Analyse II pour Ingénieurs et Scientifiques

Joachim STUBBE

31 janvier 2013

Table des matières

1 L'espace $\mathbb{R}^n$ **3**
- 1.1 L'espace vectoriel $\mathbb{R}^n$ . . . 3
- 1.2 L'espace normé $\mathbb{R}^n$ . . . 6
- 1.3 Sous-ensembles de $\mathbb{R}^n$ . . . 8
- 1.4 Suites dans $\mathbb{R}^n$ . . . 10
- 1.5 Fonctions continues . . . 12
- 1.6 Applications linéaires . . . 14
  - 1.6.1 Applications linéaires $\mathbb{R}^n \to \mathbb{R}^m$ . . . 14
  - 1.6.2 Applications linéaires entre des espaces normés . . . 15
  - 1.6.3 Applications linéaires dans un espace euclidien . . . 15
- 1.7 Addendum - Les matrices rechargées . . . 16
  - 1.7.1 Matrices carrées réelles . . . 16
  - 1.7.2 Matrices carrées complexes . . . 17
  - 1.7.3 Notions générales et théorie spectrale . . . 18
  - 1.7.4 Théorie spectrale des matrices hermitiennes . . . 19

2 Courbes dans $\mathbb{R}^n$ **20**
- 2.1 Courbes dérivables . . . 20
- 2.2 Longueurs de courbes . . . 25

3 Fonctions réelles sur $\mathbb{R}^n$ **30**
- 3.1 Introduction . . . 30
- 3.2 Limites et continuité d'une fonction réelle . . . 31
- 3.3 Dérivées partielles, différentielle et fonctions de la classe $\boldsymbol{C^1}$ . . . 33
  - 3.3.1 Dérivées partielles . . . 33
  - 3.3.2 Fonctions différentiables et différentielle . . . 36
  - 3.3.3 Fonctions de la classe $\boldsymbol{C^1}$ . . . 36
- 3.4 Hyperplan tangent . . . 37
- 3.5 Gradient et lignes de niveau . . . 38
- 3.6 Dérivées partielles d'ordre supérieur . . . 42

4 Champs vectoriels sur $\mathbb{R}^n$ **45**
- 4.1 Dérivées des Champs vectoriels . . . 45
  - 4.1.1 La règle de composition . . . 46
- 4.2 Champs vectoriels $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ . . . 47
  - 4.2.1 Le rotationnel . . . 50
  - 4.2.2 Inversibilité des champs vectoriels . . . 51
  - 4.2.3 Transformations du Gradient et du Laplacien . . . 56

4.3 Fonctions implicites 59
4.3.1 La technique de la différentiation implicite 63
4.4 Addendum - Quelques formules 64

5 Extremums locaux **65**
5.1 Extremums et points stationnaires 65
5.2 Formes quadratiques 66
5.3 Développement limité 67
5.4 Extremums locaux - conditions suffisantes 68
5.5 Extremums liés 73

6 Intégrales multiples **77**
6.1 Intégrales qui dépendent d'un paramètre 77
6.2 Intégrales doubles 81
6.2.1 Intégrale double sur un rectangle fermé 81
6.2.2 Intégrale double sur $\mathbb{R}^2$ 86
6.2.3 Changement de variables dans $\mathbb{R}^2$ 87
6.2.4 Calcul des intégrales multiples 89
6.3 Quelques applications en Physique 91
6.3.1 Le théorème de Newton 91
6.3.2 Les équations d'Euler-Lagrange 92
6.3.3 Ensemble microcanonique et l'équation d'un gaz parfait . 94

7 Equations différentielles **96**
7.1 Classification des équations différentielles 96
7.2 Existence et unicité locale des solutions 99
7.3 Résolution des certains équations différentielles du premier ordre 102
7.3.1 $\mathbf{y}' = \mathbf{f}(\mathbf{t})$ 102
7.3.2 $\mathbf{y}'(\mathbf{x}) + \mathbf{a}(\mathbf{x})\mathbf{y}(\mathbf{x}) = \mathbf{b}(\mathbf{x})$ 103
7.3.3 $\mathbf{y}' = \mathbf{f}(\mathbf{y})$ 104
7.3.4 La technique du changement des variables 106
7.3.5 Équations différentielles exactes 107
7.4 Equations différentielles du second ordre 108
7.4.1 Equations linéaires homogènes - propriétés générales . . . 108
7.4.2 Equations linéaires homogènes à coefficients constants . . 110
7.4.3 Equations linéaires inhomogènes 111
7.4.4 Equations linéaires inhomogènes à coefficients constants . 112
7.4.5 Equations autonômes et systèmes conservatifs 114
7.5 Systèmes d' équations différentielles linéares à coefficients constantes 114
7.5.1 Systèmes homogènes 114

