécessaire de spécifier des conditions supplémentaires pour déterminer la solution recherchée.

Exemple 3.2 Le mouvement d'un point matériel de masse m sous l'action d'un champ de force $F(x)$ est déterminé par l'équation de Newton

$$m \frac{d^2x}{dt^2} = F(x(t)).$$

Il s'agit d'une EDO du second ordre. Pour déterminer complètement la trajectoire de point matériel, on doit connaître sa position ainsi que sa vitesse initiale $x(0)$ et $x'(0)$. \triangleleft

Cet exemple est typique d'une classe importante de problèmes qu'on peut formuler de la manière suivante. Le **problème de Cauchy** pour l'EDO (3.1) consiste à déterminer les solutions de cette équation satisfaisant les **conditions initiales**

$$y(x_0) = c_0, \quad y'(x_0) = c_1, \quad \dots \quad y^{(n-1)}(x_0) = c_{n-1},$$

où $x_0, c_0, \dots, c_{n-1} \in \mathbb{R}$ sont des constantes fixées.

Exemple 3.3 Le mouvement d'un oscillateur harmonique est déterminé par l'équation de Newton

$$m \frac{d^2x}{dt^2}(t) = -fx(t),$$

où f est la constante de rappel de l'oscillateur. La solution du problème de Cauchy pour cette EDO avec les conditions initiales

$$x(0) = x_0, \quad x'(0) = v_0,$$

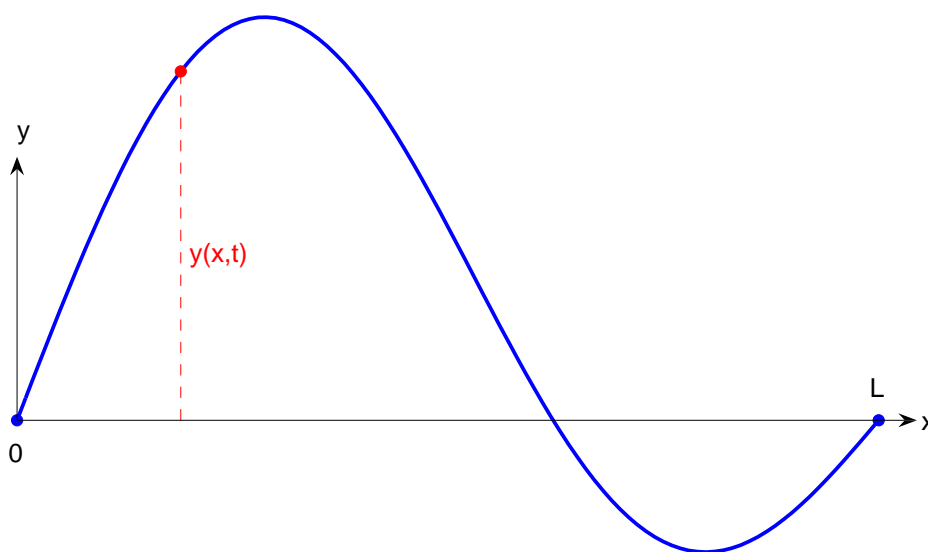
est unique et donnée par la fonction

$$x(t) = x_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t,$$

où $\omega = \sqrt{f/m}$ est la fréquence circulaire de l'oscillateur. \triangleleft

Une autre catégorie de problèmes importants est illustrée par l'exemple suivant.

Exemple 3.4 On peut décrire les petites déformations d'une corde de piano (ou de violon) tendue entre deux points séparés par une distance L comme le graphe d'une fonction $y : [0, L] \rightarrow \mathbb{R}$ satisfaisant les conditions $y(0) = y(L) = 0$ qui traduisent le fait que la corde est fixée à ses extrémités. Lorsque cette corde vibre, sa forme dépend également du temps et est donnée par une fonction $y(x, t)$.



L'évolution de la forme de cette corde est décrite par une EDP, l'équation de la corde vibrante,

$$\mu \frac{\partial^2 y}{\partial t^2}(x, t) = T \frac{\partial^2 y}{\partial x^2}(x, t), \quad (3.2)$$

assortie des conditions $y(0, t) = y(L, t) = 0$. Les constantes μ et T sont respectivement la densité linéique de la corde et la tension à laquelle elle est soumise. On peut réduire cette EDP à une EDO en cherchant des solutions périodiques en temps de la forme

$$y(x, t) = \phi(x) \cos \omega t,$$

où ω est la fréquence (inconnue) de vibration. En insérant cette expression dans l'équation (3.2), on obtient

$$T \frac{d^2 \phi}{dx^2}(x) = \mu \omega^2 \phi(x). \quad (3.3)$$

On remarque que cette équation différentielle a la même forme que celle de l'oscillateur harmonique de l'exemple 3.3. Cependant les conditions additionnelles sont différentes puisqu'on doit requérir que $\phi(0) = \phi(L) = 0$. Il ne s'agit plus ici de conditions initiales, mais de **conditions aux bords** de l'intervalle $[0, L]$ sur lequel on cherche une solution. Dans ce cas on parle de **problème de valeurs aux bords** pour l'EDO 3.3. Nous verrons dans le paragraphe 3.1.4 que toutes les solutions de cette dernière équation sont de la forme

$$\phi(x) = A \cos kx + B \sin kx,$$

où A, B sont des constantes arbitraires et $k = \omega \sqrt{\mu/T}$. La condition $\phi(0) = 0$ nous force à poser $A = 0$. La condition $\phi(L) = 0$ se traduit alors par

$$B \sin kL = 0.$$

La solution $B = 0$ n'est pas acceptable puisqu'elle correspond à la corde au repos $\phi(x) = 0$. On doit donc avoir $\sin kL = 0$, c'est-à-dire

$$k \in \{n\pi/L \mid n = 1, 2, 3, \dots\}.$$

Les conditions aux bords fixent donc les fréquences d'oscillation de la corde

$$\omega = \sqrt{\frac{T}{\mu}} k \in \left\{ n \sqrt{\frac{T}{\mu}} \frac{\pi}{L} \mid n = 1, 2, 3, \dots \right\}.$$

La fréquence de base est $\omega_0 = \sqrt{T/\mu}$. Les autres fréquences en sont des multiples ou harmoniques : $2\omega_0, 3\omega_0, \dots$ ◁

Avant de passer à l'étude de quelques classes importantes d'EDO, quelques lignes sur leur classification. On dit que (3.1) est une **EDO linéaire** lorsque F est une fonction linéaire de $y(x), y'(x), \dots, y^{(n)}(x)$. Dans ce cas on peut écrire l'équation différentielle sous la forme

$$a_0(x)y^{(n)}(x) + a_1(x)y^{(n-1)}(x) + \dots + a_n(x)y(x) = h(x). \quad (3.4)$$

Les fonctions $a_0(x), \dots, a_n(x)$ sont appelés **coefficients** de l'équation alors que la fonction $h(x)$ en est le **terme inhomogène**. L'équation est à **coefficients constants** lorsque les fonctions $a_0(x), \dots, a_n(x)$ sont constantes. Elle est **homogène** lorsque le terme inhomogène $h(x)$ est nul pour tout x , **inhomogène** dans le cas contraire.

3.1 Equations linéaires

Dans cette section, nous discutons les propriétés élémentaires de l'EDO linéaire (3.4). Pour simplifier l'écriture, nous lui associons l'**opérateur différentiel** d'ordre n (c.f. l'intermède p. 65)

$$L = a_0(x) \frac{d^n}{dx^n} + a_1(x) \frac{d^{n-1}}{dx^{n-1}} + \dots + a_n(x). \quad (3.5)$$

Nous supposons que les coefficients $a_0(x), \dots, a_n(x)$ sont des fonctions définies sur un intervalle I . Par définition, L agit sur l'espace vectoriel des fonctions n fois différentiables sur cet intervalle,

$$(Lf)(x) = a_0(x) \frac{d^n f}{dx^n}(x) + a_1(x) \frac{d^{n-1} f}{dx^{n-1}}(x) + \dots + a_n(x) f(x).$$

On peut donc écrire l'équation différentielle (3.4) sous la forme condensée

$$Ly = h.$$

3.1.1 Le problème homogène

Dans ce paragraphe, nous concentrons notre attention sur le problème homogène

$$Ly = 0, \tag{3.6}$$

associé à l'opérateur différentiel (3.5). Le résultat fondamental est le théorème d'existence et d'unicité suivant.

Théorème 3.2 *Si les coefficients $a_0(x), \dots, a_n(x)$ sont continus sur l'intervalle I et si $a_0(x)$ ne s'annule jamais sur cet intervalle, alors pour tout $x_0 \in I$ et pour tout $k \in \{0, \dots, n-1\}$ le problème de Cauchy pour l'EDO (3.6) avec conditions initiales*

$$y^{(j)}(x_0) = \delta_{jk} = \begin{cases} 1 & \text{si } j = k; \\ 0 & \text{sinon,} \end{cases}$$

admet une et une seule solution $y_k(x)$ sur l'intervalle I .

On dira que les fonctions $y_0(x), \dots, y_{n-1}(x)$ forment un **système fondamental de solutions** de l'équation homogène (3.6). On notera qu'à tout $x_0 \in I$ on peut associer un tel système.

Exemple 3.5 L'équation linéaire homogène du second ordre

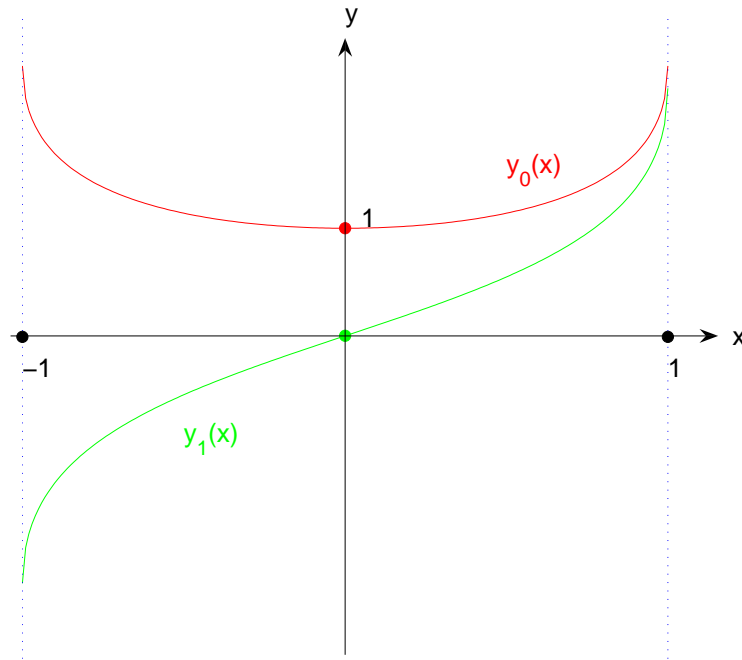
$$(Ly)(x) = (x^2 - 1)y''(x) + xy'(x) + y(x) = 0,$$

admet sur l'intervalle $I =]-1, 1[$, le système fondamental de solutions

$$\begin{aligned} y_0(x) &= \operatorname{ch}\left(\frac{\pi}{2} - \arccos(x)\right), \\ y_1(x) &= \operatorname{sh}\left(\frac{\pi}{2} - \arccos(x)\right), \end{aligned}$$

déterminé par les conditions initiales

$$\begin{aligned} y_0(0) &= 1, & y_0'(0) &= 0, \\ y_1(0) &= 0, & y_1'(0) &= 1. \end{aligned}$$



On remarque sur cet exemple que la condition $a_0(x) \neq 0$ sur l'intervalle I est nécessaire. En effet les solutions $y_0(x)$ et $y_1(x)$ sont singulières aux points $x = \pm 1$ où le coefficient $a_0(x) = x^2 - 1$ s'annule,

$$\lim_{x \rightarrow \pm 1} y_0'(x) = \pm\infty, \quad \lim_{x \rightarrow \pm 1} y_1'(x) = +\infty.$$

◁

Preuve du Théorème 3.2. Comme $a_0(x) \neq 0$ pour $x \in I$, la matrice

$$A(x) = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\frac{a_n(x)}{a_0(x)} & -\frac{a_{n-1}(x)}{a_0(x)} & -\frac{a_{n-2}(x)}{a_0(x)} & \cdots & -\frac{a_1(x)}{a_0(x)} \end{pmatrix}, \quad (3.7)$$

est bien définie et ses composantes sont des fonctions continues. On vérifie immédiatement que la fonction $y : I \rightarrow \mathbb{R}$ est solution de $Ly = 0$ si et seulement si la fonction $Y : I \rightarrow \mathbb{R}^n$ définie par

$$Y(x) = \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix},$$

vérifie le système d'EDO du premier ordre

$$Y'(x) = A(x)Y(x).$$

De même, $y_0(x), \dots, y_{n-1}(x)$ est un système fondamental de solutions de $Ly_k = 0$ avec $y_k^{(j)}(x_0) = \delta_{jk}$ si et seulement si la matrice

$$\Phi(x) = \begin{pmatrix} y_0(x) & y_1(x) & y_2(x) & \cdots & y_{n-1}(x) \\ y_0'(x) & y_1'(x) & y_2'(x) & \cdots & y_{n-1}'(x) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_0^{(n-1)}(x) & y_1^{(n-1)}(x) & y_2^{(n-1)}(x) & \cdots & y_{n-1}^{(n-1)}(x) \end{pmatrix},$$

est solution du problème de Cauchy

$$\Phi'(x) = A(x)\Phi(x), \quad \Phi(x_0) = I. \quad (3.8)$$

Soit $\Phi(x)$ une solution de ce problème de Cauchy. Nous démontrons tout d'abord que cette solution est unique. En itérant l'identité

$$\Phi(x) = \Phi(x_0) + \int_{x_0}^x \Phi'(u) du = I + \int_{x_0}^x A(u)\Phi(u) du,$$

on obtient

$$\Phi(x) = I + \int_{x_0}^x A(u) \left[I + \int_{x_0}^u A(v)\Phi(v) dv \right] du = I + \int_{x_0}^x A(u) du + \int_{x_0}^x \left[\int_{x_0}^u A(u)A(v)\Phi(v) dv \right] du.$$

Après N itérations on arrive à

$$\Phi(x) = \sum_{k=0}^{N-1} \Phi_k(x) + R_N(x), \quad (3.9)$$

où $\Phi_0(x) = I$,

$$\Phi_k(x) = \int_{x_0}^x \left[\int_{x_0}^{x_1} \left[\cdots \left[\int_{x_0}^{x_{k-1}} A(x_1)A(x_2) \cdots A(x_k) dx_k \right] \cdots \right] dx_2 \right] dx_1, \quad (3.10)$$

pour $k \geq 1$ et

$$R_N(x) = \int_{x_0}^x \left[\int_{x_0}^{x_1} \left[\cdots \left[\int_{x_0}^{x_{N-1}} A(x_1)A(x_2) \cdots A(x_N)\Phi(x_N) dx_N \right] \cdots \right] dx_2 \right] dx_1.$$

Soit $J \subset I$ un intervalle fermé contenant x_0 . Comme A et Φ sont des fonctions continues, on a

$$a_J = \max_{x \in J} \|A(x)\| < \infty, \quad \varphi_J = \max_{x \in J} \|\Phi(x)\| < \infty,$$

et donc, pour tout $x \in J$, l'estimation

$$\begin{aligned} \|R_N(x)\| &\leq \int_{x_0}^x \left[\int_{x_0}^{x_1} \left[\cdots \left[\int_{x_0}^{x_{N-1}} \|A(x_1)\| \|A(x_2)\| \cdots \|A(x_N)\| \|\Phi(x_N)\| dx_N \right] \cdots \right] dx_2 \right] dx_1 \\ &\leq a_J^N \varphi_J \int_{x_0}^x \left[\int_{x_0}^{x_1} \left[\cdots \left[\int_{x_0}^{x_{N-1}} dx_N \right] \cdots \right] dx_2 \right] dx_1 \\ &\leq a_J^N \varphi_J \frac{|x - x_0|^N}{N!}. \end{aligned}$$

On en déduit que pour tout $x \in J$

$$\lim_{N \rightarrow \infty} R_N(x) = 0,$$

et comme $J \subset I$ est arbitraire, cette relation est vérifiée pour tout $x \in I$. L'identité (3.9) nous permet de conclure que la solution du problème de Cauchy (3.8), si elle existe, est donnée par

$$\Phi(x) = \lim_{N \rightarrow \infty} \sum_{k=0}^{N-1} \Phi_k(x) = \sum_{k=0}^{\infty} \Phi_k(x), \quad (3.11)$$

cette série étant convergente pour tout $x \in I$. Ceci démontre l'unicité de la solution du problème de Cauchy. Pour démontrer son existence, il nous suffit de montrer que la fonction définie par (3.10), (3.11) est bien une solution. On vérifie immédiatement que $\Phi(x_0) = \Phi_0(x_0) = I$, $\Phi'_0(x) = 0$ et $\Phi'_k(x) = A(x)\Phi_{k-1}(x)$. De l'identité (3.10) on déduit l'estimation

$$\max_{x \in J} \|\Phi_k(x)\| \leq \frac{(a_J R)^k}{k!}, \quad R = \max_{x \in J} |x - x_0|,$$

qui nous permet de conclure que la série (3.11) ainsi que la série dérivée terme à terme

$$\sum_{k=0}^{\infty} \Phi_k'(x) = A(x) \sum_{k=0}^{\infty} \Phi_k(x),$$

sont uniformément convergentes sur tout intervalle fermé $J \subset I$. La série (3.11) définit donc une fonction Φ différentiable sur I et telle que $\Phi'(x) = A(x)\Phi(x)$. \square

Comme nous l'avons déjà remarqué (c.f. page 66), un opérateur différentiel L est une application linéaire. Une propriété qui reflète bien entendu la linéarité de l'EDO (3.4). Il en résulte que l'ensemble des solutions de l'équation homogène

$$\{y : I \rightarrow \mathbb{R} \mid y \text{ est } n \text{ fois différentiable et } Ly = 0\} = \text{Ker } L,$$

le noyau de L , est un sous-espace vectoriel de l'espace des fonctions n fois différentiables sur l'intervalle I . Si f et g sont deux solutions de l'équation $Ly = 0$ et si $\lambda, \mu \in \mathbb{R}$ alors la fonction $\lambda f + \mu g$ est aussi une solution de cette équation. C'est le principe de superposition qui joue un rôle extrêmement important dans la théorie des équations différentielles linéaires et dont les implications, par exemple à la physique, sont très profondes. En particulier, si y_0, \dots, y_{n-1} est un système fondamental de solutions

$$Ly_k = 0, \quad y_k^{(j)}(x_0) = \delta_{jk}, \quad k, j \in \{0, \dots, n-1\},$$

alors, pour tout $c_0, \dots, c_{n-1} \in \mathbb{R}^n$, la fonction

$$y(x) = \sum_{k=0}^{n-1} c_k y_k(x),$$

est une solution du problème de Cauchy

$$Ly = 0, \quad y(x_0) = c_0, \quad y'(x_0) = c_1, \quad \dots, \quad y^{(n-1)}(x_0) = c_{n-1}. \quad (3.12)$$

Sous les hypothèses du théorème 3.2 nous pouvons donc conclure que le problème de Cauchy (3.12) admet toujours une et une seule solution.

Théorème 3.3 *Sous les hypothèses du théorème 3.2, l'ensemble des solutions de l'EDO homogène $Ly = 0$ sur l'intervalle I est un espace vectoriel de dimension n . Tout système fondamental de solution de cette ODE forme une base de cet espace. Pour tout $x_0 \in I$ et pour tous c_0, c_1, \dots, c_{n-1} le problème de Cauchy pour l'ODE $Ly = 0$ avec les conditions initiales*

$$y(x_0) = c_0, \quad y'(x_0) = c_1, \quad \dots, \quad y^{(n-1)}(x_0) = c_{n-1},$$

admet une unique solution sur l'intervalle I . De plus, si y_0, \dots, y_{n-1} est un système fondamental de solution de cette ODE tel que $y_k^{(j)}(x_0) = \delta_{jk}$, alors cette solution est donnée par

$$y = \sum_{k=0}^{n-1} c_k y_k(x).$$

Dans un espace vectoriel dont les éléments sont des fonctions, le rôle du vecteur nul est tenu par la fonction constante, identiquement nulle. En particulier, les fonctions f_0, \dots, f_k sont linéairement indépendantes sur l'intervalle I si, pour tous $\alpha_0, \dots, \alpha_k \in \mathbb{R}$, la condition

$$\alpha_0 f_0(x) + \dots + \alpha_k f_k(x) = 0 \text{ pour tout } x \in I,$$

implique que $\alpha_0 = \dots = \alpha_k = 0$. Le lemme suivant nous donne une caractérisation très utile de l'indépendance linéaire d'une famille de fonctions différentiables.

Lemme 3.4 On appelle *wronskien* du système de fonctions $\{f_0, \dots, f_k\}$, k fois différentiables sur l'intervalle I , le déterminant $w_x(f_0, \dots, f_k)$ de la matrice

$$W_x(f_0, \dots, f_k) = \begin{pmatrix} f_0(x) & f_1(x) & \cdots & f_k(x) \\ f_0'(x) & f_1'(x) & \cdots & f_k'(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_0^{(k)}(x) & f_1^{(k)}(x) & \cdots & f_k^{(k)}(x) \end{pmatrix}.$$

Si il existe $x_0 \in I$ tel que $w_{x_0}(f_0, \dots, f_k) \neq 0$, alors les fonctions f_0, \dots, f_k sont linéairement indépendantes.

Démonstration. Supposons que $\alpha_0 f_0(x) + \dots + \alpha_k f_k(x) = 0$ pour tout $x \in I$. En dérivant successivement cette identité en x_0 on obtient

$$\begin{aligned} \alpha_0 f_0(x_0) + \dots + \alpha_k f_k(x_0) &= 0, \\ \alpha_0 f_0'(x_0) + \dots + \alpha_k f_k'(x_0) &= 0, \\ &\vdots \\ \alpha_0 f_0^{(k)}(x_0) + \dots + \alpha_k f_k^{(k)}(x_0) &= 0, \end{aligned}$$

ce qui signifie que le vecteur $(\alpha_0, \dots, \alpha_k)$ appartient au noyau de la matrice $W_{x_0}(f_0, \dots, f_k)$. Comme cette matrice est régulière par hypothèse, son noyau est trivial $\text{Ker}W_{x_0}(f_0, \dots, f_k) = \{(0, \dots, 0)\}$. Par conséquence $\alpha_0 = \dots = \alpha_k = 0$ et les fonctions f_0, \dots, f_k sont linéairement indépendantes. \square

Exemple 3.6 Dans l'exemple 1.46 nous avons rencontré l'équation différentielle

$$r(rf')' = f,$$

qu'on peut aussi écrire sous la forme

$$(Lf)(r) = r^2 f''(r) + rf'(r) - f(r) = 0.$$

Nous avons obtenu deux solutions de cette équation, à savoir

$$f_0(r) = r^{-1}, \quad f_1(r) = r.$$

La matrice wronskienne de ces deux solutions est

$$W_r(f_0, f_1) = \begin{pmatrix} r^{-1} & r \\ -r^{-2} & 1 \end{pmatrix},$$

et par conséquence $w_r(f_0, f_1) = \det W_r(f_0, f_1) = 2r^{-1} \neq 0$. Ces deux solutions sont donc linéairement indépendantes et forment une base du noyau de L . Toute solution de l'équation $Ly = 0$ est de la forme

$$y(r) = \frac{A}{r} + Br.$$

En particulier, pour tout $R > 0$, on peut déterminer une famille fondamentale de solutions $y_0(r)$, $y_1(r)$ telles que $y_0(R) = 1$, $y_0'(R) = 0$, $y_1(R) = 0$ et $y_1'(R) = 1$ en posant

$$y_0(r) = \frac{A_0}{r} + B_0 r, \quad y_1(r) = \frac{A_1}{r} + B_1 r,$$

et en résolvant le système

$$\begin{aligned} y_0(R) &= \frac{A_0}{R} + B_0 R = 1, \\ y_0'(R) &= -\frac{A_0}{R^2} + B_0 = 0, \\ y_1(R) &= \frac{A_1}{R} + B_1 R = 0, \\ y_1'(R) &= -\frac{A_1}{R^2} + B_1 = 1, \end{aligned}$$

pour les constantes inconnues A_0, B_0, A_1, B_1 . On obtient ainsi

$$y_0(r) = \frac{1}{2} \left(\frac{R}{r} + \frac{r}{R} \right), \quad y_1(r) = -\frac{R}{2} \left(\frac{R}{r} - \frac{r}{R} \right).$$

3.1.2 Le problème inhomogène

Pour résoudre le problème de Cauchy homogène

$$Ly = h,$$

avec conditions initiales

$$y(x_0) = c_0, \quad y'(x_0) = c_1, \quad \dots \quad y^{(n-1)}(x_0) = c_{n-1},$$

on peut suivre la méthode générale suivante, simple application du principe de superposition. On détermine tout d'abord une solution arbitraire de l'EDO inhomogène $Ly = h$ (dans la terminologie des EDO, on parle d'une [solution particulière](#)). Comme nous le démontrerons plus bas, l'existence d'une telle solution est assurée, sous les hypothèses du théorème 3.2, si le terme inhomogène $h(x)$ est continu sur l'intervalle I . Soit donc $g(x)$ une telle solution. En général elle ne satisfait pas les conditions initiales. On remarque cependant que si $f(x)$ est une solution quelconque de l'équation homogène $Ly = 0$, alors $L(f + g) = Lf + Lg = 0 + h = h$. Réciproquement, si \tilde{g} est une seconde solution particulière de l'équation inhomogène alors $L(\tilde{g} - g) = L\tilde{g} - Lg = h - h = 0$, et on peut écrire $\tilde{g} = g + f$ où $f = \tilde{g} - g$ est une solution de l'équation homogène. Nous en concluons que toute solution de l'équation inhomogène peut s'obtenir en ajoutant une solution de l'équation homogène à une solution particulière de l'équation inhomogène. On obtient donc la solution du problème de Cauchy inhomogène sous la forme

$$y(x) = g(x) + f(x),$$

où $g(x)$ est notre solution particulière de l'équation inhomogène et $f(x)$ est la solution du problème de Cauchy homogène

$$Lf = 0, \quad f(x_0) = c_0 - g(x_0), \quad \dots \quad f^{(n-1)}(x_0) = c_{n-1} - g^{(n-1)}(x_0).$$

Exemple 3.7 Considérons le problème de Cauchy inhomogène sur l'intervalle $I =]-1, 1[$,

$$(Ly)(x) = (x^2 - 1)y''(x) + xy'(x) + y(x) = x,$$

avec les conditions initiales

$$y(0) = 1, \quad y'(0) = 0.$$

On obtient une solution particulière en faisant l'Ansatz $g(x) = ax + b$. En effet,

$$(Lg)(x) = (x^2 - 1)g''(x) + xg'(x) + g(x) = xa + ax + b,$$

nous montre que $g(x) = x/2$ est bien solution de l'équation inhomogène. Nous devons maintenant résoudre le problème de Cauchy homogène

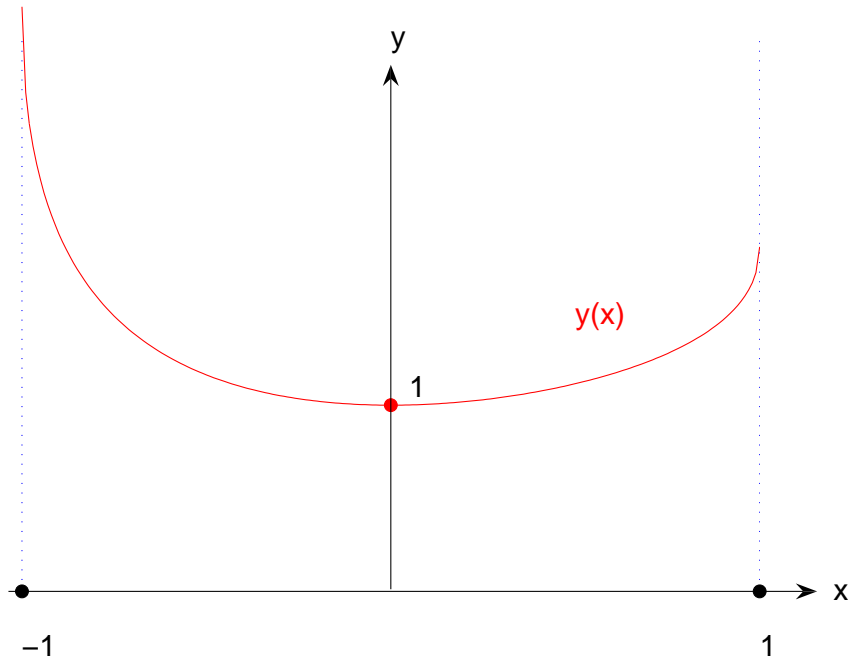
$$Lf = 0, \quad f(0) = 1 - g(0) = 1, \quad f'(0) = 0 - g'(0) = -1/2.$$

A l'aide du système fondamental de solutions obtenu dans l'exemple 3.5, nous pouvons écrire la solution de ce problème comme

$$f(x) = 1 y_0(x) - \frac{1}{2} y_1(x) = \operatorname{ch}\left(\frac{\pi}{2} - \arccos(x)\right) - \frac{1}{2} \operatorname{sh}\left(\frac{\pi}{2} - \arccos(x)\right).$$

La solution du problème de Cauchy inhomogène est donc

$$y(x) = \frac{x}{2} + \operatorname{ch}\left(\frac{\pi}{2} - \arccos(x)\right) - \frac{1}{2} \operatorname{sh}\left(\frac{\pi}{2} - \arccos(x)\right).$$



◁

Démonstration de l'existence d'une solution particulière. A l'aide de la méthode de [variation des constantes](#) que nous allons exposer maintenant, il est possible de construire une solution particulière de l'équation inhomogène $Ly = h$ si on possède une famille fondamentale de solutions de l'équation homogène $Ly = 0$. Une application de cette méthode sera faite dans l'exemple 3.16.

Soit $y_1(x), \dots, y_n(x)$ une famille fondamentale de solutions de l'équation $Ly = 0$. Comme nous l'avons vu au paragraphe précédent, toutes les solutions de cette équation sont de la forme

$$y(x) = \sum_{k=0}^{n-1} c_k y_k(x),$$

où les c_k sont des constantes. Nous cherchons une solution de l'équation inhomogène $Ly = h$ sous la forme

$$y(x) = \sum_{k=0}^{n-1} c_k(x) y_k(x),$$

où les $c_k(x)$ sont des fonctions (d'où le nom de la méthode : variation des constantes). Pour déterminer ces n fonctions inconnues, nous avons une seule équation, à savoir $Ly = h$. Nous devons donc imposer $n - 1$ conditions supplémentaires pour pouvoir déterminer ces inconnues. La première dérivée de y est donnée par

$$y'(x) = \sum_{k=0}^{n-1} c'_k(x) y_k(x) + \sum_{k=0}^{n-1} c_k(x) y'_k(x).$$

En imposant la première condition

$$\sum_{k=0}^{n-1} c'_k(x) y_k(x) = 0,$$

nous pouvons donc écrire

$$y'(x) = \sum_{k=0}^{n-1} c_k(x) y'_k(x).$$

La deuxième dérivée de y est donc

$$y''(x) = \sum_{k=0}^{n-1} c'_k(x) y'_k(x) + \sum_{k=0}^{n-1} c_k(x) y''_k(x).$$

En imposant comme seconde condition

$$\sum_{k=0}^{n-1} c'_k(x) y'_k(x) = 0,$$

nous obtenons

$$y''(x) = \sum_{k=0}^{n-1} c_k(x) y''_k(x).$$

Nous pouvons clairement continuer ce processus, en imposant les $n - 1$ conditions

$$\sum_{k=0}^{n-1} c'_k(x) y_k^{(j)}(x) = 0, \quad (3.13)$$

pour $j = 0, 1, \dots, n - 2$ on obtient les expressions

$$y^{(j)}(x) = \sum_{k=0}^{n-1} c_k(x) y_k^{(j)}(x),$$

pour $j = 0, 1, \dots, n - 1$ et

$$y^{(n)}(x) = \sum_{k=0}^{n-1} c'_k(x) y_k^{(n-1)}(x) + \sum_{k=0}^{n-1} c_k(x) y_k^{(n)}(x).$$

La dernière condition, $Ly = h$ s'écrit alors

$$\begin{aligned} h &= a_0(x)y^{(n)}(x) + a_1(x)y^{(n-1)}(x) + \dots + a_n(x)y(x) \\ &= a_0(x) \sum_{k=0}^{n-1} c'_k(x)y_k^{(n-1)}(x) + \sum_{k=0}^{n-1} c_k(x) \left(a_0(x)y_k^{(n)}(x) + \dots + a_n(x)y_k(x) \right) \\ &= a_0(x) \sum_{k=0}^{n-1} c'_k(x)y_k^{(n-1)}(x) + \sum_{k=0}^{n-1} c_k(x)(Ly_k)(x) \\ &= a_0(x) \sum_{k=0}^{n-1} c'_k(x)y_k^{(n-1)}(x). \end{aligned}$$

En résumé, nous devons résoudre le système

$$\begin{aligned} y_0(x)c'_0(x) + y_1(x)c'_1(x) + \dots + y_{n-1}(x)c'_{n-1}(x) &= 0, \\ y'_0(x)c'_0(x) + y'_1(x)c'_1(x) + \dots + y'_{n-1}(x)c'_{n-1}(x) &= 0, \\ &\vdots \\ y_0^{(n-2)}(x)c'_0(x) + y_1^{(n-2)}(x)c'_1(x) + \dots + y_{n-1}^{(n-2)}(x)c'_{n-1}(x) &= 0, \\ y_0^{(n-1)}(x)c'_0(x) + y_1^{(n-1)}(x)c'_1(x) + \dots + y_{n-1}^{(n-1)}(x)c'_{n-1}(x) &= \frac{h(x)}{a_0(x)}. \end{aligned}$$

En écrivant ces n conditions sous forme matricielle

$$\begin{pmatrix} y_0(x) & y_1(x) & \dots & y_{n-1}(x) \\ y'_0(x) & y'_1(x) & \dots & y'_{n-1}(x) \\ \vdots & \vdots & \ddots & \vdots \\ y_0^{(n-1)}(x) & y_1^{(n-1)}(x) & \dots & y_{n-1}^{(n-1)}(x) \end{pmatrix} \begin{pmatrix} c'_0(x) \\ c'_1(x) \\ \vdots \\ c'_{n-1}(x) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ h(x)/a_0(x) \end{pmatrix}, \quad (3.14)$$

on réalise que la matrice du système n'est autre que $W_x(y_0, \dots, y_{n-1})$ qui est régulière par le lemme 3.4. En dénotant par $c(x)$ le vecteur de composantes $c_k(x)$ et par $H(x)$ le vecteur du membre de droite de (3.14), on a

$$c'(x) = W_x(y_0, \dots, y_{n-1})^{-1}H(x),$$

qu'on peut intégrer immédiatement pour obtenir

$$c(x) = \int_0^x W_u(y_0, \dots, y_{n-1})^{-1}H(u) \, du,$$

ce qui achève la construction d'une solution particulière. □

3.1.3 Le cas d'ordre 1

C'est le cas le plus facile. On peut toujours le réduire à des quadratures, c'est-à-dire au calcul de quelques intégrales. Considérons tout d'abord le problème homogène

$$a_0(x)y'(x) + a_1(x)y(x) = 0.$$

En divisant par $a_0(x)y_0(x)$ on obtient l'équation

$$-\frac{a_1(x)}{a_0(x)} = \frac{y'(x)}{y(x)} = \frac{d}{dx} \log |y(x)|.$$

qui s'intègre immédiatement

$$\log |y(x)| - \log |y(x_0)| = p(x) = - \int_{x_0}^x \frac{a_1(u)}{a_0(u)} \, du.$$

On a donc

$$|y(x)| = |y(x_0)| e^{p(x)},$$

où $|y(x_0)| = c_0$ est une constante arbitraire. Comme le membre de droite de cette identité ne s'annule jamais, le signe de $y(x)$ est constant et on peut l'absorber dans la constante c_0 pour obtenir

$$y(x) = c_0 e^{p(x)}.$$

La solution fondamentale $y_0(x)$ satisfaisant la condition initiale $y_0(x_0) = 1$ est donc

$$y_0(x) = e^{p(x)}, \quad p(x) = - \int_{x_0}^x \frac{a_1(u)}{a_0(u)} \, du.$$

En particulier, la solution du problème de Cauchy

$$a_0(x)y'(x) + a_1(x)y(x) = 0, \quad y(x_0) = c_0,$$

est donnée par

$$y(x) = c_0 y_0(x) = c_0 e^{p(x)}, \quad p(x) = - \int_{x_0}^x \frac{a_1(u)}{a_0(u)} \, du.$$

Considérons maintenant le cas inhomogène

$$a_0(x)y'(x) + a_1(x)y(x) = h(x).$$

Nous cherchons une solution particulière par la méthode de variation des constantes qui consiste à faire l'Ansatz

$$y(x) = c_0(x)e^{p(x)}, \quad p(x) = - \int_{x_0}^x \frac{a_1(u)}{a_0(u)} \, du,$$

c'est-à-dire à remplacer la constante dans la solution du problème homogène par une fonction de la variable indépendante x . Nous obtenons ainsi

$$a_0(x) \left(c_0'(x) e^{p(x)} + c_0(x) p'(x) e^{p(x)} \right) + a_1(x) c_0(x) e^{p(x)} = h(x),$$

et, compte tenu du fait que $a_0(x)c_0(x)p'(x)e^{p(x)} + a_1(x)c_0(x)e^{p(x)} = 0$,

$$c_0'(x) = \frac{h(x)}{a_0(x)} e^{-p(x)}.$$

Une solution particulière de l'équation inhomogène est donc donnée par la fonction

$$g(x) = c_0(x) e^{p(x)} = \left[\int_{x_0}^x \frac{h(v)}{a_0(v)} e^{-p(v)} dv \right] e^{p(x)}.$$

En notant que cette fonction satisfait la condition initiale $g(x_0) = 0$, on obtient la solution du problème de Cauchy inhomogène

$$a_0(x)y'(x) + a_1(x)y(x) = h(x), \quad y(x_0) = c_0,$$

sous la forme

$$y(x) = \left[c_0 + \int_{x_0}^x \frac{h(v)}{a_0(v)} e^{-p(v)} dv \right] e^{p(x)}, \quad p(x) = - \int_{x_0}^x \frac{a_1(u)}{a_0(u)} du.$$

Exemple 3.8 Soit le problème de Cauchy inhomogène

$$\frac{y'(x)}{x} + y(x) = x^2, \quad y(0) = 1.$$

L'exposant $p(x)$ est donné par

$$p(x) = - \int_0^x \frac{1}{u-1} du = -\frac{x^2}{2}.$$

Une solution particulière du problème homogène est

$$g(x) = \int_0^x \frac{v^2}{v-1} e^{v^2/2} dv = 2 \int_0^{x^2/2} t e^{-t} dt = 2 + (x^2 - 2) e^{x^2/2}.$$

Nous obtenons ainsi la solution de notre problème

$$y(x) = \left[1 + 2 + (x^2 - 2) e^{x^2/2} \right] e^{-x^2/2} = 3e^{-x^2/2} + x^2 - 2.$$

◁

3.1.4 Le cas d'ordre 2

Bien que ce soit le cas le plus important pour les applications, il n'existe pas de méthode générale pour réduire la solution d'un problème de Cauchy homogène du second ordre à un nombre fini de quadratures. Cependant, beaucoup de ces problèmes ont été étudiés en détail, et leurs solutions sont répertoriées (voir par exemple [AS]). C'est le cas en particulier de toutes les EDO qui proviennent de la séparation des variables dans les équations aux dérivées partielles de la physique. On donne le nom générique de **fonctions spéciales** aux solutions de ces problèmes.

Exemple 3.9 L'équation des ondes est l'EDP

$$\frac{\partial^2 \mathbf{u}}{\partial t^2}(x, t) = \Delta \mathbf{u}(x, t),$$

qui régit la propagation de divers types d'ondes (sonore, électromagnétiques,...). Nous en avons déjà rencontré un cas particulier dans l'exemple 3.4. Comme dans ce dernier, on peut chercher des solutions périodiques en temps en faisant l'Ansatz de séparation des variables $u(x, t) = \varphi(x) \cos \omega t$. La fonction φ doit alors satisfaire l'équation d'Helmholtz

$$\Delta \varphi(x) + \omega^2 \varphi(x) = 0.$$

Une famille importantes de solutions de cette équations sont les ondes planes. En dimension 2 elles sont données par

$$\varphi_{k,\phi}(x_1, x_2) = \cos(k_1 x_1 + k_2 x_2 + \phi),$$

où $k = (k_1, k_2) \in \mathbb{R}^2$ est le **vecteur d'onde** et ϕ la **phase** (des solutions analogues existent en dimension quelconque). On vérifie immédiatement que

$$\Delta \varphi_{k,\phi}(x) = \frac{\partial^2 \varphi_{k,\phi}}{\partial x_1^2}(x_1, x_2) + \frac{\partial^2 \varphi_{k,\phi}}{\partial x_2^2}(x_1, x_2) = -(k_1^2 + k_2^2) \varphi_{k,\phi}(x_1, x_2),$$

et par conséquent que $\varphi_{k,\phi}$ satisfait l'équation d'Helmholtz sous la condition

$$\|k\|^2 = k_1^2 + k_2^2 = \omega^2.$$

La fonction $\varphi_{k,\phi}$ décrit une onde se propageant dans le plan dans la direction k en absence de tout obstacle. Lorsque la géométrie est plus complexe, on doit résoudre l'équation d'Helmholtz dans un domaine Ω et l'assujettir de conditions au bord de ce domaine.

Comme illustration, nous allons calculer les fréquences de vibration d'une membrane circulaire de rayon R fixée sur sa circonférence (la peau d'un tambour). Elles sont décrites par l'équation d'Helmholtz sur le disque de rayon R avec la condition $\varphi(x) = 0$ sur le bord de ce disque. Il est dans ce cas naturel de formuler le problème en coordonnées polaires. A l'aide de la formule (1.47), on peut alors écrire

$$\frac{\partial^2 \varphi}{\partial r^2}(r, \theta) + \frac{1}{r} \frac{\partial \varphi}{\partial r}(r, \theta) + \frac{1}{r^2} \frac{\partial^2 \varphi}{\partial \theta^2}(r, \theta) + \omega^2 \varphi(r, \theta) = 0, \quad \varphi(R, \theta) = 0. \quad (3.15)$$

Avec l'Ansatz de séparation des variables

$$\varphi(r, \theta) = f(\omega r) \cos n\theta,$$

on obtient

$$\omega^2 f''(\omega r) \cos n\theta + \frac{\omega}{r} f'(\omega r) \cos n\theta - \frac{n^2}{r^2} f(\omega r) \cos n\theta + \omega^2 f(\omega r) \cos n\theta = 0, \quad f(\omega R) \cos n\theta = 0.$$

En simplifiant par $\cos n\theta$ on arrive à l'ODE

$$\omega^2 f''(\omega r) + \frac{\omega}{r} f'(\omega r) - \frac{n^2}{r^2} f(\omega r) + \omega^2 f(\omega r) = 0, \quad f(\omega R) = 0.$$

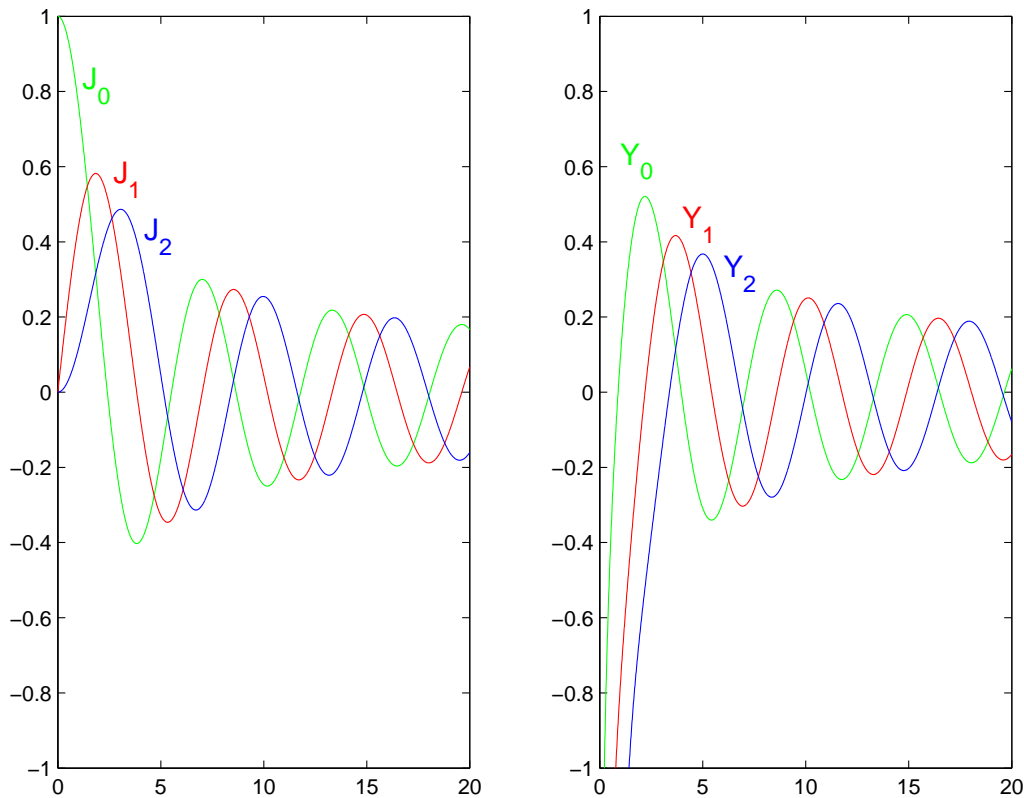
Après le changement de variable $x = \omega r$, cette dernière s'écrit

$$\omega^2 f''(x) + \frac{\omega^2}{x} f'(x) - \frac{\omega^2 n^2}{x^2} f(x) + \omega^2 f(x) = 0, \quad f(\omega R) = 0.$$

On obtient finalement, après simplification, l'équation de Bessel

$$x^2 f''(x) + x f'(x) + (x^2 - n^2) f(x) = 0.$$

Deux solutions linéairement indépendantes de cette dernière sont les fonctions de Bessel de première espèce $J_n(x)$ et de seconde espèce $Y_n(x)$.