Chapitre 1

L'espace $\mathbb{R}^n$

L'ensemble $\mathbb{R}^n$ est défini comme l'ensemble des n-tuplets ordonnés $(x_1, \ldots, x_n)$ de nombres réels. Ces n-tuplets sont appelés points de $\mathbb{R}^n$. En même temps on peut interpéter l'ensemble $\mathbb{R}^n$ comme espace vectoriel de dimension n. Dans ce chapitre nous passons d'abord en revue cette structure algébrique de $\mathbb{R}^n$. Ensuite nous introdusions une structure topologique qui nous permet d'étendre le concept du processus limite à l'espace $\mathbb{R}^n$. On retrouve ces structures de base dans des situations plus abstraites comme celles d'un espace normé ou d'un espace métrique auxquelles nous donnons également une brève introduction.

1.1 L'espace vectoriel $\mathbb{R}^n$

Notations. Une autre représentation des éléments de l'espace vectoriel $\mathbb{R}^n$ est donnée par des vecteurs-colonnes à n composantes (au lieu des n-tuplets $(x_1, \ldots, x_n)$. On les note

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

ou surmontés d'une flèche $\mathbf{x} = \vec{x}$. À chaque n-tuplet correspond un unique vecteur-colonne et vice versa. Pour le calcul nous utilisons généralement des vecteurs-colonnes pour les éléments de $\mathbb{R}^n$ et les deux notations pour les arguments des fonctions définies sur $\mathbb{R}^n$ en raison d'une meilleure lisibilité du texte.

Espace vectoriel. L'ensemble $\mathbb{R}^n$ peut également être considéré comme un espace vectoriel réel muni d'une addition

$$\mathbf{x} + \mathbf{y} = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n + y_n \end{pmatrix}$$

et d'une multiplication avec un scalaire $\lambda \in \mathbb{R}$ définie par

$$\lambda \mathbf{x} = \lambda \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 \\ \cdot \\ \cdot \\ \cdot \\ \lambda x_n \end{pmatrix}$$

Produit scalaire. L'espace vectoriel $\mathbb{R}^n$ est également muni d'un produit scalaire $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ défini par

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{n} x_k y_k.$$

De façon générale un produit scalaire $\langle \cdot, \cdot \rangle$ vérifie les trois propriétés suivantes :

1. Positivité : $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ pour tout $\mathbf{x}$ et $\langle \mathbf{x}, \mathbf{x} \rangle = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
2. Symétrie : $\langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$
3. Linéarité dans chaque argument : $\langle \alpha \mathbf{x} + \beta \mathbf{y}, \mathbf{z} \rangle = \alpha \langle \mathbf{x}, \mathbf{z} \rangle + \beta \langle \mathbf{y}, \mathbf{z} \rangle$ pour tous $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ et $\alpha, \beta \in \mathbb{R}$

L'algèbre linéaire : Un vecteur $\mathbf{x}$ est aussi une matrice $n \times 1$, sa transposée, notée $\mathbf{x}^T = (x_1, ..., x_n)$ est alors une matrice $1 \times n$. Par conséquent, on peut interpréter le produit scalaire de deux vecteurs $\mathbf{x}, \mathbf{y}$ comme produit matriciel de $\mathbf{x}^T$ et $\mathbf{y}$:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} = (x_1, \ldots, x_n) \cdot \begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}.$$

Dimension de $\mathbb{R}^n$**.** Les vecteurs

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_k = \begin{pmatrix} 0 \\ \cdot \\ 1 \\ \cdot \\ 0 \end{pmatrix}, \ldots, \mathbf{e}_n = \begin{pmatrix} 0 \\ \cdot \\ \cdot \\ 0 \\ 1 \end{pmatrix}$$

forment une base orthonormale de $\mathbb{R}^n$ et

$$\mathbf{x} = \sum_{k=1}^{n} x_k \mathbf{e}_k$$

pour tout $\mathbf{x} \in \mathbb{R}^n$ où $x_k = \langle \mathbf{x}, \mathbf{e}_k \rangle \in \mathbb{R}$. Alors,

$$\dim \mathbb{R}^n = n < \infty.$$

On dit que $\mathbb{R}^n$ est un espace vectoriel euclidien de dimension n. Tout espace vectoriel réel de dimension n peut être identifié avec $\mathbb{R}^n$.