La fonction Y_n est singulière en $x = 0$, on ne peut donc pas l'accepter comme solution de notre problème. La fonction J_n est régulière, elle est donnée par l'intégrale

$$J_n(x) = \frac{1}{\pi} \int_0^\pi \cos(nt - x \sin t) dt,$$

on peut aussi la représenter à l'aide de la série

$$J_n(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k!(k+n)!} \left(\frac{x}{2}\right)^{2k+n},$$

(voir [W] pour tout savoir sur ces fonctions très importantes). On a donc

$$f(\omega r) = J_n(\omega r),$$

et la condition au bord, $f(\omega R) = 0$, fixe la fréquence de vibration

$$\omega \in \left\{ \frac{x_{n,k}}{R} \mid k = 1, 2, 3, \dots \right\},$$

où $x_{n,1}, x_{n,2}, x_{n,3}, \dots$ sont les zéros de J_n . Les premiers termes de cette suite de fréquences sont

n	k	$x_{n,k}$
0	1	2.4048
1	1	3.8317
2	1	5.1356
0	2	5.5201
3	1	6.3802
1	2	7.0156

on remarque en particulier que, contrairement aux fréquences de vibrations de la corde de l'exemple 3.4 qui sont en progression arithmétique ($\omega_0, 2\omega_0, 3\omega_0, \dots$), il n'existe aucune relation simple entre les termes de cette suite. ◁

Exemple 3.10 Le polynôme de Legendre de degré l ,

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2 - 1)^l,$$

où l est un entier, est une solution de l'équation de Legendre

$$(1 - x^2)y''(x) - 2xy'(x) + l(l + 1)y(x) = 0,$$

qui apparaît naturellement lorsqu'on écrit l'équation de Laplace $\Delta f = 0$ dans \mathbb{R}^3 en coordonnées sphériques. La fonction de Legendre associée de degré l et d'ordre m

$$P_l^m(x) = (1 - x^2)^{m/2} \frac{d^m}{dx^m} P_l(x),$$

est une solution de l'équation de Legendre associée

$$(1 - x^2)y''(x) - 2xy'(x) + \left(l(l + 1) - \frac{m^2}{1 - x^2} \right) y(x) = 0.$$

◁

Exemple 3.11 La fonction hypergéométrique

$$F(a, b; c; x) = \sum_{n=0}^{\infty} \frac{a(a+1) \cdots (a+n-1) b(b+1) \cdots (b+n-1) x^n}{c(c+1) \cdots (c+n-1) n!},$$

est une solution de l'équation hypergéométrique

$$x(1-x)y''(x) + (c - (a+b+1)x)y'(x) - aby(x) = 0. \quad (3.16)$$

Il est souvent possible de transformer une ODE du second ordre en une de ces équations bien connues. L'équation

$$A(x - \alpha)(x - \beta)y''(x) + B(x - \gamma)y'(x) + Cy(x) = 0, \quad (3.17)$$

est équivalent à l'équation hypergéométrique. En faisant le changement de coordonnées $x = (\beta - \alpha)u + \alpha$ et en posant $y(x) = y((\beta - \alpha)u + \alpha) = z(u)$ on a

$$z'(u) = y'((\beta - \alpha)u + \alpha)(\beta - \alpha), \quad z''(u) = y''((\beta - \alpha)u + \alpha)(\beta - \alpha)^2,$$

ainsi que

$$(x - \alpha) = (\beta - \alpha)u, \quad (x - \beta) = (\beta - \alpha)(u - 1), \quad (x - \gamma) = (\beta - \alpha) \left(u - \frac{\gamma - \alpha}{\beta - \alpha} \right).$$

L'équation (3.17) devient donc

$$\begin{aligned} 0 &= A(\beta - \alpha)^2 u(u - 1)y''((\beta - \alpha)u + \alpha) \\ &+ B(\beta - \alpha) \left(u - \frac{\gamma - \alpha}{\beta - \alpha} \right) y'((\beta - \alpha)u + \alpha) + Cy((\beta - \alpha)u + \alpha) \\ &= -A \left[u(1 - u)f''(u) + \left(\frac{B}{A} \frac{\gamma - \alpha}{\beta - \alpha} - \frac{B}{A} u \right) f'(u) - \frac{C}{A} f(u) \right]. \end{aligned}$$

En comparant cette dernière expression avec l'équation (3.16) on voit que si

$$ab = \frac{C}{A}, \quad a + b + 1 = \frac{B}{A}, \quad c = \frac{B}{A} \frac{\gamma - \alpha}{\beta - \alpha},$$

alors la fonction hypergéométrique $f(u) = F(a, b; c; u)$ est une solution. On résout facilement pour les paramètres a, b, c le système ci-dessus. Avec

$$a = \frac{(B - A) + \sqrt{(B - A)^2 - 4AC}}{2A}, \quad b = \frac{(B - A) - \sqrt{(B - A)^2 - 4AC}}{2A}, \quad c = \frac{B}{A} \frac{\gamma - \alpha}{\beta - \alpha}.$$

une solution de (3.17) est donc

$$y(x) = F \left(a, b; c; \frac{x - \alpha}{\beta - \alpha} \right).$$

◁

Exemple 3.12 L'équation

$$xy''(x) - y'(x) + x^3y(x) = 0, \quad (3.18)$$

s'exprime plus simplement après un changement de variable $x^2 = u$. En posant $y(x) = z(x^2) = z(u)$ on a

$$y'(x) = 2xz'(x^2) = 2\sqrt{u}z'(u),$$

et

$$y''(x) = \frac{d}{dx}2xz'(x^2) = 4x^2z''(x^2) + 2z'(x^2) = 4uz''(u) + 2z'(u).$$

L'équation (3.18) devient donc

$$\sqrt{u}(4uz''(u) + 2z'(u)) - 2\sqrt{u}z'(u) + u\sqrt{u}z(u) = 0,$$

c'est-à-dire

$$z''(u) + \frac{1}{4}z(u) = 0.$$

Une solution de cette équation est $z(u) = \cos(u/2)$. On en déduit qu'une solution de l'équation (3.18) est $y(x) = \cos(x^2/2)$. \triangleleft

Lorsqu'on connaît une solution y_0 de l'équation homogène du second ordre

$$a_0(x)y''(x) + a_1(x)y'(x) + a_2(x)y(x) = 0. \quad (3.19)$$

la méthode de variation des constantes permet d'obtenir une solution particulière de l'équation inhomogène

$$a_0(x)y''(x) + a_1(x)y'(x) + a_2(x)y(x) = h(x). \quad (3.20)$$

Nous cherchons cette solution sous la forme $y(x) = c(x)y_0(x)$. En insérant cet Ansatz dans (3.20),

$$\begin{aligned} h(x) &= a_0(x)(c''(x)y_0(x) + 2c'(x)y_0'(x) + c(x)y_0''(x)) \\ &+ a_1(x)(c'(x)y_0(x) + c(x)y_0'(x)) + a_2(x)c(x)y_0(x), \end{aligned}$$

d'où, en tenant compte du fait que y_0 est solution de (3.19),

$$a_0(x)y_0(x)c''(x) + (2a_0(x)y_0'(x) + a_1(x)y_0(x))c'(x) = h(x),$$

qui est une EDO du premier ordre pour la fonction $d(x) = c'(x)$. En divisant par $a_0(x)y_0(x)$ on la met sous la forme

$$d'(x) + \left(2\frac{y_0'(x)}{y_0(x)} + \frac{a_1(x)}{a_0(x)}\right)d(x) = \frac{h(x)}{a_0(x)y_0(x)}.$$

En appliquant la méthode du paragraphe précédent, on obtient

$$d(x) = \frac{e^{p(x)}}{y_0(x)^2} \int_{x_0}^x \frac{h(v)y_0(v)}{a_0(v)} e^{-p(v)} dv, \quad p(x) = - \int_{x_0}^x \frac{a_1(u)}{a_0(u)} du.$$

Cette dernière relation est immédiatement intégrée pour obtenir

$$c(x) = \int_{x_0}^x \frac{e^{p(w)}}{y_0(w)^2} \left[\int_{x_0}^w \frac{h(v)y_0(v)}{a_0(v)} e^{-p(v)} dv \right] dw,$$

et par conséquence

$$y(x) = y_0(x) \int_{x_0}^x \frac{e^{p(w)}}{y_0(w)^2} \left[\int_{x_0}^w \frac{h(v)y_0(v)}{a_0(v)} e^{-p(v)} dv \right] dw. \quad (3.21)$$

La même technique permet d'obtenir, à partir de la solution y_0 , une seconde solution de l'équation homogène (3.19). Il suffit pour cela de poser $h = 0$ dans l'argument précédent. On obtient alors la seconde solution sous la forme

$$y_1(x) = y_0(x) \int_{x_0}^x \frac{e^{p(w)}}{y_0(w)^2} dw. \quad (3.22)$$

Exemple 3.13 On se propose de résoudre le problème de Cauchy pour l'EDO inhomogène

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

avec conditions initiales

$$y(0) = 1, \quad y'(0) = 1.$$

L'équation homogène correspondante peut s'écrire sous la forme

$$(1+x^2)y''(x) = 2y(x),$$

et on vérifie sans peine qu'elle admet la solution $y_0(x) = 1 + x^2$. Cherchons tout d'abord à construire une seconde solution de l'équation homogène de façon à obtenir un système fondamental de solution. En appliquant (3.22) nous obtenons

$$y_1(x) = (1+x^2) \int_0^x \frac{dw}{(1+w^2)^2} = (1+x^2) \left[\frac{1}{2} \frac{w}{1+w^2} + \frac{1}{2} \arctan w \right]_{w=0}^{w=x} = \frac{x}{2} + \frac{1+x^2}{2} \arctan x.$$

On vérifie facilement que y_0, y_1 forme un système fondamental de solutions de l'équation homogène

$$y_0(0) = 1, \quad y_0'(0) = 0, \quad y_1(0) = 0, \quad y_1'(0) = 1.$$

La formule (3.21) nous donne une solution particulière du problème inhomogène

$$\begin{aligned} g(x) &= (1+x^2) \int_0^x \left[\int_0^w v(1+v^2) dv \right] \frac{dw}{(1+w^2)^2} \\ &= (1+x^2) \int_0^x \left[\frac{w^2}{2} + \frac{w^4}{4} \right] \frac{dw}{(1+w^2)^2} \\ &= (1+x^2) \frac{1}{4} \int_0^x \left[1 - \frac{1}{(1+w^2)^2} \right] dw \\ &= \frac{1}{4} \left[(1+x^2) \left(x - \frac{1}{2} \arctan x \right) - \frac{x}{2} \right], \end{aligned}$$

dont on vérifie qu'elle satisfait les conditions initiales

$$g(0) = 0, \quad g'(0) = 0.$$

La solution du problème de Cauchy est donc

$$y(x) = g(x) + y_0(x) + y_1(x) = 1 + \frac{5}{8}x + x^2 + \frac{1}{4}x^2 + \frac{3}{8}(1+x^2) \arctan x.$$

◁

3.1.5 Equations à coefficients constants

Dans le cas particulier d'une ODE linéaire à coefficients constants $Ly = h$ avec

$$L = a_0 \frac{d^n}{dx^n} + a_1 \frac{d^{n-1}}{dx^{n-1}} + \dots + a_n. \tag{3.23}$$

il est possible de construire explicitement la solution du problème de Cauchy

$$y(x_0) = c_0, \quad y'(x_0) = c_1, \quad \dots \quad y^{(n-1)}(x_0) = c_{n-1}. \tag{3.24}$$

C'est cette construction que nous discutons dans ce paragraphe. Commençons par remarquer que l'opérateur différentiel L est invariant par translation : si $z(x) = y(x + x_0)$ alors $z^{(k)}(x) = y^{(k)}(x + x_0)$ et

$$(Lz)(x) = (Ly)(x + x_0).$$

Ainsi, y est solution du problème de Cauchy $(Ly)(x) = h(x)$ avec conditions initiales (3.24) si et seulement si z est solutions de $(Lz)(x) = \tilde{h}(x) = h(x + x_0)$ avec conditions initiales

$$z(0) = c_0, \quad z'(0) = c_1, \quad \dots \quad z^{(n-1)}(0) = c_{n-1}.$$

On peut donc, sans restriction de la généralité, supposer que $x_0 = 0$.

Pour $\lambda \in \mathbb{C}$, la fonction $y_\lambda(x) = e^{\lambda x}$ satisfait

$$\frac{d^n y_\lambda}{dx^n}(x) = \lambda^n y_\lambda(x).$$

Il en résulte que

$$Ly_\lambda = P_L(\lambda)y_\lambda,$$

où

$$P_L(z) = a_0 z^n + a_1 z^{n-1} + a_2 z^{n-2} + \dots + a_n,$$

est le **polynôme caractéristique** de l'opérateur différentiel à coefficient constant L . Si λ est une racine du polynôme caractéristique, on a donc

$$Ly_\lambda = P_L(\lambda)y_\lambda = 0,$$

et on obtient un système $\{y_\lambda \mid \lambda \in \mathbb{C}, p_L(\lambda) = 0\}$ de solutions de l'EDO homogène. Les coefficients du polynôme P_L étant réels, ses racines peuvent être réelles ou former des paires complexes conjuguées. Dans le premier cas y_λ est une fonction réelle. Dans le second cas, à la paire de racines $\lambda = a \pm ib$ correspond les deux solutions complexes

$$y_\lambda(x) = e^{(a+ib)x} = e^{ax}(\cos bx + i \sin bx), \quad y_{\bar{\lambda}}(x) = e^{(a-ib)x} = e^{ax}(\cos bx - i \sin bx).$$

L'équation $Ly = 0$ étant linéaire, on peut remplacer ces deux solutions par les solutions réelles

$$\frac{1}{2}(y_\lambda(x) + y_{\bar{\lambda}}(x)) = e^{ax} \cos bx, \quad \frac{1}{2i}(y_\lambda(x) - y_{\bar{\lambda}}(x)) = e^{ax} \sin bx.$$

Il est cependant pratique d'accepter les solutions complexes qui sont plus facile à manipuler, sachant qu'à la fin des calculs, la solution recherchée sera toujours réelle.

Le théorème fondamental de l'algèbre nous assure que $P_L(z)$, qui est de degré n , admet n racines comptées avec leur multiplicité. Si toutes les racines sont simples, le système de solution $\{y_\lambda \mid \lambda \in \mathbb{C}, p_L(\lambda) = 0\}$ qui comporte n fonctions est donc un bon candidat pour former une base du noyau de L . Cependant, si certaines de ces racines sont dégénérées (c'est-à-dire de multiplicité supérieure à 1), P_L a moins de n racines distinctes ! Dans ce cas, le système de solutions $\{y_\lambda \mid \lambda \in \mathbb{C}, p_L(\lambda) = 0\}$ est trop petit pour former une base du noyau de L .

Nous supposons tout d'abord que le polynôme caractéristique P_L admet n racines distinctes $\lambda_0, \dots, \lambda_{n-1}$. Dans ce cas, $y_{\lambda_0}, \dots, y_{\lambda_{n-1}}$ sont n solutions de l'équation homogène $Ly = 0$. Pour déterminer si ces solutions sont linéairement indépendantes, formons la matrice

$$W_0(y_{\lambda_0}, \dots, y_{\lambda_{n-1}}) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_0 & \lambda_1 & \dots & \lambda_{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_0^{n-1} & \lambda_1^{n-1} & \dots & \lambda_{n-1}^{n-1} \end{pmatrix}.$$

C'est une **matrice de Vandermonde** dont le déterminant est donné par

$$w_0(y_{\lambda_1}, \dots, y_{\lambda_n}) = \prod_{0 \leq i < j < n} (\lambda_j - \lambda_i).$$

Les racines λ_i étant distinctes, ce déterminant est non nul et le lemme 3.4 nous permet de conclure que les solutions $y_{\lambda_0}, \dots, y_{\lambda_{n-1}}$ sont linéairement indépendantes. Elles forment donc une base du noyau de L .

La solution du problème de Cauchy $Ly = 0$ avec les conditions initiales $y^{(k)}(0) = c_k$ pour $k = 0, \dots, n-1$ est donc de la forme

$$y(x) = \sum_{j=0}^{n-1} a_j y_{\lambda_j}(x),$$

où les coefficients inconnus a_j sont déterminés par les conditions initiales

$$y^{(k)}(0) = \sum_{j=0}^{n-1} a_j y_{\lambda_j}^{(k)}(0) = \sum_{j=0}^{n-1} a_j \lambda_j^k = c_k.$$

C'est un système de n équations ($k = 0, \dots, n-1$) pour les n inconnues a_0, \dots, a_{n-1} . Sous forme matricielle ce système s'écrit

$$W_0(y_{\lambda_0}, \dots, y_{\lambda_{n-1}})a = c.$$

et la solution en est

$$a = W_0(y_{\lambda_0}, \dots, y_{\lambda_{n-1}})^{-1}c.$$

On obtient ainsi la solution du problème de Cauchy homogène

$$y(x) = \sum_{ij=1}^n [W_0(y_{\lambda_0}, \dots, y_{\lambda_{n-1}})^{-1}]_{ji} c_i y_{\lambda_j}(x).$$

Exemple 3.14 L'équation du mouvement de l'oscillateur harmonique amorti est donnée par

$$m \frac{d^2x}{dt^2}(t) = -fx(t) - \gamma \frac{dx}{dt}(t),$$

A la force de rappel $-fx$, s'ajoute une force de friction

$$F_{\text{fric}} = -\gamma \frac{dx}{dt},$$

proportionnelle et opposée à la vitesse de l'oscillateur ($\gamma > 0$ est la constante de friction). La trajectoire de l'oscillateur est déterminée par le problème de Cauchy avec conditions initiales

$$x(0) = x_0, \quad \frac{dx}{dt}(0) = v_0.$$

L'opérateur différentiel de ce problème est

$$L = m \frac{d^2}{dt^2} + \gamma \frac{d}{dt} + f,$$

et son polynôme caractéristique

$$P_L(z) = mz^2 + \gamma z + f.$$

Ce dernier admet les racines

$$\begin{aligned} \lambda_0 &= \frac{-\gamma - \sqrt{\gamma^2 - 4mf}}{2m} = -\Gamma - \sqrt{\Gamma^2 - \omega_0^2}, \\ \lambda_1 &= \frac{-\gamma + \sqrt{\gamma^2 - 4mf}}{2m} = -\Gamma + \sqrt{\Gamma^2 - \omega_0^2}, \end{aligned}$$

où nous avons introduit le coefficient d'amortissement

$$\Gamma = \frac{\gamma}{2m},$$

et où

$$\omega_0 = \sqrt{\frac{f}{m}},$$

est la fréquence circulaire de l'oscillateur non amorti. Ces racines dégènèrent lorsque $\Gamma = \omega_0$, on parle alors d'amortissement critique. Nous traiterons ce cas lorsque nous aurons discuté la dégénérescence

des zéros du polynôme caractéristique. Pour un amortissement sous-critique $\Gamma < \omega_0$, on a deux racines complexes conjuguées

$$\lambda_0 = -\Gamma - i\Omega, \quad \lambda_1 = \bar{\lambda}_0 = -\Gamma + i\Omega, \quad \Omega = \sqrt{\omega_0^2 - \Gamma^2} > 0,$$

alors que pour un amortissement sur-critique $\Gamma > \omega_0$ les deux racines sont négatives

$$\lambda_0 = -\Gamma - \Delta, \quad \lambda_1 = -\Gamma + \Delta, \quad 0 < \Delta = \sqrt{\Gamma^2 - \omega_0^2} < \Gamma.$$

La matrice wronskienne est

$$W_0(\mathbf{y}_{\lambda_0}, \mathbf{y}_{\lambda_1}) = \begin{pmatrix} 1 & 1 \\ \lambda_0 & \lambda_1 \end{pmatrix},$$

son déterminant est donc $w_0(\mathbf{y}_{\lambda_0}, \mathbf{y}_{\lambda_1}) = \lambda_1 - \lambda_0 = 2\sqrt{\Gamma^2 - \omega_0^2}$ et son inverse

$$W_0(\mathbf{y}_{\lambda_0}, \mathbf{y}_{\lambda_1})^{-1} = \frac{1}{2\sqrt{\Gamma^2 - \omega_0^2}} \begin{pmatrix} \lambda_1 & -1 \\ -\lambda_0 & 1 \end{pmatrix}.$$

La solution du problème de Cauchy est

$$x(t) = a_1 e^{\lambda_0 t} + a_2 e^{\lambda_1 t},$$

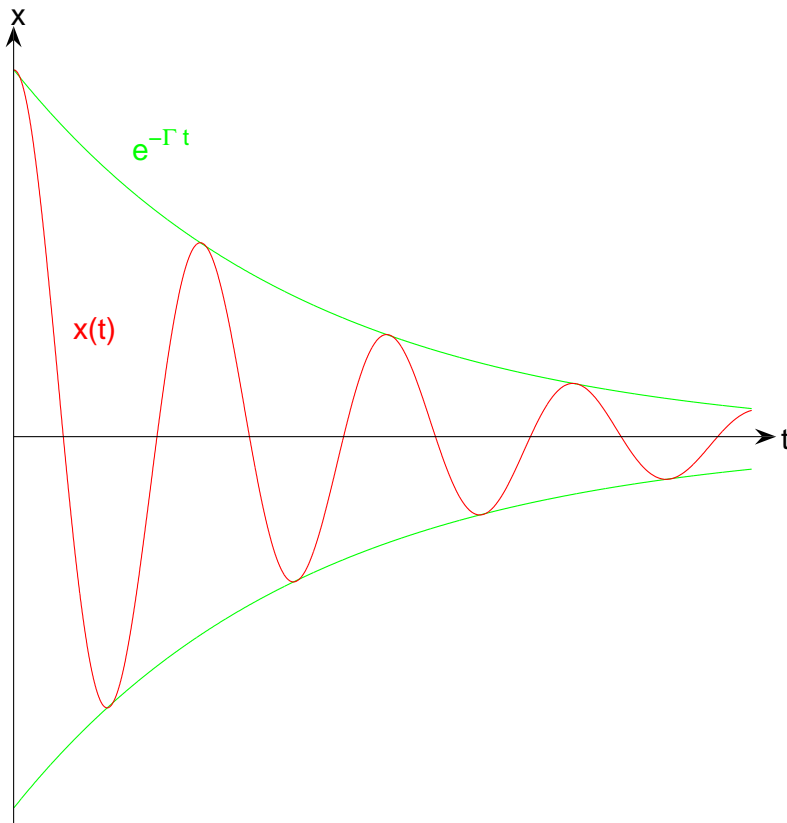
où

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \frac{1}{2\sqrt{\Gamma^2 - \omega_0^2}} \begin{pmatrix} \lambda_1 & -1 \\ -\lambda_0 & 1 \end{pmatrix} \begin{pmatrix} x_0 \\ v_0 \end{pmatrix}.$$

Après un peu d'algèbre élémentaire on obtient, dans le cas sous-critique $\Gamma < \omega_0$

$$x(t) = e^{-\Gamma t} \left[x_0 \cos \Omega t + \frac{\Gamma x_0 + v_0}{\Omega} \sin \Omega t \right], \quad (3.25)$$

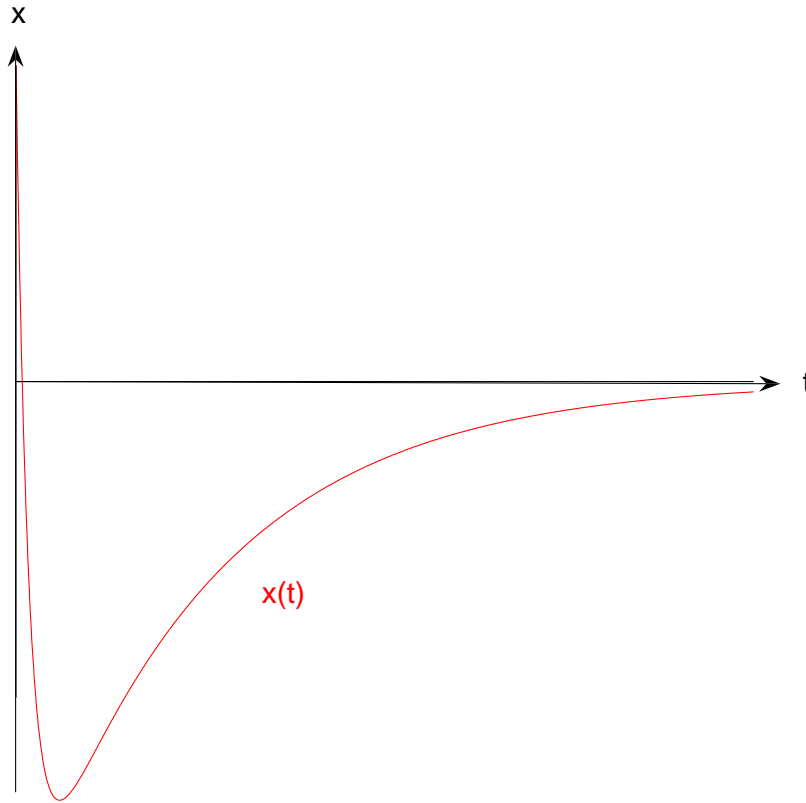
On observe que le mouvement est la superposition d'une enveloppe exponentiellement décroissante $e^{-\Gamma t}$ et d'une oscillation de fréquence circulaire $\Omega = \sqrt{\omega_0^2 - \Gamma^2} < \omega_0$.



Dans le cas sur-critique $\Gamma > \omega_0$ on a

$$x(t) = e^{-\Gamma t} \left[x_0 \operatorname{ch} \Delta t + \frac{\Gamma x_0 + v_0}{\Delta} \operatorname{sh} \Delta t \right], \quad (3.26)$$

les oscillations ont disparu, le mouvement est complètement dominé par la force de friction.



◁

Lorsqu'une racine λ du polynôme caractéristique P_L est dégénérée, de multiplicité r_λ , nous devons trouver $r_\lambda - 1$ solutions supplémentaires pour combler le déficit causé par cette dégénérescence. Dans ce cas, on a

$$P_L(\lambda) = P_L'(\lambda) = \dots = P_L^{(r_\lambda-1)}(\lambda) = 0, \quad P_L^{(r_\lambda)}(\lambda) \neq 0,$$

et en notant que

$$\left. \frac{d^k}{dt^k} y_{\lambda+t}(x) \right|_{t=0} = \left. \frac{d^k}{dt^k} e^{\lambda x} e^{tx} \right|_{t=0} = x^k y_\lambda(x),$$

on obtient, pour $k = 1, \dots, r_\lambda - 1$,

$$\begin{aligned} \left. \frac{d^k}{dt^k} L y_{\lambda+t} \right|_{t=0} &= \left. \frac{d^k}{dt^k} P_L(\lambda + t) y_{\lambda+t} \right|_{t=0} \\ &= \sum_{j=0}^k \binom{k}{j} \underbrace{P_L^{(j)}(\lambda)}_{=0} x^{k-j} y_\lambda(x) = 0. \end{aligned}$$

car $j \leq k < r$

La fonction $(t, x) \mapsto y_{\lambda+t}(x) = e^{\lambda x} e^{tx}$ étant de classe C^∞ , ses dérivées partielles ne dépendent pas de l'ordre de dérivation. On en conclut que

$$\left. \frac{d^k}{dt^k} L y_{\lambda+t} \right|_{t=0} = L \left. \frac{d^k}{dt^k} y_{\lambda+t} \right|_{t=0},$$

et par conséquence que les r_λ fonctions

$$x^k y_\lambda(x), \quad k = 0, \dots, r_\lambda - 1,$$

sont dans le noyau de L . On peut montrer que le système de fonctions

$$\bigcup_{\lambda \in \mathbb{C}, P_L(\lambda)=0} \{x^k y_\lambda(x) \mid k = 0, \dots, r_\lambda - 1\},$$

est une base du noyau de L . On peut donc écrire la solution du problème de Cauchy homogène comme

$$y(x) = \sum_{\lambda} \sum_{k=0}^{r_\lambda-1} a_{\lambda k} x^k y_\lambda(x),$$

où la somme sur λ se fait sur les racines distinctes de P_L . Les n coefficients inconnus $a_{\lambda k}$ sont déterminés par les conditions initiales

$$y^{(j)}(0) = \sum_{\lambda} \sum_{k=0}^{r_\lambda-1} a_{\lambda k} \left. \frac{d^j}{dx^j} x^k y_\lambda \right|_{x=0} = \sum_{\lambda} \sum_{k=0}^{r_\lambda-1} a_{\lambda k} j(j-1) \cdots (j-k+1) \lambda^{j-k} = c_j.$$

Exemple 3.15 Nous pouvons maintenant traiter le cas de l'amortissement critique de l'exemple 3.14. Lorsque $\Gamma = \omega_0$, le polynôme caractéristique admet une seule racine double $\lambda = -\Gamma$. Les deux fonctions

$$e^{-\Gamma t}, \quad t e^{-\Gamma t},$$

forment dans ce cas une base du noyau de L . La solution du problème de Cauchy est de la forme

$$x(t) = (a + bt)e^{-\Gamma t},$$

et les constantes a, b sont déterminées par les conditions initiales

$$x(0) = a = x_0, \quad x'(0) = b - a\Gamma = v_0.$$

On obtient donc

$$x(t) = (x_0 + (v_0 + \Gamma x_0)t)e^{-\Gamma t}.$$

Il est intéressant de noter qu'on retrouve cette formule en prenant, dans les expressions (3.25) ou (3.26) la limite $\Omega \rightarrow 0$, respectivement $\Delta \rightarrow 0$. \triangleleft

Finalement, le problème homogène $Ly = h$ peut être résolu à l'aide de la méthode de variation des constantes (c.f. p. 111). Nous nous contenterons de l'illustrer par un exemple.

Exemple 3.16 L'oscillateur harmonique amorti sous l'influence d'une force extérieure F obéit à l'ODE

$$m \frac{d^2 x}{dt^2}(t) = -fx(t) - \gamma \frac{dx}{dt}(t) + F(t). \quad (3.27)$$

Avec la notation de l'exemple 3.14, cette équation s'écrit

$$Lx = F.$$

Dans le cas non critique, une base du noyau de L est donnée par les deux fonctions $e^{\lambda_0 t}$ et $e^{\lambda_1 t}$. Nous cherchons donc une solution particulière du problème inhomogène sous la forme

$$x_{in}(t) = a(t)e^{\lambda_0 t} + b(t)e^{\lambda_1 t}, \quad (3.28)$$

où $a(t)$ et $b(t)$ sont des fonctions inconnues. En dérivant cet Ansatz,

$$x'_{in}(t) = (a'(t)e^{\lambda_0 t} + b'(t)e^{\lambda_1 t}) + (a(t)\lambda_0 e^{\lambda_0 t} + b(t)\lambda_1 e^{\lambda_1 t}),$$

on impose la première condition

$$a'(t)e^{\lambda_0 t} + b'(t)e^{\lambda_1 t} = 0, \quad (3.29)$$

pour obtenir

$$x'_{\text{in}}(t) = a(t)\lambda_0 e^{\lambda_0 t} + b(t)\lambda_1 e^{\lambda_1 t}. \quad (3.30)$$

Une seconde dérivation donne

$$x''_{\text{in}}(t) = (a'(t) + \lambda_0 a(t))\lambda_0 e^{\lambda_0 t} + (b'(t) + \lambda_1 b(t))\lambda_1 e^{\lambda_1 t}. \quad (3.31)$$

En insérant les expressions (3.28) (3.30) et (3.31) dans l'équation (3.27), et en tenant compte du fait que $e^{\lambda_0 t}$ et $e^{\lambda_1 t}$ sont des solutions de l'équation homogène on obtient la seconde condition

$$a'(t)\lambda_0 e^{\lambda_0 t} + b'(t)\lambda_1 e^{\lambda_1 t} = \frac{F(t)}{m}.$$

Avec la condition (3.29), cette dernière équation forme le système

$$\begin{pmatrix} e^{\lambda_0 t} & e^{\lambda_1 t} \\ \lambda_0 e^{\lambda_0 t} & \lambda_1 e^{\lambda_1 t} \end{pmatrix} \begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} = \begin{pmatrix} 0 \\ F(t)/m \end{pmatrix}.$$

En inversant la matrice de ce système on l'écrit sous la forme

$$\begin{pmatrix} a'(t) \\ b'(t) \end{pmatrix} = \frac{1}{\lambda_1 - \lambda_0} \begin{pmatrix} \lambda_1 e^{-\lambda_0 t} & -e^{-\lambda_0 t} \\ -\lambda_0 e^{-\lambda_1 t} & e^{-\lambda_1 t} \end{pmatrix} \begin{pmatrix} 0 \\ F(t)/m \end{pmatrix} = \frac{1}{\lambda_1 - \lambda_0} \begin{pmatrix} -e^{-\lambda_0 t} F(t)/m \\ e^{-\lambda_1 t} F(t)/m \end{pmatrix}.$$

Nous sommes maintenant en mesure d'intégrer pour obtenir les fonctions recherchées

$$a(t) = -\frac{1}{\lambda_1 - \lambda_0} \int_0^t e^{-\lambda_0 u} \frac{F(u)}{m} du,$$

et

$$b(t) = \frac{1}{\lambda_1 - \lambda_0} \int_0^t e^{-\lambda_1 u} \frac{F(u)}{m} du.$$

Nous avons ainsi obtenu une solution particulière de l'équation inhomogène

$$x_{\text{in}}(t) = \frac{1}{\lambda_1 - \lambda_0} \int_0^t \left[e^{\lambda_1(t-u)} - e^{\lambda_0(t-u)} \right] \frac{F(u)}{m} du, \quad (3.32)$$

telle que

$$x'_{\text{in}}(t) = \frac{1}{\lambda_1 - \lambda_0} \int_0^t \left[\lambda_1 e^{\lambda_1(t-u)} - \lambda_0 e^{\lambda_0(t-u)} \right] \frac{F(u)}{m} du.$$

Notez qu'il n'y a pas de contribution à cette dérivée provenant de la borne supérieure d'intégration car l'intégrand de (3.32) s'annule en $u = t$. On se convainc aisément que c'est la manifestation de la condition (3.29). Notre solution particulière satisfait donc les conditions initiales

$$x_{\text{in}}(0) = 0, \quad x'_{\text{in}}(0) = 0.$$

On obtient la solution du problème de Cauchy inhomogène $Lx = F$, $x(0) = x_0$, $x'(0) = v_0$ en ajoutant x_{in} à la solution du problème homogène obtenue précédemment. Dans le cas sous-critique, on a donc

$$x(t) = e^{-\Gamma t} \left[x_0 \cos \Omega t + \frac{\Gamma x_0 + v_0}{\Omega} \sin \Omega t \right] + \int_0^t e^{-\Gamma(t-u)} \frac{\sin \Omega(t-u)}{\Omega} \frac{F(u)}{m} du,$$

alors que dans le cas sur-critique

$$x(t) = e^{-\Gamma t} \left[x_0 \text{ch } \Delta t + \frac{\Gamma x_0 + v_0}{\Delta} \text{sh } \Delta t \right] + \int_0^t e^{-\Gamma(t-u)} \frac{\text{sh } \Delta(t-u)}{\Delta} \frac{F(u)}{m} du.$$

Il est particulièrement intéressant d'étudier la réponse de l'oscillateur en régime sous-critique avec un forçage périodique

$$F(t) = F_0 \sin \omega t,$$

d'amplitude F_0 et de fréquence ω . L'évaluation de l'intégrale (3.32) est un simple exercice dont on peu écrire le résultat sous la forme

$$x(t) = x_{\text{trans}}(t) + x_{\infty}(t), \quad (3.33)$$

avec

$$x_{\text{trans}}(t) = e^{-\Gamma t} \left[\left(x_0 + A_\infty(\omega) \frac{F_0}{m} \right) \cos \Omega t + \left(\frac{\Gamma x_0 + v_0}{\Omega} + C_\infty(\omega) \frac{\omega F_0}{\Omega m} \right) \sin \Omega t \right],$$

$$x_\infty(t) = -\frac{F_0}{m} (A_\infty(\omega) \cos \omega t + B_\infty(\omega) \sin \omega t),$$

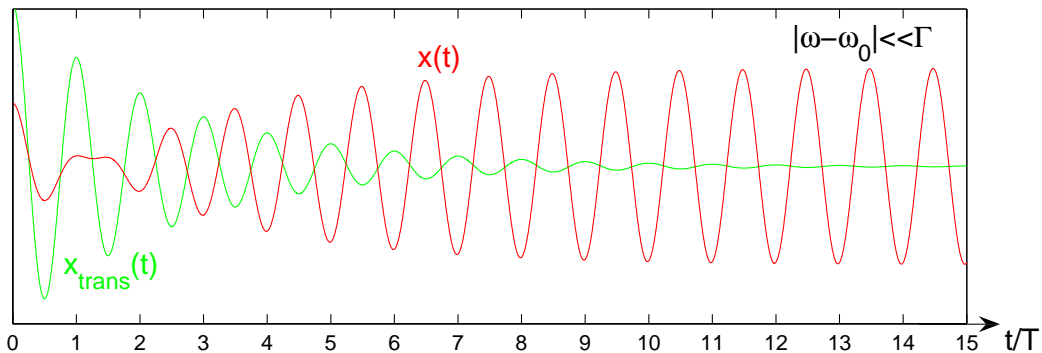
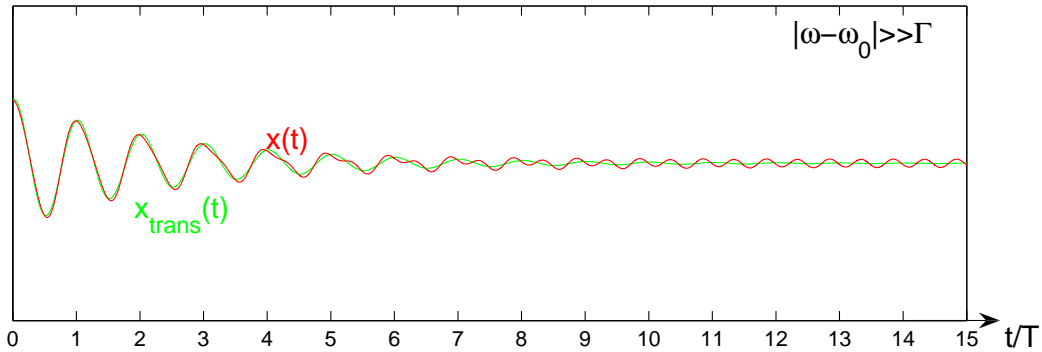
où

$$A_\infty(\omega) = \frac{2\Gamma\omega}{(\Gamma^2 + (\Omega + \omega)^2)(\Gamma^2 + (\Omega - \omega)^2)},$$

$$B_\infty(\omega) = \frac{\omega^2 - \Omega^2 - \Gamma^2}{(\Gamma^2 + (\Omega + \omega)^2)(\Gamma^2 + (\Omega - \omega)^2)},$$

$$C_\infty(\omega) = \frac{\omega^2 - \Omega^2 + \Gamma^2}{(\Gamma^2 + (\Omega + \omega)^2)(\Gamma^2 + (\Omega - \omega)^2)}.$$

Dans (3.33) nous avons séparé deux composantes du mouvement clairement identifiées. La première, $x_{\text{trans}}(t)$, de fréquence Ω (c'est la fréquence de l'oscillateur amorti non forcé) garde en mémoire la position et la vitesse initiale. C'est un phénomène transitoire qui décroît exponentiellement. Lorsque $t \rightarrow \infty$, seule survit la seconde composante, $x_\infty(t)$, purement périodique, dont la fréquence est celle du forçage. La figure suivante représente deux trajectoires typiques, avec les mêmes conditions initiales mais des fréquences de forçage différentes. Dans la première $|\omega - \omega_0| \gg \Gamma$ alors que $|\omega - \omega_0| \ll \Gamma$ dans la seconde (notez que l'abscisse est en nombre de périodes $T = 2\pi/\Omega$ de l'oscillateur amorti non forcé). Nous allons tenter de comprendre qualitativement ce qui distingue ces deux cas.



On peut écrire la composante périodique sous la forme

$$x_\infty(t) = -\frac{F_0}{m} (A_\infty(\omega) \cos \omega t + B_\infty(\omega) \sin \omega t) = -\frac{F_0}{m} D(\omega) \cos(\omega t - \phi(\omega)),$$

où l'amplitude est donné par

$$D(\omega) = \sqrt{A_\infty^2(\omega) + B_\infty^2(\omega)} = \frac{1}{\sqrt{(\Gamma^2 + (\Omega + \omega)^2)(\Gamma^2 + (\Omega - \omega)^2)}},$$

et le déphasage par

$$\phi(\omega) = \arcsin\left(\frac{B_\infty(\omega)}{\sqrt{A_\infty^2(\omega) + B_\infty^2(\omega)}}\right) = \arcsin\left(\frac{\omega^2 - \omega_0^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4\Gamma^2\omega^2}}\right).$$

Dans le régime asymptotique la puissance moyenne dissipée par l'oscillateur est égale au travail de la force $F(t)$ par unité de temps calculé sur une période $\tau = 2\pi/\omega$ de cette force.

$$P(\omega) = \frac{1}{\tau} \int_0^\tau x'_\infty(t)F(t) dt.$$

Comme

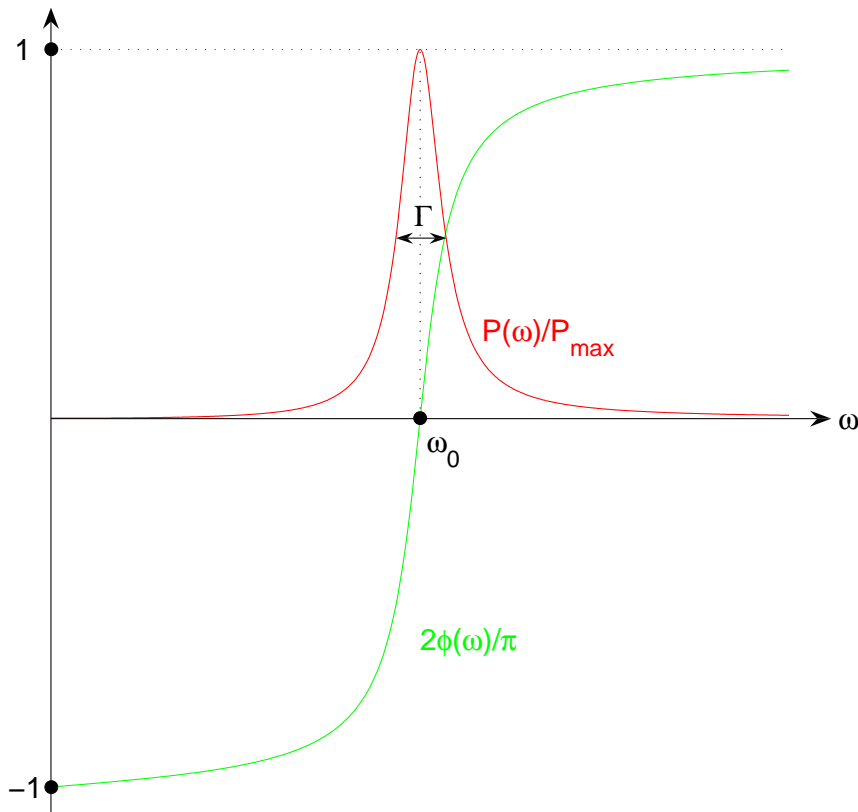
$$x'_\infty(t)F(t) = \frac{F_0^2}{m} D(\omega)\omega \sin(\omega t - \phi(\omega)) \sin \omega t = \frac{F_0^2}{2m} D(\omega)\omega (\cos \phi(\omega) - \cos(2\omega t - \phi(\omega))),$$

on a donc

$$P(\omega) = \frac{F_0^2}{2m} D(\omega)\omega \cos \phi(\omega) = \frac{F_0^2}{m} \frac{\Gamma\omega^2}{(\Gamma^2 + (\Omega + \omega)^2)(\Gamma^2 + (\Omega - \omega)^2)}.$$

A partir de cette formule on vérifie facilement que la puissance dissipée est maximale lorsque $\omega = \omega_0$,

$$P_{\max} = \max_{\omega > 0} P(\omega) = P(\omega_0) = \frac{F_0^2}{4m\Gamma}.$$



Il ressort clairement de la figure ci-dessus que lorsque $|\omega - \omega_0| \ll \Gamma$ le déphasage $\phi(\omega)$ est petit. La vitesse de l'oscillateur $x'_\infty(t)$ est alors en phase avec la force $F(t)$. Cette dernière effectue un travail optimal contre la force de friction, ce qui lui permet de maintenir une amplitude d'oscillation importante. Par contre, lorsque $|\omega - \omega_0| \gg \Gamma$, le déphasage est proche de $\pm\pi/2$. Dans ce cas la force s'oppose à la vitesse pendant une grande partie de la période τ . Son travail moyen sur une période est faible et par conséquence l'amplitude des oscillations est petite. \triangleleft

3.2 Equations non linéaires

En général, on ne peut pas réduire la solution d'une EDO non-linéaire

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)), \quad (3.34)$$

à des quadratures, c'est-à-dire au calcul d'un nombre fini d'intégrales, et ceci même pour des équations du premier ordre. Il est nécessaire d'utiliser des méthodes d'approximations. En pratique, on doit le plus souvent recourir à la résolution numérique de telles équations. En raison de l'une des plus importantes découvertes scientifiques du XX^{ème} siècle, celle du "chaos déterministe" (c.f. paragraphe 3.2.5), ce domaine relève dans certains cas du grand art. Nous n'aborderons pas ce sujet dans ce cours (voir [HNW],[HW] pour une introduction approfondie), nous nous contenterons d'explorer quelques cas simples qui se résolvent par quadrature et de formuler quelques résultats généraux pour les équations du premier ordre.