Espaces euclidiens de dimension infinie. L'ensemble $l_2(\mathbb{R})$ consiste en les suites numériques $(x_k)_{k\in\mathbb{N}}$ telles que $\sum_{k=0}^{\infty} x_k^2 < \infty$. C'est un espace vectoriel muni de l'addition $(x_k)_{k\in\mathbb{N}} + (y_k)_{k\in\mathbb{N}} = (x_k + y_k)_{k\in\mathbb{N}}$. On définit un produit scalaire par

$$\langle (x_k)_{k\in\mathbb{N}}, (y_k)_{k\in\mathbb{N}} \rangle = \sum_{k=0}^{\infty} x_k y_k.$$

Il est plus commode de représenter les suites formellement par des vecteurs dans "$\mathbb{R}^\infty$". On peut donc écrire

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=0}^{\infty} x_k y_k.$$

L'ensemble $C([a,b])$ des fonctions continues $f : [a,b] \longrightarrow \mathbb{R}$ est un espace vectoriel. Sur $C([a,b])$ on peut définir un produit scalaire par

$$\langle f, g \rangle = \int_a^b f(x)g(x)\,dx.$$

L'espace vectoriel complexe $\mathbb{C}^n$. L'espace $\mathbb{C}^n$ est l'ensemble de vecteurs complexes $\mathbf{x}$ muni de l'addition habituelle et d'une multiplication externe par des nombres complexe. C'est un espace vectoriel complexe (c'est-à-dire sur le corps $\mathbb{C}$). On peut définir un produit scalaire par

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{k=1}^{n} \bar{x}_k y_k. \tag{1.1}$$

Le produit scalaire a la propriété d'être linéaire dans sa deuxième composante et anti-linéaire dans sa première composante. De plus, pour tous $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$

$$\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}.$$

Les vecteurs $\mathbf{e}_k$, $k = 1 \ldots n$, forment une base orthonormale. L'espace $\mathbb{C}^n$ est un espace hermitien de dimension n. En mécanique quantique non relativiste l'état de spin d'une particule de spin s est représenté par un vecteur dans $\mathbb{C}^{2s+1}$.

L'espace vectoriel réel $\mathbb{C}^n$. Un espace hermitien est toujours un espace euclidien si la multiplication externe est restreinte aux nombres réels. On prend le produit scalaire

$$\langle \mathbf{x}, \mathbf{y} \rangle_{\mathbb{R}} = \mathrm{Re}(\langle \mathbf{x}, \mathbf{y} \rangle) = \mathrm{Re}\big(\sum_{k=1}^{n} \bar{x}_k y_k \big). \tag{1.2}$$

Les vecteurs $\mathbf{e}_k, i\mathbf{e}_k$, $k = 1 \ldots n$, forment une base orthonormale par rapport à ce produit scalaire. Par conséquent, l'espace vectoriel réel $\mathbb{C}^n$ est de dimension $2n$. Nous pouvons donc identifier $\mathbb{C}^n$ avec l'espace euclidien $\mathbb{R}^{2n}$.

1.2 L'espace normé $\mathbb{R}^n$

Pour pouvoir étendre les methodes d'analyse développées en Analyse I à l'espace $\mathbb{R}^n$, celui doit être muni d'une structure topologique. Sur le corps $\mathbb{R}$ nous avons utilisé la valeur absolue $|\cdot|$ pour définir une distance $d(x,y) = |x-y|$. On a défini la convergence et la continuité dans $\mathbb{R}$ par cette distance d. On veut bien généraliser la valeur absolue et la distance à l'espace $\mathbb{R}^n$. Pour cela nous allons introduire les notions d'une norme et d'une métrique.