La première distinction entre EDO linéaires et non linéaires est assez évidente : le principe de superposition ne s'applique pas à ces dernières. Si $y_1(x)$ et $y_2(x)$ sont deux solutions de (3.34), leur somme $y_1(x) + y_2(x)$ n'en n'est généralement plus une. De même, si $c \in \mathbb{R}$ est une constante, $cy_1(x)$ n'est généralement plus une solution. Toutes les stratégies basées sur le principe de superposition et qui se sont révélées très utiles dans la section 3.1 ne fonctionnent plus dans le cas non linéaire.

Nous commençons par étudier une classe particulièrement simple d'EDO non linéaires du premier ordre pour mettre en évidence d'autres distinctions plus subtiles.

3.2.1 Equations séparables

Une EDO du premier ordre est **séparable** si elle peut s'écrire sous la forme

$$y'(x) = f(x)g(y(x)). \quad (3.35)$$

Si en plus f est constante, c'est-à-dire si

$$y'(x) = g(y(x)),$$

on dit que c'est une équation **autonome**. C'est un type d'équation très courant. Le problème de Cauchy correspondant avec condition initiale

$$y(x_0) = y_0, \quad (3.36)$$

se résout par quadrature de la manière suivante. Soit ϕ la primitive de la fonction $1/g$ qui s'annule en y_0 ,

$$\phi(u) = \int_{y_0}^u \frac{dv}{g(v)},$$

et F la primitive de la fonction f qui s'annule en x_0 ,

$$F(x) = \int_{x_0}^x f(v) dv.$$

Si $y(x)$ est une solution du problème de Cauchy et $G(x) = \phi(y(x)) - F(x)$ alors $G(x_0) = \phi(y(x_0)) - F(x_0) = \phi(y_0) - 0 = 0$ et

$$\frac{d}{dx} G(x) = \phi'(y(x))y'(x) - F'(x) = \frac{y'(x)}{g(y(x))} - f(x) = 0.$$

On en conclut que $G(x) = 0$ c'est-à-dire que

$$\phi(y(x)) = F(x). \quad (3.37)$$

Si la fonction ϕ est inversible, on applique son inverse aux deux membres de l'identité précédente pour obtenir

$$y(x) = \phi^{-1}(F(x)).$$

Exemple 3.17 Le problème de Cauchy

$$y'(x) = g(y(x)) = 1 + y(x)^2, \quad y(0) = \alpha, \quad (3.38)$$

est un cas particulièrement simple d'équation séparable. Dans ce cas

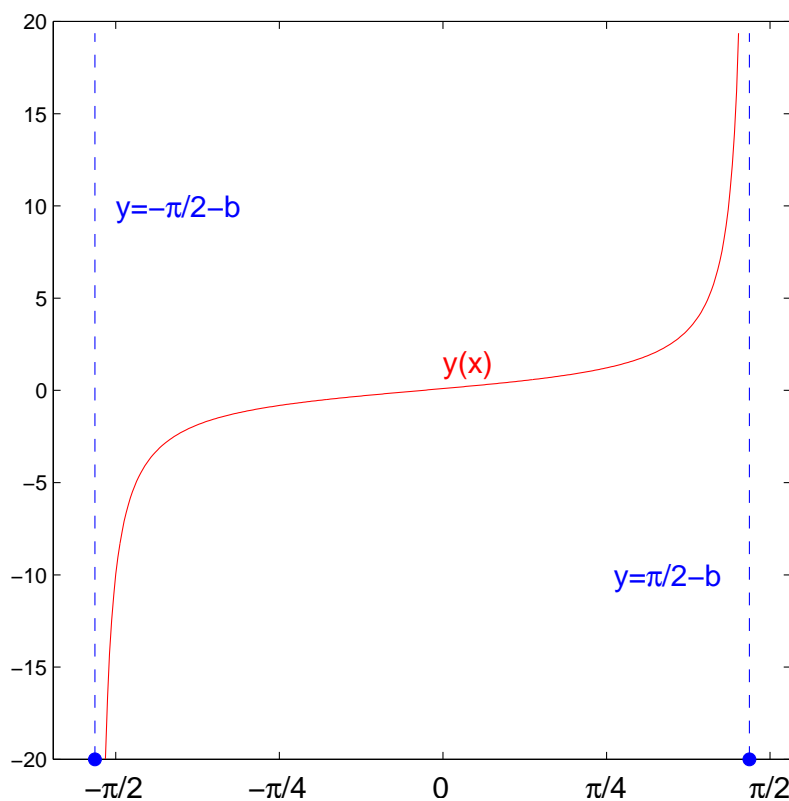
$$\phi(u) = \int_{\alpha}^u \frac{dv}{1+v^2} = \arctan u - \arctan \alpha, \quad F(x) = \int_0^x 1 \, dx = x,$$

et donc, en posant $b = \arctan \alpha \in]-\pi/2, \pi/2[$,

$$\arctan y(x) = b + x, \quad (3.39)$$

d'où on déduit

$$y(x) = \tan(b + x).$$



On rencontre ici un phénomène nouveau! Alors que les fonctions $f(x) = 1$ et $g(y) = 1 + y^2$ qui apparaissent dans l'équation différentielle sont parfaitement régulières (elle sont de classe C^∞ sur \mathbb{R} tout entier), la solution $y(x)$ de cette équation n'existe pas pour tous les x . Elle est singulière en $x = \pm\pi/2 - b$, plus précisément

$$\lim_{x \downarrow -\pi/2 - b} y(x) = -\infty, \quad \lim_{x \uparrow \pi/2 - b} y(x) = \infty.$$

Il est donc impossible d'étendre cette solution en dehors de l'intervalle $I =]-\pi/2 - b, \pi/2 - b[$. On dira que I est un **intervalle d'existence maximal** pour la solution $y(x)$ du problème de Cauchy (3.38) et que la fonction $I \ni x \mapsto \tan(x + b)$ en est une **solution maximale**.

Si on recherche l'origine du problème dans cet exemple, on remarque qu'il survient au moment d'inverser la fonction $\arctan u$. Cette dernière est strictement croissante, donc injective. Cependant elle n'est pas surjective, plus précisément son image est l'intervalle $]-\pi/2, \pi/2[$. La validité de la relation (3.39) cesse lorsque son membre de droite sort de cet intervalle!

Ce comportement est clairement dû au fait que la fonction $1/g(y)$ est intégrable à l'infini

$$\lim_{u \rightarrow \infty} \phi(u) = \lim_{u \rightarrow \infty} \int_{\alpha}^u \frac{dv}{g(v)} = \int_{\alpha}^{\infty} \frac{dv}{g(v)} = \frac{\pi}{2} - b < \infty,$$

et

$$\lim_{u \rightarrow -\infty} \phi(u) = \lim_{u \rightarrow -\infty} \int_a^u \frac{dv}{g(v)} = - \lim_{u \rightarrow -\infty} \int_u^a \frac{dv}{g(v)} = - \int_{-\infty}^a \frac{dv}{g(v)} = -b - \pi/2 > -\infty.$$

Comme ϕ est monotone croissante, il en résulte que $-\pi/2 - b < \phi(u) < \pi/2 - b$ pour tout $u \in \mathbb{R}$, c'est-à-dire que l'image de ϕ n'est pas \mathbb{R} tout entier. Il est instructif de comparer cette situation avec le cas d'une équation linéaire $y'(x) = f(x)(Ay(x) + B)$, dont nous savons, par le théorème 3.3, que les solutions existent globalement (c'est-à-dire sur tout intervalle I sur lequel la fonction f est continue). Dans ce cas on a $g(y) = Ay + B$ et l'intégrale de $1/g$ diverge logarithmiquement en $\pm\infty$ et en $-B/A$. Supposons par exemple que $A > 0$ et $a > -B/A$. Alors, pour $u > -B/A$, on a

$$\phi(u) = \int_a^u \frac{dv}{Av + B} = \frac{1}{A} \log \frac{Au + B}{Aa + B},$$

et donc

$$\lim_{u \downarrow -B/A} \phi(u) = -\infty, \quad \lim_{u \uparrow \infty} \phi(u) = +\infty$$

L'image de la fonction $\phi :] -B/A, \infty[\rightarrow \mathbb{R}$ est \mathbb{R} tout entier, elle est donc bijective et son inverse est définie partout

$$\phi^{-1}(x) = \frac{e^{x(Aa + B)} - B}{A}.$$

Il en résulte que la solution du problème de Cauchy avec condition initiale $y(0) = a$,

$$y(x) = \phi^{-1}(F(x)),$$

existe sur tout intervalle $I \ni 0$ sur lequel la fonction f est continue, et donc la fonction

$$F(x) = \int_0^x f(u) du,$$

est bien définie.

◁

Nous retiendrons donc une première leçon : même si la fonction $f(x, y)$ est très régulière, une solution de l'EDO non linéaire $y'(x) = f(x, y(x))$ n'existe pas nécessairement pour tous les x . Le mieux qu'on puisse attendre est l'existence d'une solution locale, dans un voisinage du point initial x_0 .

Exemple 3.18 Considérons le problème de Cauchy

$$y'(x) = \sqrt{|y(x)|}, \quad y(0) = 0.$$

Si $y(x)$ est une solution de ce problème, alors $y'(x) \geq 0$ et donc $y(x)$ est non-décroissante. Comme $y(0) = 0$, on déduit que $y(x) \geq 0$ et par conséquence que $|y(x)| = y(x)$. Dans ce cas on a donc

$$\phi(u) = \int_0^u \frac{dv}{\sqrt{v}} = 2\sqrt{u}, \quad F(x) = x,$$

et la relation

$$2\sqrt{y(x)} = x,$$

se résout pour obtenir une solution du problème de Cauchy

$$y(x) = \frac{x^2}{4}.$$

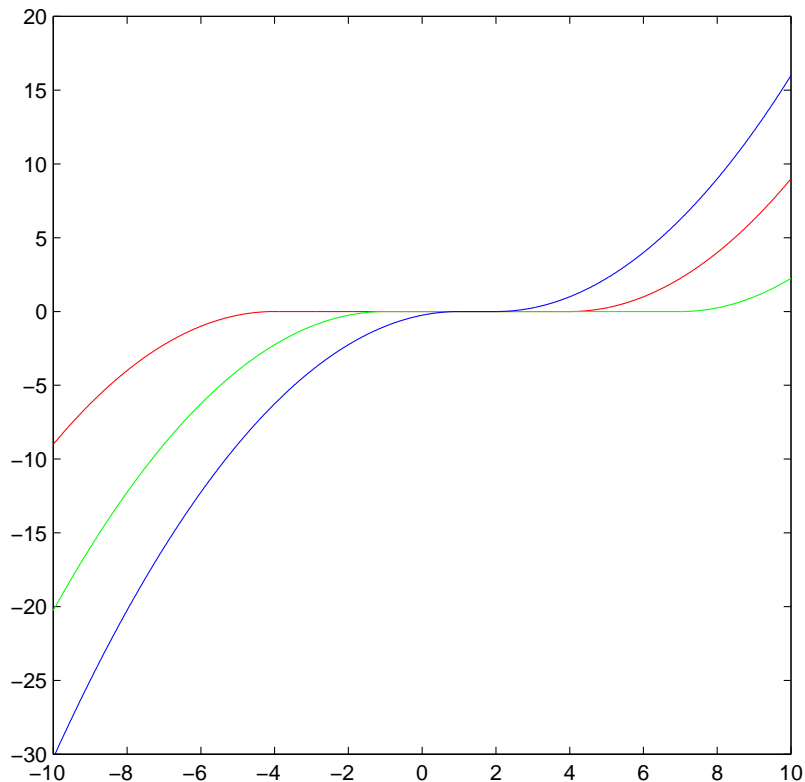
On remarque cependant que la fonction constante $y(x) = 0$ est également une solution de ce problème ! Plus encore : pour tous $a \leq 0 \leq b$, la fonction

$$y(x) = \begin{cases} \frac{-(x-a)^2}{4} & \text{pour } x \in]-\infty, a[, \\ 0 & \text{pour } x \in [a, b], \\ \frac{(x-b)^2}{4} & \text{pour } x \in]b, \infty[, \end{cases}$$

est partout différentiable et sa dérivée est

$$y'(x) = \begin{cases} \frac{a-x}{2} & \text{pour } x \in]-\infty, a[, \\ 0 & \text{pour } x \in [a, b], \\ \frac{x-b}{2} & \text{pour } x \in]b, \infty[. \end{cases} = \sqrt{|y(x)|}.$$

Toutes ces fonctions sont donc des solutions de notre problème. La figure ci-dessous en montre trois.



◁

Nous avons reçu notre deuxième leçon : l'unicité de la solution du problème de Cauchy pour une EDO non linéaire $y'(x) = f(y(x), x)$ n'est pas assurée, même si la fonction f est continue. Nous sommes donc amené à formuler le théorème suivant.

Théorème 3.5 Soit f une fonction continue sur l'intervalle $I =]a, b[$ et g une fonction continue sur l'intervalle $J =]c, d[$.

1. Pour tout $(x_0, y_0) \in I \times J$ il existe un intervalle ouvert $O \subset I$ contenant x_0 et une solution $y : O \rightarrow J$ du problème de Cauchy

$$y'(x) = f(x)g(y(x)), \quad y(x_0) = y_0.$$

2. Si g est de classe C^1 sur J , cette solution est unique

Démonstration. Nous démontrerons l'unicité de la solution dans le paragraphe 3.2.2. Nous pouvons démontrer l'existence en analysant soigneusement la construction de la solution que nous avons décrite au début de ce paragraphe.

Considérons tout d'abord le cas particulier où la condition initiale y_0 est telle que $g(y_0) = 0$. Il est alors clair que la fonction constante $y(x) = y_0$ pour tout $x \in I$ est une solution du problème de Cauchy. On dit dans ce cas que y_0 est un **point fixe** de l'EDO $y'(x) = f(x)g(y(x))$.

Nous pouvons donc supposer que $g(y_0) \neq 0$. Comme g est continue, il existe un intervalle ouvert $U \subset J$ contenant y_0 et tel que $g(y) \neq 0$ pour tout $y \in U$. Par conséquent la fonction

$$\phi(u) = \int_{y_0}^u \frac{dv}{g(v)},$$

est de classe C^1 sur U , où sa dérivée $\phi'(u) = 1/g(u)$ a un signe constant. On en déduit que ϕ est strictement monotone et donc injective sur U . Il en résulte que la fonction réciproque ϕ^{-1} existe et est de classe C^1 sur l'image $V = \phi(U)$, un intervalle ouvert contenant $0 = \phi(y_0)$. Sa dérivée est donnée par

$$\frac{d}{dw} \phi^{-1}(w) = \frac{1}{\phi'(\phi^{-1}(w))} = g(\phi^{-1}(w)).$$

La fonction f étant continue sur I , sa primitive

$$F(x) = \int_{x_0}^x f(u) du,$$

y est de classe C^1 . Comme $F(x_0) = 0$, l'image réciproque $F^{-1}(V)$ de l'intervalle ouvert V contenant 0 est un ouvert contenant x_0 . En particulier il existe un intervalle ouvert $O =]x_0 - \varepsilon, x_0 + \varepsilon[$ tel que $F(O) \subset V$. On en conclut que la fonction

$$y(x) = \phi^{-1} \circ F(x),$$

est de classe C^1 sur O et que

$$y'(x) = g(\phi^{-1}(F(x)))F'(x) = g(y(x))f(x),$$

pour tout $x \in O$. □

Remarque 3.6 D'un point de vue pratique, on peut résumer la méthode de résolution du problème de Cauchy (3.35) (3.36) en écrivant

$$\frac{dy}{dx} = f(y)g(x),$$

sous la forme

$$\frac{dy}{f(y)} = g(x)dx,$$

puis en intégrant cette relation

$$\int_{y_0}^{y(x)} \frac{dy}{f(y)} = \int_{x_0}^x g(x)dx.$$

Exemple 3.19 Si on admet que chaque individu d'une colonie de bactéries scissipares se reproduit par division cellulaire une fois toutes les τ secondes, la population $n(t)$ de cette colonie satisfait $n(t + \tau) = 2n(t)$. En écrivant cette relation sous la forme

$$\frac{n(t + \tau) - n(t)}{\tau} = \frac{n(t)}{\tau},$$

on est amené à la remplacer, pour des temps $t \gg \tau$, par l'EDO linéaire

$$\frac{dn}{dt}(t) = \frac{n(t)}{\tau}.$$

Si n_0 désigne la population au temps $t = 0$, la solution du problème de Cauchy correspondant donne la population au temps t ,

$$n(t) = e^{t/\tau} n_0.$$

Cette loi de croissance a été obtenue en négligeant un premier facteur important : le fait qu'une bactérie n'est pas éternelle. En supposant que son temps de vie moyen est de τ^\dagger secondes, on enregistrera en moyenne $n(t)/\tau^\dagger$ décès par seconde dans la colonie. L'équation gouvernant sa croissance devient alors

$$\frac{dn}{dt}(t) = \left(\frac{1}{\tau} - \frac{1}{\tau^\dagger} \right) n(t) = Rn(t),$$

où nous avons introduit le taux de croissance R . Comme habituellement $\tau^\dagger \gg \tau$, on peut supposer que $R > 0$. La prise en compte du temps de vie fini d'un individu ne change donc pas la nature de l'évolution de la colonie, mais modifie simplement son taux de croissance.

Pour obtenir un modèle plus réaliste, nous devons considérer un second facteur déterminant : la capacité d'une bactérie à se reproduire dépend de la disponibilité des ressources nécessaires à la construction son matériel cellulaire. Il est clair que plus la population est grande, plus faibles sont les ressources moyennes accessibles à chaque individu. Le temps de reproduction τ est donc une fonction croissante du nombre d'individus. Autrement dit $1/\tau$ décroît avec n . En première approximation on peut donc postuler que

$$\frac{1}{\tau} = \frac{1}{\tau_0} - kn,$$

où τ_0, k sont des constantes positives. En posant

$$R_0 = \frac{1}{\tau_0} - \frac{1}{\tau^\dagger},$$

on arrive ainsi à l'EDO

$$\frac{dn}{dt}(t) = (R_0 - kn(t))n(t) = g(n(t)), \quad (3.40)$$

La fonction $g(n) = (R_0 - kn)n$ a deux zéros, $n = 0$ et $n = n_{\max} = R_0/k$. L'équation (3.40) a donc deux point fixes. Pour $n_0 \in]0, n_{\max}[$, on a

$$\phi(u) = \int_{n_0}^u \frac{dv}{(R_0 - kv)v} = \frac{1}{R_0} \int_{n_0}^u \left(\frac{1}{v} + \frac{1}{n_{\max} - v} \right) dv = \frac{1}{R_0} \log \left(\frac{u}{n_0} \frac{n_{\max} - n_0}{n_{\max} - u} \right). \quad (3.41)$$

On peut donc écrire la relation $\phi(n(t)) = t$ sous la forme

$$e^{R_0 t} = \frac{n(t)}{n_0} \frac{n_{\max} - n_0}{n_{\max} - n(t)},$$

qui se résout facilement pour obtenir

$$n(t) = \frac{n_0 n_{\max}}{n_0 + (n_{\max} - n_0)e^{-R_0 t}}.$$

On remarque en particulier que

$$\lim_{t \rightarrow \infty} n(t) = n_{\max}, \quad \lim_{t \rightarrow -\infty} n(t) = 0,$$

et qu'on a donc une solution globale définie sur \mathbb{R} tout entier. Dans le contexte de la dynamique de population qui nous a mené à l'EDO (3.40), les conditions initiales $n(0) = n_0 < 0$ ou $n(0) = n_0 > n_{\max}$ n'ont pas de sens ($\tau < 0$ dans ce dernier cas!). Néanmoins l'examen de ces cas est intéressant du point de vue mathématique. Si $n_0 < 0$, on trouve que pour tout $u < 0$,

$$\phi(u) = \int_{n_0}^u \frac{dv}{(R_0 - kv)v},$$

est toujours donné par l'expression (3.41). On peut écrire la solution sous la forme

$$n(t) = -n_{\max} \left[\left(1 + \frac{n_{\max}}{|n_0|} \right) e^{-R_0 t} - 1 \right]^{-1},$$

et on remarque que

$$\lim_{t \uparrow t_{\max}} n(t) = -\infty, \quad t_{\max} = \frac{1}{R_0} \log \left(1 + \frac{n_{\max}}{|n_0|} \right),$$

alors que

$$\lim_{t \rightarrow -\infty} n(t) = 0.$$

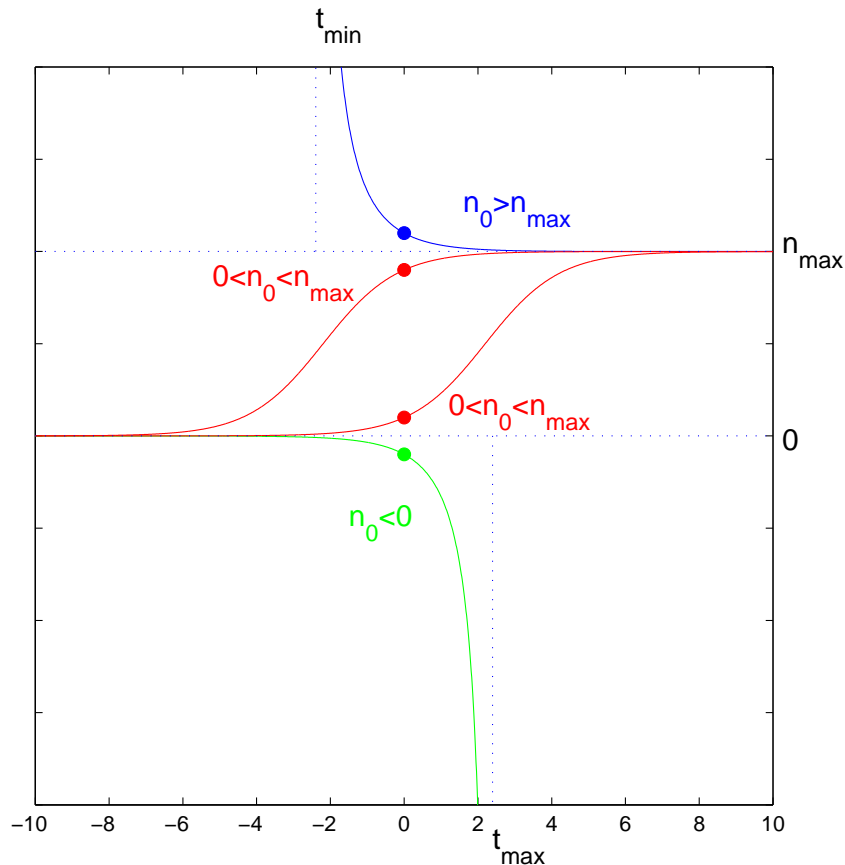
La solution n'existe pas globalement dans ce cas et son intervalle d'existence maximal est donné par $] - \infty, t_{\max}[$. Pour $n_0 > n_{\max}$, on trouve de manière analogue que

$$\lim_{t \rightarrow \infty} n(t) = n_{\max},$$

alors que

$$\lim_{t \downarrow t_{\min}} n(t) = +\infty, \quad t_{\min} = \frac{1}{R_0} \log \left(1 - \frac{n_{\max}}{n_0} \right).$$

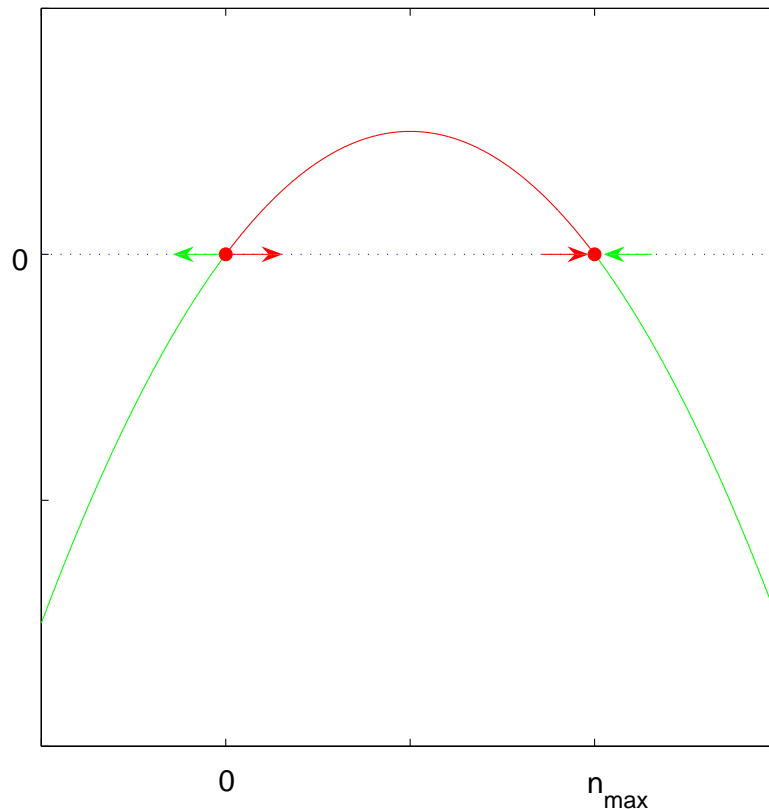
On a donc un intervalle maximal $]t_{\min}, \infty[$. Les trois types de solutions de l'EDO (3.40) sont représenté sur la figure suivante.



On peut faire deux observations importantes sur cette figure.

1. Les points fixes 0 et n_{\max} séparent les différents types de solutions. C'est un principe général pour les EDO autonomes $y'(x) = g(y(x))$. Lorsqu'on étudie une telle équation, on commence toujours par chercher ses points fixes, c'est-à-dire les zéros de la fonction $g(y)$.
2. La nature des deux point fixes 0 et n_{\max} est différente. Toutes les solutions pour des conditions initiales n_0 proches de n_{\max} convergent vers n_{\max} lorsque $t \rightarrow \infty$. On dit d'un tel point fixe qu'il est **stable**. On remarque au contraire que toutes les solutions avec des conditions initiales $n_0 \neq 0$ mais proches de 0 s'écartent de ce point fixe. On dit qu'il est **instable**. On note aussi que cette situation est inversée lorsque $t \rightarrow -\infty$.

On peut comprendre ce comportement en remarquant que les points fixes 0 et n_{\max} sont des zéros simples de la fonction g , c'est-à-dire que g change de signe en ces points (voir la figure suivante). Sur l'intervalle $] -\infty, 0[$ la fonction g est négative. Si $n(t)$ est une solution de l'EDO $n'(t) = g(n(t))$ avec condition initiale $n(0) < 0$, alors $n'(0) = g(n(0)) < 0$. La fonction $n(t)$ est donc décroissante en $t = 0$, ce qui implique que $n(t) < n(0) < 0$ pour des petits $t > 0$. La solution s'éloigne du point fixe 0 en restant dans l'intervalle $] -\infty, 0[$. On en déduit que $n'(t) < 0$ pour tout les $t > 0$ pour lesquels la solution existe. De l'autre côté du point fixe 0 , sur l'intervalle $]0, n_{\max}[$, la fonction g est positive. Les solutions avec condition initiale $n(0) > 0$ sont donc croissantes en $t = 0$ et on a $n(t) > n(0) > 0$. Ici aussi la solution s'éloigne du point fixe. La situation est similaire près du point fixe n_{\max} , mais cette fois les solutions sont croissantes à gauche du point fixe et décroissantes à droite. Elles s'approchent donc du point fixe.



◁

Nous pouvons déduire de cette discussion un principe général :

Théorème 3.7 *Un point fixe y_0 de l'EDO autonome $y'(x) = g(y(x))$ qui est un zéro simple de g est instable si g est négative à droite de y_0 et positive à gauche, c'est-à-dire si $g'(y_0) > 0$. Il est stable si au contraire $g'(y_0) < 0$.*

3.2.2 Deux théorèmes généraux

Dans ce paragraphe nous énonçons deux résultats plus généraux que le théorème 3.5. Le premier, que nous ne démontrerons pas, assure l'existence d'une solution maximale du problème de Cauchy pour une EDO non linéaire du premier ordre sous une simple hypothèse de continuité (voir le chapitre 2 de [H] pour une démonstration. Un résultat similaire existe pour des équations d'ordre supérieur). Le second montre que sous une condition de régularité supplémentaire, cette solution est unique. Sa démonstration est très simple et se base sur une inégalité différentielle que nous analyserons en détail.

Théorème 3.8 *Soit f une fonction continue sur le rectangle ouvert $R =]a, b[\times]c, d[$ où $-\infty \leq a < b \leq \infty$ et $-\infty \leq c < d \leq \infty$. Pour tout $(x_0, y_0) \in R$ il existe une solution $y(x)$ du problème de Cauchy*

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0, \tag{3.42}$$

définie sur un intervalle $J =]a', b'[$ tel que $(x, y(x))$ s'approche du bord de R lorsque x s'approche d'un bord de J , c'est-à-dire :

1. Soit $a' = a$, soit $\lim_{x \downarrow a'} y(x) = c$, soit $\lim_{x \downarrow a'} y(x) = d$.
2. Soit $b' = b$, soit $\lim_{x \uparrow b'} y(x) = c$, soit $\lim_{x \uparrow b'} y(x) = d$.

Comme nous l'avons appris dans l'exemple 3.18, l'hypothèse de continuité du théorème précédent ne suffit pas pour assurer l'unicité de la solution $y(x)$.

Théorème 3.9 Si, en plus des hypothèses du théorème 3.8, la fonction f admet une dérivée partielle par rapport à y continue sur le rectangle R , alors la solution maximale du problème de Cauchy (3.42) est unique.

Remarque 3.10 L'unicité de la solution d'une équation séparable formulée dans le théorème 3.5 est une conséquence de ce résultat.

Pour démontrer le théorème 3.9 nous utiliserons le résultat suivant qui est à un intérêt intrinsèque.

Théorème 3.11 (Lemme de Gronwall) Si la fonction $f(x)$ de classe C^1 sur l'intervalle $]a, b[$ et continue sur l'intervalle $[a, b]$ satisfait l'inégalité différentielle

$$|f'(x)| \leq M|f(x)|, \quad (3.43)$$

pour tout $x \in]a, b[$, alors elle satisfait l'inégalité

$$|f(x)| \leq |f(a)|e^{M(x-a)},$$

pour tout $x \in [a, b]$.

Démonstration. Pour $x_0 \in]a, b[$ fixé, on a

$$f(x) = f(x_0) + \int_{x_0}^x f'(u) du,$$

pour tout $x \in]a, b[$. Pour $x \in [x_0, b[$ on en tire

$$|f(x)| \leq |f(x_0)| + \int_{x_0}^x |f'(u)| du,$$

et l'inégalité (3.43) nous donne

$$|f(x)| \leq |f(x_0)| + M \int_{x_0}^x |f(u)| du.$$

En introduisant la fonction

$$g_\varepsilon(x) = |f(x_0)| + M \int_{x_0}^x |f(u)| du + \varepsilon,$$

on peut donc écrire

$$|f(x)| \leq g_\varepsilon(x), \quad (3.44)$$

pour tout $\varepsilon > 0$. On en déduit que

$$g'_\varepsilon(x) = M|f(x)| \leq M g_\varepsilon(x).$$

Comme $g_\varepsilon(x) \geq \varepsilon > 0$, on peut écrire cette dernière inégalité comme

$$\frac{g'_\varepsilon(x)}{g_\varepsilon(x)} = \frac{d}{dx} \log g_\varepsilon(x) \leq M.$$

On obtient ainsi

$$\log g_\varepsilon(x) - \log g_\varepsilon(x_0) = \int_{x_0}^x \frac{d}{du} \log g_\varepsilon(u) du \leq \int_{x_0}^x M du = M(x - x_0).$$

En prenant l'exponentielle des deux côtés de cette inégalité on arrive à

$$g_\varepsilon(x) \leq g_\varepsilon(x_0)e^{M(x-x_0)} = (|f(x_0)| + \varepsilon)e^{M(x-x_0)},$$

et l'inégalité (3.44) nous permet d'écrire

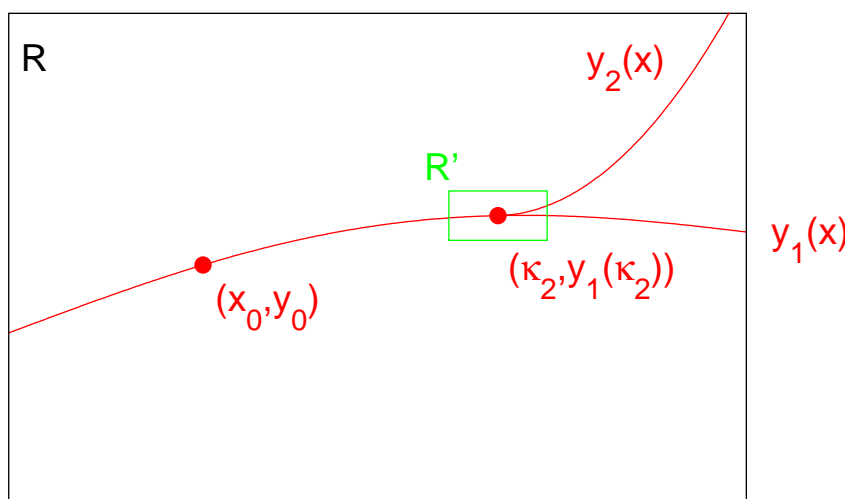
$$|f(x)| \leq (|f(x_0)| + \varepsilon)e^{M(x-x_0)}.$$

Comme cette inégalité est vraie pour tout $\varepsilon > 0$, on peut y prendre la limite $\varepsilon \downarrow 0$ pour obtenir

$$|f(x)| \leq |f(x_0)|e^{M(x-x_0)},$$

pour tout $x \in [x_0, b[$. Les deux membres de cette inégalité étant continus en x_0 , on peut y prendre la limite $x_0 \downarrow a$. Le même argument de continuité montre que cette inégalité reste vérifiée en $x = b$. \square

Preuve du théorème 3.9. Soient $y_1(x)$ et $y_2(x)$ deux solutions maximales du problème de Cauchy (3.42) définies sur des intervalles ouverts J_1 et J_2 . Comme $x_0 \in J = J_1 \cap J_2$, J est un intervalle ouvert non-vide. Comme $y_1(x_0) = y_2(x_0)$, l'ensemble $K = \{x \in J \mid y_1(x) = y_2(x)\}$ n'est pas vide et donc $\kappa_1 = \inf K \in [a, x_0]$, $\kappa_2 = \sup K \in [x_0, b]$ et $y_1(x) = y_2(x)$ pour tout $x \in]\kappa_1, \kappa_2[$. Si $(x, y_1(x))$ s'approche du bord de R lorsque x s'approche de κ_1 et de κ_2 , alors $J_1 = J_2 = K$ et les solutions y_1 et y_2 sont identiques. Supposons donc que $(x, y_1(x))$ ne s'approche pas du bord de R lorsque $x \uparrow \kappa_2$ (le cas $x \downarrow \kappa_1$ se traite de façon similaire). Il existe alors un rectangle fermé $R' = [\kappa_2 - \varepsilon, \kappa_2 + \varepsilon] \times [\alpha, \beta] \subset R$ tel que les solutions y_1 et y_2 soient définies sur $U =]\kappa_2 - \varepsilon, \kappa_2 + \varepsilon[$ et $y_1(x), y_2(x) \in [\alpha, \beta]$ pour tout $x \in U$. De plus $y_1(\kappa_2) = \lim_{x \uparrow \kappa_2} y_1(x) = \lim_{x \uparrow \kappa_2} y_2(x) = y_2(\kappa_2)$ alors que pour tout $\varepsilon > 0$ est assez petit $y_1(\kappa_2 + \varepsilon) \neq y_2(\kappa_2 + \varepsilon)$ (voir la figure suivante).



On en déduit qu'il existe un intervalle ouvert $V =]\kappa_2, \kappa_2 + c[\subset U$ tel que $y_1(x) \neq y_2(x)$ pour tout $x \in V$. Posons $z(x) = y_2(x) - y_1(x)$. Pour tout $x \in V$ on a

$$z'(x) = y_2'(x) - y_1'(x) = f(x, y_2(x)) - f(x, y_1(x)),$$

et le théorème des accroissements finis nous permet d'écrire

$$z'(x) = \frac{\partial f}{\partial y}(x, \theta(x))z(x), \quad z(\kappa_2) = 0,$$

où $\theta(x)$ est dans l'intervalle entre $y_1(x)$ et $y_2(x)$. Comme $y_1(x)$ et $y_2(x)$ sont tous deux dans l'intervalle $[\alpha, \beta]$ on a

$$|z'(x)| \leq M' |z(x)|,$$

pour tout $x \in V$, où

$$M' = \max_{(x,y) \in R'} \left| \frac{\partial f}{\partial y}(x, y) \right| < \infty,$$

puisque la dérivée partielle de f est continue. Le théorème 3.11 nous amène donc à la contradiction

$$|z(x)| \leq |z(\kappa_2)| e^{M'(x-\kappa_2)} = 0,$$

pour tout $x \in V$. \square

3.2.3 Analyse qualitative

Bien qu'en général il ne soit pas possible d'exprimer la solution $y(x)$ du problème de Cauchy non linéaire

$$y'(x) = f(x, y(x)), \quad y(x_0) = y_0, \quad (3.45)$$

à l'aide de fonctions élémentaires, de fonctions spéciales ou d'intégrales, il est souvent possible d'obtenir des informations qualitatives sur cette solution.

Nous avons remarqué, lors de notre discussion de l'exemple 3.19, qu'une connaissance des points fixes d'une EDO autonome permettrait de comprendre qualitativement le comportement global des solutions de cette équation. Cette technique ne s'applique pas aux EDO non-autonomes puisque dans ce cas l'équation $f(x, y) = 0$ définit implicitement une fonction $y = \xi(x)$ dont la dérivée est donnée par le théorème de la fonction implicite 1.48

$$\xi'(x) = -\frac{f_x(x, \xi(x))}{f_y(x, \xi(x))}.$$

Comme en général $f_x(x, \xi(x)) \neq 0$, la fonction $\xi(x)$ ne satisfait pas l'EDO $\xi'(x) = f(x, \xi(x)) = 0$.

Si une EDO non autonome n'a généralement pas de point fixe, elle peut néanmoins admettre des points ou plus généralement des ensembles absorbants vers lesquels ses solutions sont attirées ou qui, au contraire, les repoussent. Nous allons illustrer à l'aide d'un exemple comment utiliser de telles informations pour comprendre de façon qualitative le comportement des solutions.

Exemple 3.20 La fonction $f(x, y) = x - y^3$ est continue dans \mathbb{R}^2 tout entier, de même que sa dérivée partielle

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

Pour tout $(x_0, y_0) \in \mathbb{R}^2$, le problème de Cauchy

$$y'(x) = x - y(x)^3, \quad y(x_0) = y_0,$$

admet donc une unique solution maximale, c'est-à-dire que par chaque point $(x_0, y_0) \in \mathbb{R}^2$ passe le graphe d'une et d'une seule solution de l'EDO $y'(x) = x - y(x)^3$. A priori, cette solution peut être de l'un des quatre types suivants :

- (i) définie sur \mathbb{R} tout entier (sans asymptote verticale);
- (ii) définie sur un intervalle $]a, \infty[$ avec $\lim_{x \downarrow a} y(x) = \pm\infty$ (une asymptote verticale à gauche);
- (iii) définie sur un intervalle $] -\infty, b[$ avec $\lim_{x \uparrow b} y(x) = \pm\infty$ (une asymptote verticale à droite);
- (iv) définie sur un intervalle $]a, b[$ avec $\lim_{x \downarrow a} y(x) = \pm\infty$ et $\lim_{x \uparrow b} y(x) = \pm\infty$ (deux asymptotes verticales, à gauche et à droite).

Afin de ne pas alourdir inutilement notre discussion, nous utiliserons le terme "solution" pour désigner soit une fonction $J \ni x \mapsto y(x)$, solution maximale de l'EDO $y'(x) = x - y(x)^3$, soit le graphe $\{(x, y(x)) \in \mathbb{R}^2 \mid x \in J\}$ de cette solution, soit la courbe plane $J \ni x \mapsto (x, y(x))$. Lorsque nous dirons que la solution y passe par le point (x_0, y_0) il faudra donc comprendre que ce point est sur le graphe de la solution y , c'est-à-dire que $x_0 \in J$ et $y_0 = y(x_0)$, soit encore que la courbe $x \mapsto (x, y(x))$ passe par le point (x_0, y_0) .

Remarquons tout d'abord qu'une solution est automatiquement de classe C^∞ . En effet elle est différentiable (par définition d'une solution d'une EDO!) et par conséquence continue. Comme $y'(x) = x - y(x)^2$, on en déduit que $y'(x)$ est différentiable et que

$$y''(x) = \frac{\partial}{\partial x} f(x, y(x)) = 1 - 3y(x)^2 y'(x) = 1 - 3y(x)^2 (x - y(x)^3) = g(x, y(x)). \quad (3.46)$$

On peut clairement répéter cet argument et montrer que toutes les dérivées de y peuvent s'exprimer comme des polynômes dans les deux variables x et $y(x)$.

Notons ensuite que le plan \mathbb{R}^2 est divisé en deux régions par la courbe Ψ_+ d'équation $x = y^3$, lieu des zéros de la fonction $f(x, y)$. Nous appellerons P_+ la portion du plan se situant au dessus de cette courbe, où la fonction f est négative,

$$\Psi_+ = \{(x, y) \mid f(x, y) = 0\}, \quad P_+ = \{(x, y) \mid f(x, y) < 0\}.$$

Une deuxième courbe importante est le lieu des zéros de la fonction $g(x, y)$. Son équation, qu'on peut déduire de (3.46), est

$$x = y^3 + \frac{1}{3y^2}.$$

Elle est formée de deux nappes Ψ_- et Φ , séparées par l'asymptote horizontale $y = 0$ et située toutes deux en dessous de la courbe Ψ_+ puisque

$$y^3 + \frac{1}{3y^2} > y^3,$$

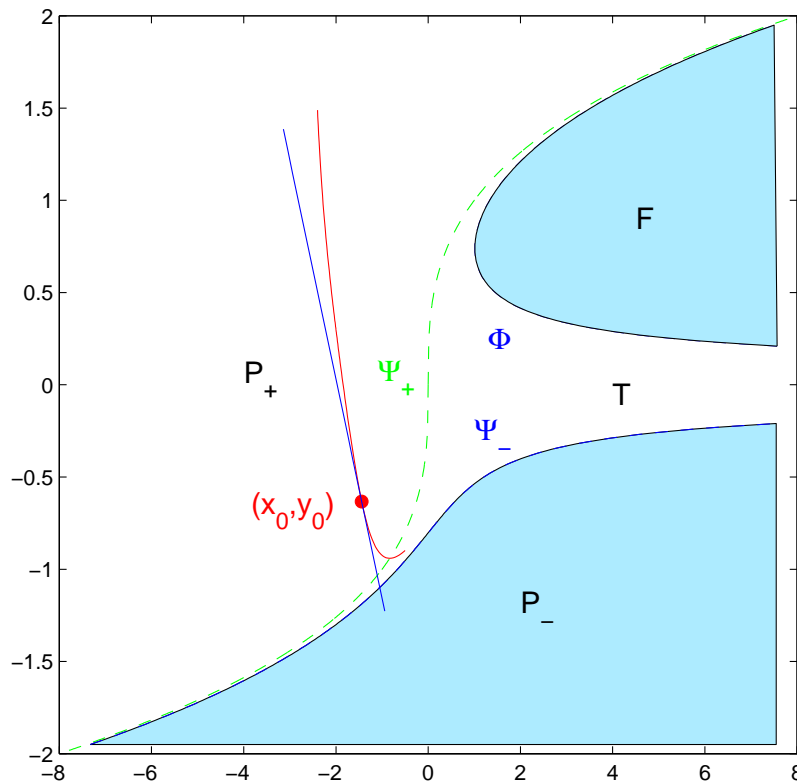
(voir la figure ci-dessous). Nous appelons P_- la région située en dessous de Ψ_- , F celle limitée par Φ et T celle délimitée par les trois courbes Φ_- , Φ_+ et Ψ_+ ,

$$\Psi_- = \{(x, y) \mid g(x, y) = 0, y < 0\}, \quad P_- = \{(x, y) \mid g(x, y) < 0, y < 0\},$$

$$\Phi = \{(x, y) \mid g(x, y) = 0, y > 0\}, \quad F = \{(x, y) \mid g(x, y) < 0, y > 0\},$$

$$T = \{(x, y) \mid f(x, y) > 0, g(x, y) > 0\}.$$

On remarque en particulier que $g(x, y) > 0$ dans P_+ et sur sa frontière Ψ_+ .