Définition - norme et espace normé. Soit E un espace vectoriel réel. Une fonction $||\cdot|| : E \longrightarrow \mathbb{R}_+$ est appelée une norme sur E si $||\cdot||$ vérifie les trois proprétés suivantes :

1. Positivité : $||\mathbf{x}|| \geq 0$ pour tout $\mathbf{x} \in E$ et $||\mathbf{x}|| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$
2. Homogénéité : $||\lambda \cdot \mathbf{x}|| = |\lambda| \cdot ||\mathbf{x}||$ pour tout $\lambda \in \mathbb{R}$ et $\mathbf{x} \in E$
3. Inégalité triangulaire : $||\mathbf{x} + \mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$ pour tout $\mathbf{x}, \mathbf{y} \in E$

Le couple $(E, ||\cdot||)$ est appelé un espace normé.

La norme euclidienne sur $\mathbb{R}^n$. La fonction $||\cdot||_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par

$$||\mathbf{x}||_2 = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle} = \left(\sum_{k=1}^{n} x_k^2 \right)^{\frac{1}{2}} \tag{1.3}$$

est une norme sur $\mathbb{R}^n$. On l'appelle la norme euclidienne (canonique) sur $\mathbb{R}^n$. L'inégalité triangulaire est une conséquence de l'inégalité de Cauchy-Schwarz

$$\langle \mathbf{x}, \mathbf{y} \rangle^2 \leq \langle \mathbf{x}, \mathbf{x} \rangle \langle \mathbf{y}, \mathbf{y} \rangle = ||\mathbf{x}||^2 ||\mathbf{y}||^2. \tag{1.4}$$

Autres normes sur $\mathbb{R}^n$.

$$||\mathbf{x}||_1 = \sum_{k=1}^{n} |x_k| \tag{1.5}$$

$$||\mathbf{x}||_\infty = \max_{1 \leq k \leq n} |x_k| \tag{1.6}$$

On dit que deux normes $||\cdot||$ et $|||\cdot|||$ sont équivalentes s'ils existent des constantes $C_1, C_2 > 0$ telles que pour tout $\mathbf{x} \in E$:

$$C_1 ||\mathbf{x}|| \leq |||\mathbf{x}||| \leq C_2 ||\mathbf{x}||$$

Évidemment, les trois normes $||\cdot||_1, ||\cdot||_2, ||\cdot||_\infty$ sont équivalentes, puisque

$$\frac{1}{\sqrt{n}} ||\mathbf{x}||_1 \leq ||\mathbf{x}||_2 \leq ||\mathbf{x}||_1$$

et

$$||\mathbf{x}||_\infty \leq ||\mathbf{x}||_2 \leq \sqrt{n}\, ||\mathbf{x}||_\infty.$$

Soit $||\cdot||$ une norme sur $\mathbb{R}^n$. Alors, il existe une constante $C > 0$ telle que

$$||\mathbf{x}|| \leq C ||\mathbf{x}||_2 \quad \text{pour tout } \mathbf{x} \in \mathbb{R}^n. \tag{1.7}$$

En fait, en écrivant $\mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i$, par les propriétés d'une norme et l'inégalité de Cauchy-Schwarz :

$$||\mathbf{x}|| \leq \sum_{i=1}^{n} |x_i| \cdot ||\mathbf{e}_i|| \leq \left(\sum_{i=1}^{n} x_i^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} ||\mathbf{e}_i||^2 \right)^{\frac{1}{2}} = C||\mathbf{x}||_2$$

avec $C = \left(\sum_{i=1}^{n} ||\mathbf{e}_i||^2 \right)^{\frac{1}{2}}$. Nous allons démontrer plus loin que toutes les normes sur $\mathbb{R}^n$ sont équivalentes.

Le notion d'une distance nous amène vers un concept encore plus général que celui d'un espace normé : l'espace métrique est un ensemble muni d'une distance mais sans demander une structure algébrique particulière (comme la structure linéaire d'un espace vectoriel).

Définition - métrique et espace métrique. Soit M un ensemble. Une fonction $d(\cdot, \cdot) : M \times M \longrightarrow \mathbb{R}_+$ est appellée une métrique sur M si d vérifie les trois propriétés suivantes :

1. Positivité : $d(\mathbf{x}, \mathbf{y}) \geq 0$ pour tous $\mathbf{x}, \mathbf{y} \in M$ et $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$
2. Symétrie : $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$
3. Inégalité triangulaire : $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$

pour $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$. Le couple (M, d) est appelé un espace métrique.

Une métrique sur les espace normés. Soit $(E, ||\cdot||)$ un espace normé. Alors la fonction $d(\cdot, \cdot) : E \times E \longrightarrow \mathbb{R}_+$ définie par

$$d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}|| \tag{1.8}$$

est une métrique. Donc tout espace normé est un espace métrique.

Métrique euclidienne dans $\mathbb{R}^n$.

$$d(\mathbf{x}, \mathbf{y}) = d_2(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2 = \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} \tag{1.9}$$

Cette métrique est appelé la métrique euclidienne ou la distance euclidienne.

Interprétation géométrique. Dans $\mathbb{R}^2$ ou $\mathbb{R}^3$ la métrique euclidienne $d_2(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||_2$ signifie la distance euclidienne (habituelle) de deux points $\mathbf{x}$ et $\mathbf{y}$. La norme euclidienne $||\mathbf{x}||_2$ correspond également à la longeur du vecteur $\mathbf{x}$. Le produit scalaire $\langle \mathbf{x}, \mathbf{y} \rangle$ donne une mesure de l'angle entre les vecteurs $\mathbf{x}$ et $\mathbf{y}$: On désigne $\theta = \sphericalangle(\mathbf{x}, \mathbf{y})$, alors

$$\langle \mathbf{x}, \mathbf{y} \rangle = ||\mathbf{x}||_2 ||\mathbf{y}||_2 \cos\theta.$$

En particulier, si $\mathbf{x}$ et $\mathbf{y}$ sont des vecteurs orthogonaux, i.e. $\theta = \pm\pi/2$, alors $\langle \mathbf{x}, \mathbf{y} \rangle = 0$.

L'espace normé $C([a,b])$. Il existe d'autres espaces normés que $\mathbb{R}^n$. Par exemple, l'ensemble $C([a,b])$ des fonctions continues $f : [a,b] \longrightarrow \mathbb{R}$ est un espace vectoriel. Sur $C([a,b])$ on peut définit les normes suivantes :

$$||f|| = ||f||_\infty = \max\{|f(x)| : x \in [a,b]\},$$

$$||f||_2 = \sqrt{\int_a^b |f(x)|^2\, dx}$$

ou encore

$$||f||_1 = \int_a^b |f(x)|\, dx.$$

Ces normes ne sont plus équivalentes puisque $C([a,b])$ est un espace vectoriel de dimension infinie. La norme "naturelle" sur $C([a,b])$ est $||f|| = \max\{|f(x)| : x \in [a,b]\}$, la norme de la convergence uniforme.

Exemple - espace métrique. Soit $M \subset \mathbb{R}^n$ muni de la distance euclidienne. Alors (M, d_2) est un espace métrique.

1.3 Sous-ensembles de $\mathbb{R}^n$

Pour définir certaines notions et propriétés des sous-ensembles de $\mathbb{R}^n$ on utilise seulement le fait que l'espace $\mathbb{R}^n$ est muni d'une métrique, par exemple la distance euclidienne. C'est pourquoi on préfère les donner pour un espace métrique (M, d) quelconque.

Boule ouverte. Soit $\mathbf{a} \in M$ et $r > 0$. L'ensemble

$$B(\mathbf{a}, r) = \{\mathbf{x} \in M : d(\mathbf{x}, \mathbf{a}) < r\}$$

est appelé la boule ouverte de centre $\mathbf{a}$ et de rayon r.

Voisinage. Un sous-ensemble $V \subset M$ est dit voisinage du point $\mathbf{a} \in M$ s'il existe un $\epsilon > 0$ tel que

$$B(\mathbf{a}, \epsilon) \subset V.$$

En particulier, $B(\mathbf{a}, \epsilon)$ est un voisinage de $\mathbf{a}$ dit ϵ-voisinage de $\mathbf{a}$.