Soit y une solution passant par le point $(x_0, y_0) \in P_+$ (voir la figure ci-dessus). Sa tangente en ce point a une pente négative car $f(x_0, y_0) < 0$ dans P_+ . La partie de la courbe y contenue dans P_+ est convexe car dans cette région $g(x, y) > 0$. La courbe y doit donc se situer en dessous de cette tangente dans la région P_+ . On déduit deux choses : premièrement que $y(x) \rightarrow \infty$ lorsque x diminue, deuxièmement que y doit sortir de la région P_+ et par conséquent couper la courbe Ψ_+ en un point $(x_1, y(x_1))$ pour pénétrer dans la région T . En ce point $y'(x_1) = f(x_1, y(x_1)) = 0$ et $y''(x_1) = g(x_1, y(x_1)) > 0$, x_1 donc un minimum local de la fonction y . Pour préciser notre première conclusion, supposons que la solution y n'admette pas d'asymptote verticale à gauche, c'est-à-dire que $y(x)$ soit défini pour tout $x < x_0$. On en déduit que si $a < x_0$ est assez négatif,

$$y(x) > 0 \text{ et } y'(x) = x - y(x)^3 < -y(x)^3,$$

pour tout $x < a$. On peut écrire cette inégalité comme

$$\frac{d}{dx} \frac{1}{y(x)^2} = -2 \frac{y'(x)}{y(x)^3} > 1,$$

et l'intégrer pour obtenir

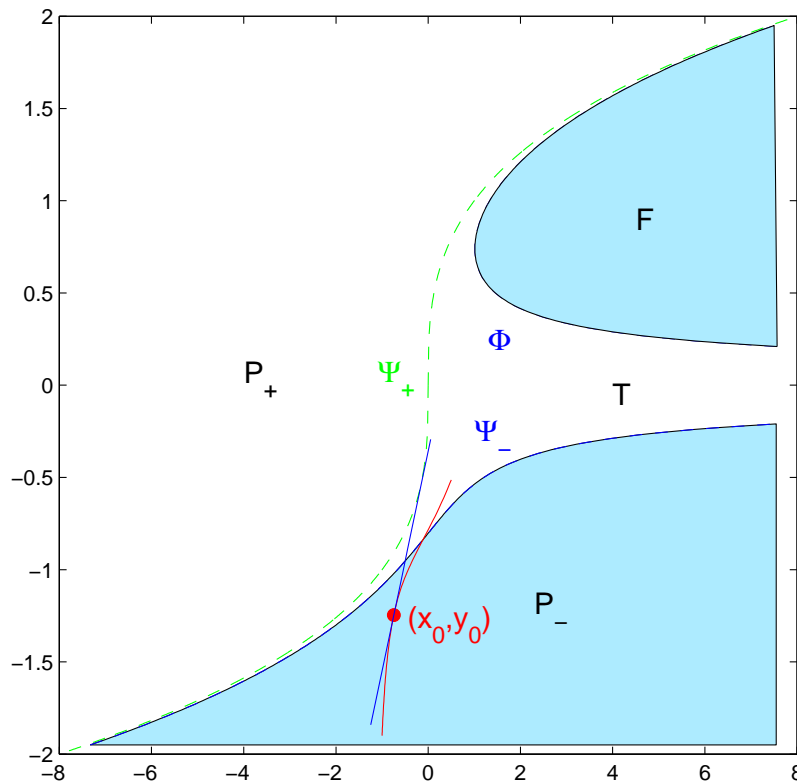
$$\frac{1}{y(a)^2} - \frac{1}{y(x)^2} = \int_x^a \frac{d}{du} \frac{1}{y(u)^2} du > \int_x^a du = a - x,$$

pour tout $x < a$. On arrive ainsi à

$$y(x) > \frac{y(a)^2}{1 - (a-x)y(a)^2},$$

dont on déduit que $y(x) \rightarrow \infty$ lorsque $x \rightarrow a - y(a)^{-2} < a < x_0$, une contradiction avec notre hypothèse selon laquelle $y(x)$ est défini pour tout $x < a$. Nous avons ainsi démontré que toutes les solutions maximales passant par un point de P_+ ont une asymptote verticale à gauche et coupent la courbe Ψ_+ .

En suivant un raisonnement analogue, on montre que toutes les solutions passant par un point de P_- ont une asymptote verticale à gauche et traversent la courbe Ψ_- pour pénétrer la région T . (voir la figure ci-dessous).



Nous allons maintenant montrer que les courbes Ψ_{\pm} et Φ sont semi-perméables. Une solution ne peut traverser Ψ_+ que du haut vers le bas, c'est-à-dire qu'elle peut quitter la région P_+ à travers cette courbe, mais ne peut pas y pénétrer. De même une solution ne peut traverser Ψ_- que du bas vers le haut, elle peut donc sortir de la région P_- à travers cette courbe, mais ne peut pas y entrer. Finalement, une solution peut entrer dans la région F à travers la courbe Φ , mais ne peut pas en sortir. On peut résumer cette situation en disant que les régions P_{\pm} sont répulsives alors que la région F est attractive.

Toutes ces propriétés peuvent se comprendre de la manière suivante. Soit $F(x, y)$ une fonction de classe C^1 et $\Gamma = \{(x, y) \mid F(x, y) = c\}$ une courbe de niveau de F . Cette courbe est la frontière de la région $\Omega = \{(x, y) \mid F(x, y) < c\}$. Supposons qu'une solution y pénètre Ω à travers Γ , c'est-à-dire qu'il existe $(x_0, y_0) = (x_0, y(x_0)) \in \Gamma$ et $\varepsilon > 0$ tels que $F(x, y(x)) > c$ pour $x \in]x_0 - \varepsilon, x_0[$ et $F(x, y(x)) < c$ pour $x \in]x_0, x_0 + \varepsilon[$. On en déduit que la fonction $x \mapsto F(x, y(x))$ est décroissante en x_0 et donc

$$\left. \frac{d}{dx} F(x, y(x)) \right|_{x=x_1} < 0.$$

Comme

$$\frac{d}{dx}F(x, y(x)) = F_x(x, y(x)) + F_y(x, y(x))y'(x) = F_x(x, y(x)) + F_y(x, y(x))f(x, y(x)),$$

une solution ne peut pénétrer Ω en un point $(x_0, y_0) \in \Gamma$ que si

$$F_x(x_0, y_0) + F_y(x_0, y_0)f(x_0, y_0) < 0.$$

De manière totalement analogue, on conclut qu'une solution ne peut sortir de Ω par $(x_0, y_0) \in \Gamma$ que si

$$F_x(x_0, y_0) + F_y(x_0, y_0)f(x_0, y_0) > 0.$$

Appliquons ce principe à la fonction $f(x, y)$ dont la courbe de niveau $\{(x, y) \mid f(x, y) = 0\}$ est précisément Ψ_+ . Une solution ne peut pénétrer la région $P_+ = \{(x, y) \mid f(x, y) < 0\}$ en un point $(x, y) \in \Psi_+$ que si la condition

$$f_x(x, y) + f_y(x, y)f(x, y) < 0,$$

est satisfaite. Comme $f_x(x, y) + f_y(x, y)f(x, y) = g(x, y)$ et comme nous avons déjà remarqué que g est positive sur Ψ_+ , ceci est impossible.

Appliquons le même principe à la fonction $g(x, y)$. Sa courbe de niveau $\{(x, y) \mid g(x, y) = 0\}$ est formée des deux nappes Ψ_- et Φ et $P_- \cup F$ est la région $\{(x, y) \mid g(x, y) < 0\}$. Une solution ne peut pénétrer cette région en $(x, y) \in \Psi_+ \cup \Phi$ que si

$$g_x(x, y) + g_y(x, y)f(x, y) < 0,$$

c'est-à-dire

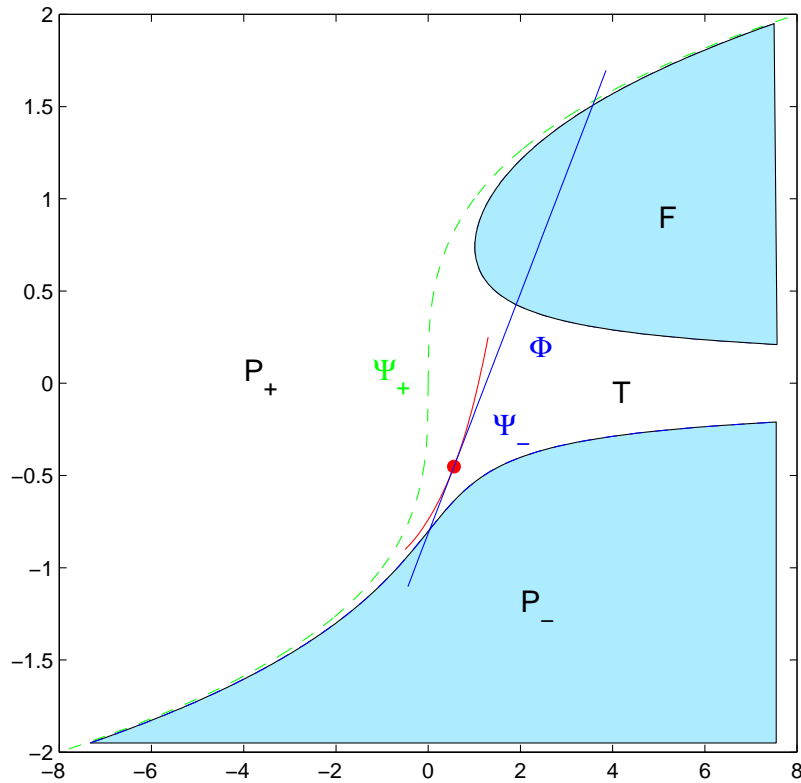
$$-3y^2 + (-6xy + 15y^4)(x - y^3) < 0 \text{ avec } x = y^3 + \frac{1}{3y^2}.$$

Après quelques manipulations élémentaires, cette condition s'écrit

$$-\frac{2}{3y^3} < 0.$$

Elle n'est vérifiée que si $y > 0$. On en conclut qu'une solution ne peut pénétrer la région $P_- \cup F$ que sur la nappe Φ où $y > 0$, c'est-à-dire en passant de T à F . Réciproquement elle ne peut quitter cette région que sur la nappe Ψ_- où $y < 0$, donc en sortant de P_- pour pénétrer T .

Toutes les solutions passant par un point de $P_- \cup P_+$ ont donc une asymptote verticale à gauche et pénètrent dans la région T . D'autre part il est aisé de montrer que toute solution passant par un point de T doit forcément sortir de cette région pour entrer dans F . Supposons en effet que la solution y passant par un point de T ne pénètre jamais F (voir la figure suivante). Comme elle ne peut entrer dans P_- elle reste dans T où g est positive. Son graphe est donc convexe et situé au dessus de sa tangente. On en déduit immédiatement qu'il doit pénétrer la région P_+ , ce qui est impossible.



Comme aucune solution ne peut suivre une des trois courbes Ψ_-, Ψ_+, Φ , toutes les solutions doivent passer par un point de l'ensemble $P_- \cup P_+ \cup T \cup F$. Nous pouvons en conclure que toutes les solutions pénètrent "l'ensemble final" F en passant par "l'ensemble transitoire" T.

Le comportement asymptotique, lorsque $x \rightarrow +\infty$, de toutes les solutions est donc déterminé par le comportement des solutions passant par un point de F. On note en particulier qu'aucune solution n'a d'asymptote à droite. Il est un peu plus difficile de montrer que ces solutions s'approchent de la branche supérieure de la courbe Ψ , plus précisément que

$$y(x) = x^{1/3}(1 - \theta(x)),$$

où la fonction $\theta(x)$ tend vers zéro lorsque $x \rightarrow \infty$,

$$0 < \theta(x) \leq Cx^{-5/3}. \tag{3.47}$$

Une démonstration de cette estimation est faite dans l'appendice de ce paragraphe.

Le comportement asymptotique d'une solution lorsque $x \rightarrow -\infty$ dépend de la région par laquelle elle pénètre dans T. Nous avons déjà discuté deux cas possibles : si y pénètre dans T en provenance de P₊ elle a une asymptote verticale

$$\lim_{x \downarrow x_-} y(x) = +\infty,$$

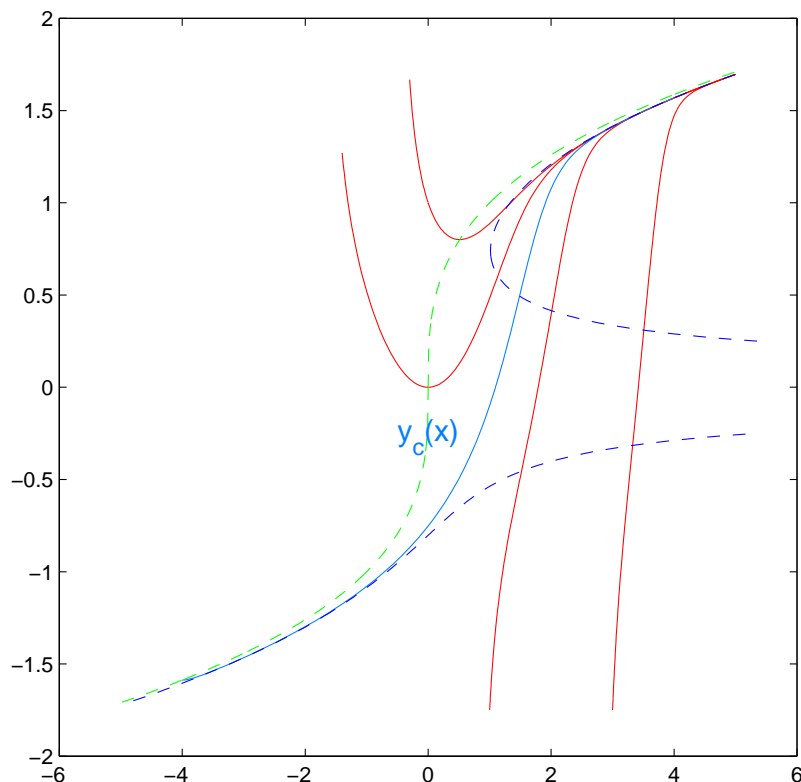
alors que si elle arrive de P₋ elle a une asymptote verticale

$$\lim_{x \downarrow x_-} y(x) = -\infty.$$

La seule possibilité restante est celle d'une solution qui ne sortirait jamais de T lorsque $x \rightarrow -\infty$. Une telle solution reste donc "coincée" entre les courbes Ψ_- et Ψ_+ dans cette limite. On peut montrer qu'il existe une et une seule solution $y_c(x)$ de ce type. Comme deux solutions ne peuvent jamais se couper (en raison de l'unicité de la solution du problème de Cauchy !), y_c sépare les solutions en provenance de P₊ de celles sortant de P₋ : si y_1 passe par un point de P₋ et y_2 par un point de P₊ alors, pour tout x dans l'intervalle d'existence commun de ces trois solutions

$$y_1(x) < y_c(x) < y_2(x).$$

La figure suivante, obtenue par intégration numérique, confirme bien notre discussion qualitative.



◁

Preuve de l'estimation (3.47). Pour déterminer le comportement lorsque $x \rightarrow \infty$ des solutions dans F , on note qu'elles satisfont toutes l'inégalité $0 < y(x) < x^{1/3}$. On peut donc écrire

$$y(x) = (1 - \theta(x))x^{1/3},$$

où la fonction $\theta(x)$ satisfait l'inégalité $0 < \theta(x) < 1$. En dérivant cette identité on obtient une équation différentielle pour la fonction θ ,

$$\theta'(x) = \frac{1}{3}(1 - \theta(x))\frac{1}{x} - (1 - (1 - \theta(x))^3)x^{2/3}.$$

En écrivant

$$1 - (1 - \theta(x))^3 = \theta(x)(3 - 3\theta(x) + \theta(x)^2),$$

et en remarquant que

$$3 - 3\theta + \theta^2 > 1,$$

pour $\theta \in]0, 1[$, on arrive à l'inégalité différentielle

$$\theta'(x) \leq -x^{2/3}\theta(x) + \frac{1}{3x}.$$

Cette dernière peut s'écrire sous la forme

$$\frac{d}{dx}\theta(x)e^{\phi(x)} < \frac{e^{\phi(x)}}{3x},$$

où la fonction ϕ est une primitive de $x^{2/3}$,

$$\phi(x) = \frac{3}{5}x^{5/3}.$$

En intégrant cette inégalité on obtient

$$\theta(x)e^{\phi(x)} - \theta(x_0)e^{\phi(x_0)} \leq \int_{x_0}^x \frac{e^{\phi(u)}}{3u} du,$$

ou encore

$$\theta(x) \leq \theta(x_0)e^{-(\phi(x)-\phi(x_0))} + \int_{x_0}^x \frac{e^{-(\phi(x)-\phi(u))}}{3u} du.$$

Le premier terme du membre de droite de cette inégalité décroît rapidement, plus vite que toutes les puissances négative de x , lorsque $x \rightarrow \infty$. Pour estimer le second terme nous séparons l'intégrale en deux parties

$$\int_{x_0}^x \frac{e^{-(\phi(x)-\phi(u))}}{3u} du = \int_{x_0}^a \frac{e^{-(\phi(x)-\phi(u))}}{3u} du + \int_a^x \frac{e^{-(\phi(x)-\phi(u))}}{3u} du,$$

où a est une variable que nous ajusterons plus tard. La fonction ϕ étant monotone croissante, on peut estimer la première intégrale par

$$\frac{1}{3} \int_{x_0}^a \frac{e^{-(\phi(x)-\phi(u))}}{u} du \leq \frac{1}{3} e^{-(\phi(x)-\phi(a))} \int_{x_0}^a \frac{du}{u} \leq \frac{1}{3} e^{-(\phi(x)-\phi(a))} \log \frac{a}{x_0}.$$

alors que la seconde est bornée par

$$\frac{1}{3} \int_a^x \frac{e^{-(\phi(x)-\phi(u))}}{u} du \leq \frac{1}{3} \int_a^x \frac{du}{u} = \frac{1}{3} \log \frac{x}{a}.$$

Nous avons ainsi obtenu l'estimation

$$\theta(x) \leq e^{-(\phi(x)-\phi(a))} \left(\theta(x_0) + \frac{1}{3} \log \frac{a}{x_0} \right) + \frac{1}{3} \log \frac{x}{a}.$$

En choisissant $a = x(1 - x^{-\alpha})$ et en supposant que $x > 1$ on a

$$\phi(x) - \phi(a) = \frac{3}{5} x^{5/3} \left(1 - (1 - x^{-\alpha})^{5/3} \right),$$

et comme $(1 - x^{-\alpha})^{5/3} \leq 1 - x^{-\alpha}$ pour $x > 1$,

$$\phi(x) - \phi(a) \geq \frac{3}{5} x^{5/3-\alpha}.$$

D'autre part

$$\log \frac{a}{x_0} = \log \frac{x}{x_0} + \log(1 - x^{-\alpha}) \leq \log \frac{x}{x_0},$$

alors que

$$\log \left(\frac{x}{a} \right) = -\log(1 - x^{-\alpha}) = \int_{1-x^{-\alpha}}^1 \frac{du}{u} \leq \frac{1}{1-x^{-\alpha}} \int_{1-x^{-\alpha}}^1 du = \frac{x^{-\alpha}}{1-x^{-\alpha}} \leq C_\alpha x^{-\alpha},$$

pour $x > 2$. On a donc

$$\theta(x) \leq \exp \left(-\frac{3}{5} x^{5/3-\alpha} \right) \left(\theta(x_0) + \frac{1}{3} \log \frac{x}{x_0} \right) + \frac{C_\alpha}{3} x^{-\alpha}.$$

Pour tout $0 < \alpha < 5/3$, le premier terme du membre de droite de cette inégalité décroît plus vite que toutes les puissances négatives de x . On en conclut que pour tout $\varepsilon > 0$ il existe une constante C'_ε telle que

$$\theta(x) \leq C'_\varepsilon x^{-5/3+\varepsilon}.$$

□

3.2.4 Intégrales premières

Dans certains cas il est possible de réduire la solution d'une EDO du premier ordre à la résolution d'une équation non linéaire. Supposons en effet que la fonction $F(x, y)$, de classe C^1 , soit constante sur toutes les solutions de l'équation différentielle

$$y'(x) = f(x, y(x)), \tag{3.48}$$

c'est-à-dire que pour toute solution $y(x)$ de cette équation et pour tout x_0, x dans le domaine d'existence de cette solution $F(x, y(x)) = F(x_0, y(x_0))$. On dit dans ce cas que F est une **intégrale première** de l'équation (3.48). La solution du problème de Cauchy pour cette équation avec condition initiale $y(x_0) = y_0$ peut être obtenue en résolvant l'équation

$$F(x, y(x)) = F(x_0, y_0),$$

pour $y(x)$.

Exemple 3.21 L'équation

$$y'(x) = f(x, y(x)) = -2 \frac{y(x)}{x} + x^2 y(x)^2, \quad (3.49)$$

admet l'intégrale première

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

En effet, si $y(x)$ est une solution de (3.49), alors

$$\frac{d}{dx} F(x, y(x)) = F_x(x, y(x)) + F_y(x, y(x))y'(x) = F_x(x, y(x)) + F_y(x, y(x))f(x, y(x)). \quad (3.50)$$

Comme

$$F_x(x, y) = 1 - \frac{2}{x^3 y},$$

et

$$F_y(x, y)f(x, y) = -\frac{1}{x^2 y^2} \left(-2 \frac{y}{x} + x^2 y^2 \right) = \frac{2}{x^3 y} - 1,$$

on a bien

$$\frac{d}{dx} F(x, y(x)) = 0.$$

Pour toute solution $y(x)$ de (3.49) il existe donc une constante c telle que

$$F(x, y(x)) = c.$$

On en déduit que

$$y(x) = \frac{1}{x^2(c-x)}.$$

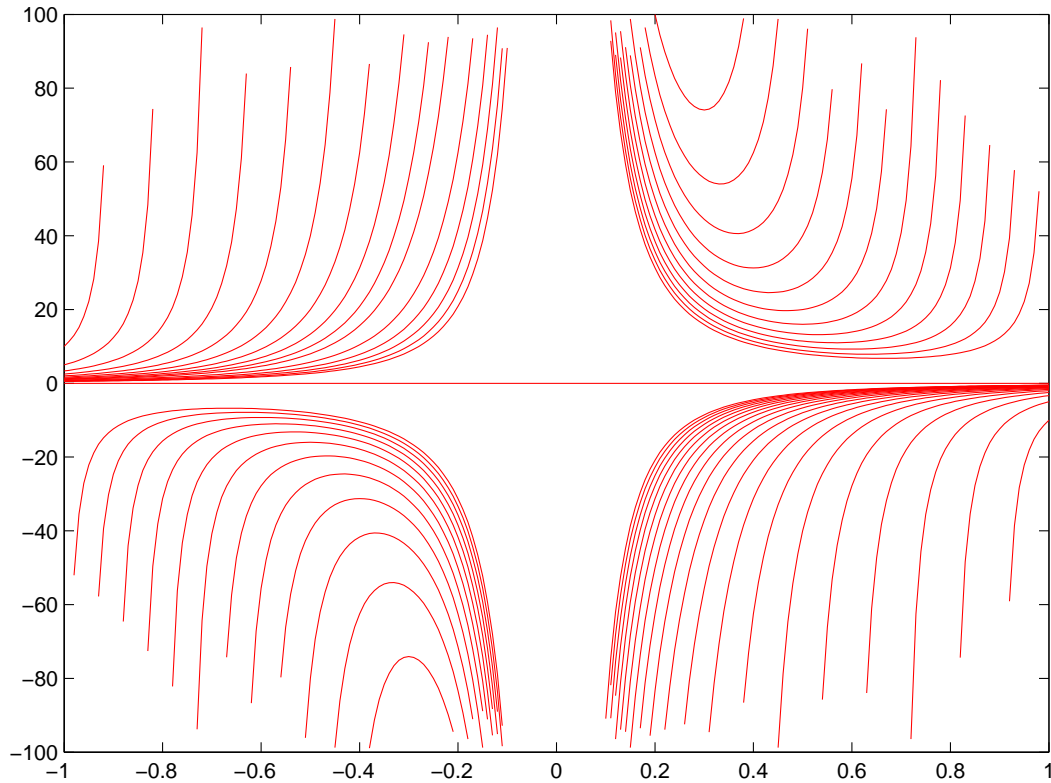
Pour résoudre le problème de Cauchy avec condition initiale $y(x_0) = y_0$ il suffit de déterminer la constante c à l'aide de l'équation

$$y_0 = \frac{1}{x_0^2(c-x_0)},$$

ce qui donne

$$y(x) = \left(\frac{x_0}{x} \right)^2 \frac{y_0}{1 - x_0^2 y_0 (x - x_0)}.$$

Quelques-unes de ces solutions sont représentées sur la figure suivante.



On note que dans le premier et le troisième quadrant les solutions ont deux asymptotes verticales alors que dans le second et le quatrième quadrant elles ont une asymptote verticale et une asymptote horizontale.

Il est important de remarquer aussi que l'intégrale première $F(x, y)$ est singulière en $y = 0$. Elle ne permet donc pas d'obtenir la solution pourtant évidente $y(x) = 0$ de l'équation différentielle (3.49). <

3.2.4.1 EDO exacte

En comparant l'EDO

$$A(x, y(x)) + B(x, y(x))y'(x) = 0, \tag{3.51}$$

à l'identité (3.50), on réalise que si il existe une fonction $F(x, y)$ telle que

$$A(x, y) = F_x(x, y), \quad B(x, y) = F_y(x, y), \tag{3.52}$$

cette fonction est une intégrale première de (3.51). On dit alors que cette dernière est une **EDO exacte**. Le théorème de Clairaut 1.32 nous donne une condition nécessaire sur $A(x, y)$ et $B(x, y)$ pour qu'une telle fonction $F(x, y)$ existe. En effet, si $A(x, y)$ et $B(x, y)$ sont de classe C^1 et si (3.52) est vérifiée, alors $F(x, y)$ est de classe C^2 et

$$A_y(x, y) = F_{xy}(x, y) = F_{yx}(x, y) = B_x(x, y).$$

Réciproquement, si $A(x, y)$ et $B(x, y)$ sont de classe C^1 dans le rectangle $R =]a, b[\times]c, d[$ et y satisfont $A_y(x, y) = B_x(x, y)$, il existe une fonction $F(x, y)$ de classe C^2 dans R satisfaisant (3.52). On peut construire cette fonction de la manière suivante. On commence par résoudre l'équation $F_x = A$ en posant

$$F(x, y) = \int_{x_0}^x A(u, y) du + g(y),$$

où $x_0 \in]a, b[$ et g est une fonction que nous devons déterminer en résolvant l'équation $F_y = B$, c'est-à-dire

$$g'(y) = B(x, y) - \frac{\partial}{\partial y} \int_{x_0}^x A(u, y) du = B(x, y) - \int_{x_0}^x A_y(u, y) du.$$

En remarquant que, contrairement aux apparences, le membre de droite de cette relation ne dépend pas de x

$$\frac{\partial}{\partial x} \left(B(x, y) - \int_{x_0}^x A_y(u, y) du \right) = B_x(x, y) - A_y(x, y) = 0,$$

on peut donc écrire

$$g(y) = \int_{y_0}^y B(x, v) dv - \int_{y_0}^y \left[\int_{x_0}^x A_y(u, v) du \right] dv,$$

où $y_0 \in]c, d[$, et par conséquence

$$F(x, y) = \int_{x_0}^x A(u, y) du + \int_{y_0}^y B(x, v) dv - \int_{y_0}^y \left[\int_{x_0}^x A_y(u, v) du \right] dv.$$

Remarque 3.12 Une équation séparable

$$y'(x) = f(x)g(y(x)),$$

est exacte lorsqu'on l'écrit sous la forme

$$f(x) - \frac{1}{g(y(x))} y'(x) = 0.$$

On vérifie immédiatement qu'une intégrale première est donnée par

$$F(x, y) = \int_a^x f(u) du - \int_b^y \frac{dv}{g(v)},$$

où a, b sont des constantes arbitraires. La solution du problème de Cauchy avec condition initiale $y(x_0) = y_0$ est donc déterminée par l'équation

$$F(x, y(x)) = F(x_0, y_0),$$

qui est identique à (3.37).

Exemple 3.22 L'équation

$$(2x + y(x))y(x) + x(x + 2y(x))y'(x) = 0, \tag{3.53}$$

est exacte, en effet

$$A_y(x, y) = \frac{\partial}{\partial y} (2x + y)y = 2(x + y) = \frac{\partial}{\partial x} x(x + 2y) = B_x(x, y).$$

On cherche un intégrale première $F(x, y)$ en résolvant

$$F_x(x, y) = A(x, y) = 2xy + y^2,$$

ce qui nous donne

$$F(x, y) = x^2y + xy^2 + g(y).$$

La fonction g est déterminée par

$$x^2 + 2xy + g'(y) = F_y(x, y) = B(x, y) = x^2 + 2xy.$$

On a donc $g'(y) = 0$ et une intégrale première est donnée par la fonction

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

On obtient la solution du problème de Cauchy pour l'EDO (3.53) avec la condition initiale $y(x_0) = y_0$ en résolvant l'équation

$$xy^2 + x^2y = c = x_0y_0^2 + x_0^2y_0,$$

d'où

$$y(x) = \frac{-x^2 \pm \sqrt{x^4 + 4cx}}{2x}.$$

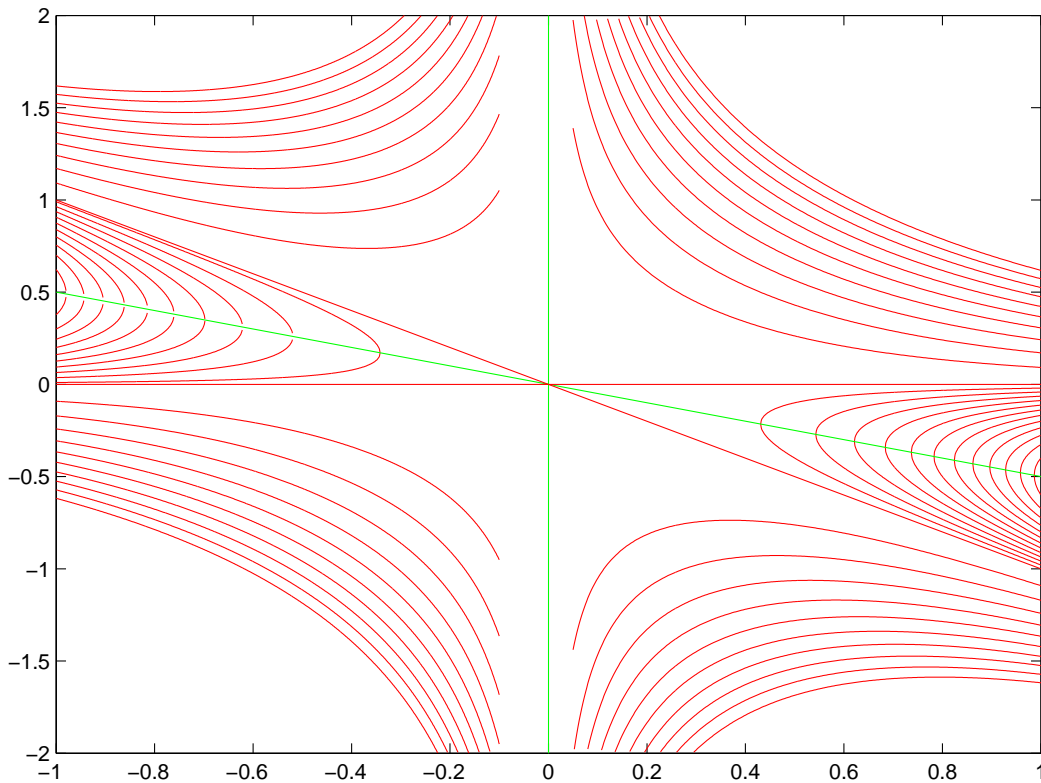
Pour déterminer laquelle de ces deux solutions est la bonne, on remarque qu'on doit avoir

$$y_0 = \frac{-x_0^2 \pm \sqrt{x_0^4 + 4cx_0}}{2x_0} = \frac{-x_0^2 \pm |x_0^2 + 2x_0y_0|}{2x_0}.$$

On a donc

$$y(x) = \begin{cases} \frac{-x^2 - \sqrt{x^4 + 4cx}}{2x} & \text{si } x_0^2 + 2x_0y_0 < 0; \\ \frac{-x^2 + \sqrt{x^4 + 4cx}}{2x} & \text{si } x_0^2 + 2x_0y_0 > 0. \end{cases} \quad (3.54)$$

La figure suivante montre quelques-unes de ces solutions.



On note que le plan est divisé en 8 secteurs que nous numérotions I à VIII à partir du premier quadrant et dans le sens positif. Ils sont délimités par les 4 droites d'équation $y = 0$, $x = 0$, $y = -x$ et $y = -x/2$. Dans le secteur I (le premier quadrant $x > 0$ et $y > 0$), les solutions ont une asymptote verticale $x = 0$ et une asymptote horizontale $y = 0$. Dans le secteur II ($0 < -x < y$), elles ont une asymptote oblique $y = -x$ et une asymptote verticale $x = 0$. Dans le secteur III ($0 < -x/2 < y < -x$), les solutions ont une asymptote oblique $y = -x$ et une singularité en $y = -x/2$ où la dérivée diverge. En effet le radicant dans (3.54) s'annule en $x = a = -(4c)^{1/3}$ et par conséquent

$$\lim_{x \uparrow a} y(x) = -\frac{a}{2}, \quad \lim_{x \uparrow a} y'(x) = -\infty.$$

On retrouve le même phénomène dans le secteur IV ($0 < y < -x/2$) où

$$\lim_{x \uparrow a} y(x) = -\frac{a}{2}, \quad \lim_{x \uparrow a} y'(x) = +\infty.$$

Pour comprendre l'origine de la singularité de la dérivée, il faut écrire l'équation différentielle sous la forme $y'(x) = f(x, y(x))$, avec

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

et remarquer que la fonction f est singulière sur les droites $x + 2y = 0$ et $x = 0$, ce qui se traduit par la divergence de $y'(x)$ lorsque le graphe de la solution $y(x)$ s'approche de ces droites.

La discussion des secteurs V, VI, VII et IIX est similaire. On remarquera finalement que les deux droites $y = 0$ et $y = -x$ sont des solutions ! Ce sont d'ailleurs les seules solutions définies sur \mathbb{R} tout entier. \triangleleft

3.2.4.2 Facteur intégrant

Une fonction $Q(x, y)$ est un **facteur intégrant** pour l'équation différentielle (3.51) si l'équation équivalente

$$Q(x, y(x))A(x, y(x)) + Q(x, y(x))B(x, y(x))y'(x) = 0,$$

est exacte, c'est-à-dire si elle admet une intégrale première $F(x, y)$ telle que

$$F_x(x, y) = Q(x, y)A(x, y), \quad F_y(x, y) = Q(x, y)B(x, y). \quad (3.55)$$

Le facteur intégrant est donc solution de l'EDP

$$Q_y(x, y)A(x, y) + Q(x, y)A_y(x, y) = Q_x(x, y)B(x, y) + Q(x, y)B_x(x, y). \quad (3.56)$$

Cette dernière n'étant généralement pas plus facile à résoudre que l'EDO (3.51), elle est de peu d'utilité (en fait on utilise plutôt cette relation entre EDO et EDP pour réduire la résolution de l'EDP à celle d'une EDO). Il existe cependant des cas particuliers où cette approche fonctionne. Par exemple lorsqu'il est possible de trouver un facteur intégrant ne dépendant que d'une seule variable.

Exemple 3.23 Considérons le problème de Cauchy

$$(x^2 - y(x)^2)y'(x) = 2xy(x), \quad y(x_0) = y_0.$$

On peut écrire cette EDO sous la forme (3.51) avec

$$A(x, y) = 2xy, \quad B(x, y) = y^2 - x^2.$$

On a donc

$$A_y(x, y) = 2x \neq -2x = B_x(x, y),$$

c'est-à-dire que notre EDO n'est pas exacte. Cherchons un facteur intégrant $Q(x, y)$. Ce dernier doit donc satisfaire l'EDP (3.56) qui s'écrit, dans notre cas

$$2xyQ_y(x, y) + 2xQ(x, y) = (y^2 - x^2)Q_x(x, y) - 2xQ(x, y).$$

On remarque que si $Q_x(x, y) = 0$, c'est-à-dire si le facteur intégrant ne dépend pas de x , cette équation devient une EDO pour la fonction $Q(y)$,

$$yQ'(y) = -2Q(y).$$

Cette équation autonome admet la solution $Q(y) = y^{-2}$. On obtient donc une intégrale première $F(x, y)$ en intégrant les relations (3.55) qui deviennent ici

$$F_x(x, y) = \frac{2x}{y}, \quad F_y(x, y) = 1 - \frac{x^2}{y^2}.$$

On déduit de la première relation que

$$F(x, y) = \frac{x^2}{y} + g(y),$$

et la fonction $g(y)$ est déterminée par la seconde relation

$$F_y(x, y) = -\frac{x^2}{y^2} + g'(y) = 1 - \frac{x^2}{y^2}.$$

On en conclut qu'une intégrale première est donnée par

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

La solution du problème de Cauchy s'obtient donc en résolvant l'équation

$$\frac{x^2}{y(x)} + y(x) = \frac{x_0^2}{y_0} + y_0,$$

c'est-à-dire

$$y(x)^2 - \left(\frac{x_0^2 + y_0^2}{y_0} \right) y(x) + x^2 = 0. \quad (3.57)$$

Les deux solutions de cette équation du second degré sont

$$y(x) = \frac{1}{2} \left(\frac{x_0^2 + y_0^2}{y_0} \pm \sqrt{\left(\frac{x_0^2 + y_0^2}{y_0} \right)^2 - 4x^2} \right).$$

On doit choisir celle qui satisfait la condition initiale

$$y(x_0) = \frac{1}{2} \left(\frac{x_0^2 + y_0^2}{y_0} \pm \sqrt{\left(\frac{x_0^2 + y_0^2}{y_0} \right)^2 - 4x_0^2} \right) = \frac{1}{2} \left(\frac{x_0^2 + y_0^2}{y_0} \pm \left| \frac{x_0^2 - y_0^2}{y_0} \right| \right) = y_0$$

et donc le signe + lorsque

$$\frac{x_0^2 - y_0^2}{y_0} < 0,$$

et le signe - lorsque

$$\frac{x_0^2 - y_0^2}{y_0} > 0.$$

Dans le cas limite où

$$\frac{x_0^2 - y_0^2}{y_0} = 0,$$

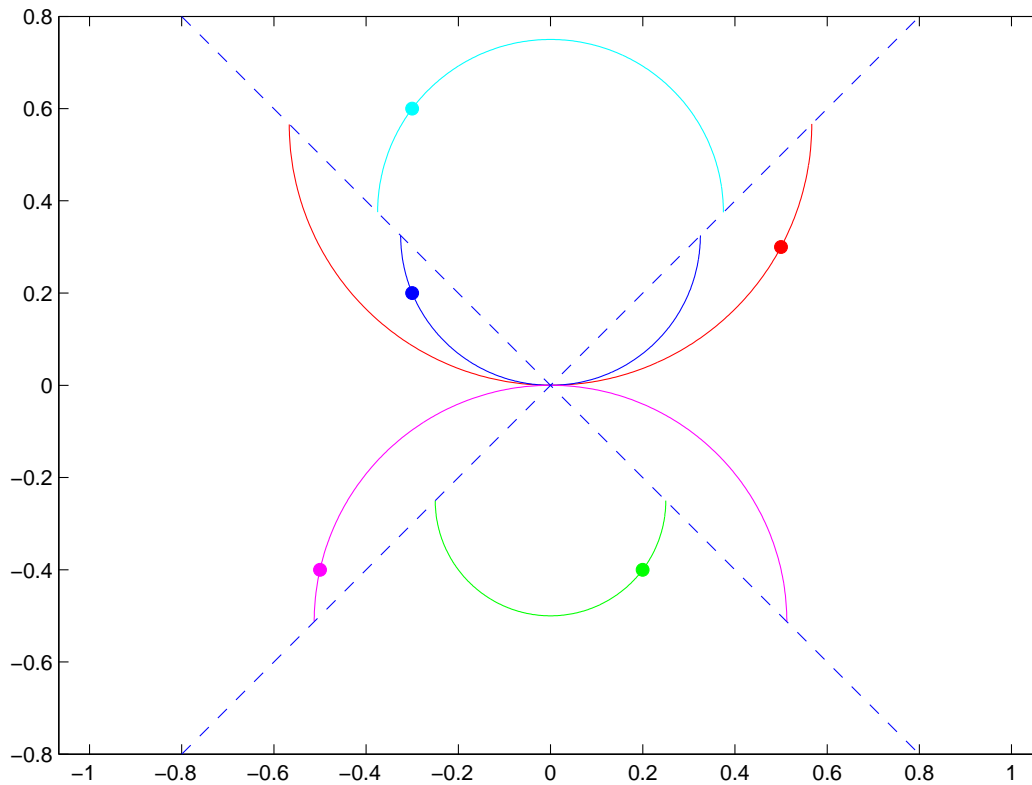
les deux solutions sont acceptables. Dans tous cas cependant, ces solutions ne sont définies que pour $|x| \leq |c|$, où

$$c = \frac{1}{2} \frac{x_0^2 + y_0^2}{y_0}.$$

En effet, en écrivant l'équation (3.57) sous la forme

$$(y - c)^2 + x^2 = c^2,$$

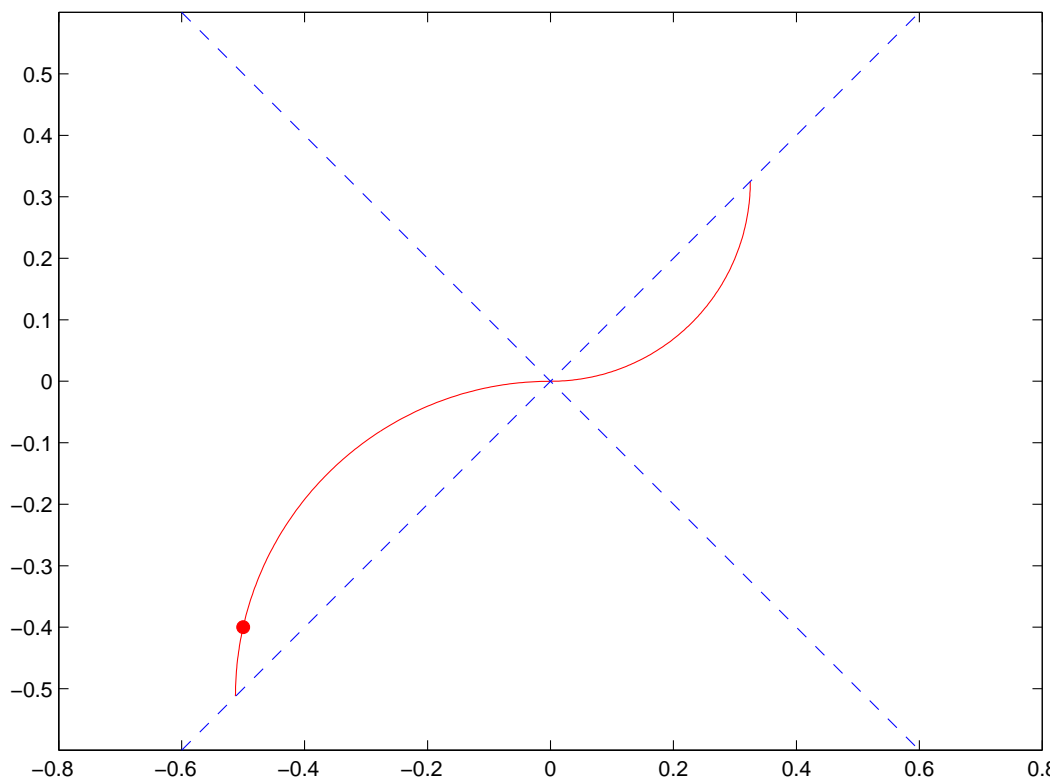
on en déduit que le graphe de la solution $y(x)$ est un demi-cercle de rayon $|c|$ centré en $(0, c)$. Quelques solutions sont représentées sur la figure suivante.



On retrouve le phénomène, déjà observé dans l'exemple 3.22, de la divergence de la dérivée $y'(x)$ sur les droites d'équation $y = \pm x$. En effet, si on écrit l'équation différentielle sous la forme $y'(x) = f(x, y(x))$ on observe que

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

est singulière sur les droites $y = \pm x$. On remarque aussi que plusieurs solutions passent par le point $x = y = 0$, ce qui indique l'existence de multiples solutions du problème de Cauchy ! En fait l'intégrale première $F(x, y)$ est singulière sur la droite $y = 0$, c'est ce qui explique pourquoi nous avons raté la solution évidente $y(x) = 0$ de l'équation différentielle dans la discussion ci-dessus ! Comme dans l'exemple 3.18, on peut obtenir de nouvelles solutions en recollant des morceaux de solutions. La figure ci-dessous en est une illustration. Elle montre une solution maximale du problème de Cauchy avec condition initiale $y(-1/2) = -2/5$.



On remarque donc que la solution est unique jusqu'à $x = 0$, mais qu'elle peut être continuée au delà de ce point de multiples façons. \triangleleft

On retiendra des deux exemples précédents qu'il faut toujours chercher les singularités de la fonction

$$f(x, y) = -\frac{A(x, y)}{B(x, y)},$$

lorsqu'on discute l'EDO (3.51), et que si on obtient une intégrale première $F(x, y)$, il est également nécessaire d'examiner ses singularités.

3.2.5 Chaos

Bibliographie

- [AS] M. Abramowitz, I.A. Stegun : *Handbook of Mathematical Functions*. Dover, New York, 1964. Disponible en ligne sur <http://www.math.ucla.edu/~cbm/aands/>.
- [A] V.I. Arnol'd : *Ordinary Differential Equations*. The MIT Presse, New York, 1998.
- [HNW] E. Hairer, S.P. Nørsett, G. Wanner : *Solving Ordinary Differential Equations I. Nonstiff Problems*. Springer, Berlin, 1993.
- [HW] E. Hairer, G. Wanner : *Solving Ordinary Differential Equations II. Stiff and Differential-Algebraic Problems*. Springer, Berlin, 1996.
- [H] P. Hartman : *Ordinary Differential Equations*. Wiley, New York, 1964.
- [LM1] F. Liret, D. Martinais : *Analyse Première Année*. Dunod, Paris, 2003.
- [LM2] F. Liret, D. Martinais : *Analyse Deuxième Année*. Dunod, Paris, 2004.
- [W] G.N. Watson : *A Treatise on the Theory of Bessel Functions*. Cambridge University Press, Londres, 1922.