Axiome de séparation de Hausdorff. Tout espace métrique satisfait l'axiome de séparation de Hausdorff : Pour tous $\mathbf{x}, \mathbf{y}$ distincts il existe deux voisinages $B(\mathbf{x}, \epsilon)$ et $B(\mathbf{y}, \epsilon)$ tels que $B(\mathbf{x}, \epsilon) \cap B(\mathbf{y}, \epsilon) = \emptyset$.

Sous-ensemble ouvert. Un sous-ensemble $S \subset M$ est dit ouvert si pour tout $x \in S$ il existe un voisinage $B(\mathbf{x}, \epsilon)$ tel que $B(\mathbf{x}, \epsilon) \subset S$. L'ensemble vide $\emptyset$ et l'espace M sont ouverts. Toute boule $B(\mathbf{a}, r)$ est un ensemble ouvert. Toute réunion quelconque d'ouverts est un ensemble ouvert. Toute intersection finie d'ouverts est un ouvert. Pour l'espace $\mathbb{R}^n$ la notion d'un ouvert ne dépend pas de la norme puisque toutes les normes sont équivalentes. Par exemple, soient

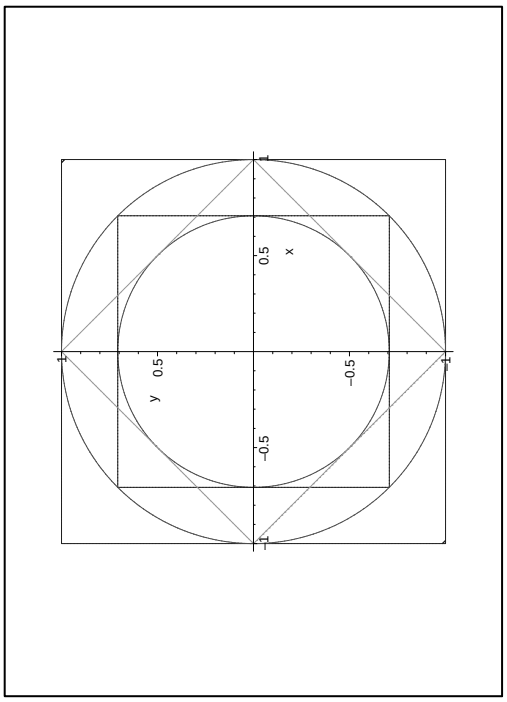
$$B(\mathbf{a}, \epsilon) = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x} - \mathbf{a}||_2 < \epsilon\}$$

les ϵ-voisinages par rapport à la norme euclidienne et

$$B'(\mathbf{a}, \epsilon) = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x} - \mathbf{a}||_\infty < \epsilon\}$$

les ϵ-voisinages par rapport à la norme du maximum, alors on a les inclusions suivantes :

$$B'(\mathbf{a}, \epsilon/\sqrt{n}) \subset B(\mathbf{a}, \epsilon) \subset B'(\mathbf{a}, \epsilon).$$



Des boules pour les normes $1, 2, \infty$

Sous-ensemble férmé. Un sous-ensemble $S \subset M$ est dit fermé si $M \setminus S$ est ouvert. L'ensemble vide $\emptyset$ et l'espace M sont fermés (et ouverts).

L'intérieur et le bord d'un ensemble. Soit $S \subset M$ et $a \in S$. On dit que a est dans l'intérieur de S s'il existe un voisinage $B(\mathbf{a}, \epsilon)$ tel que $B(\mathbf{a}, \epsilon) \subset S$. L'ensemble des points intérieurs à S est appelé l'intérieur de S et noté $\overset{\circ}{S}$. Un point $a \in M$ est appelé point frontière de S si tout voisinage $B(\mathbf{a}, \epsilon)$ contient des points de S et des points de $M \setminus S$. L'ensemble des points frontières à S est appelé le bord de S et noté ∂S.

Exemple. La boule unité ouverte dans $\mathbb{R}^n$ par rapport à la norme $||\cdot||$ est définie par

$$B_1 = B(\mathbf{0}, 1) = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}|| < 1\}.$$

Son bord est la sphère

$$\partial B_1 = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}|| = 1\}.$$

L'adhérence d'un ensemble. Soit $S \subset M$ et $a \in M$. On dit que a est adhérent à S si pour tout voisinage $B(\mathbf{a}, \epsilon)$:

$$B(\mathbf{a}, \epsilon) \cap S \neq \emptyset.$$

L'ensemble des points adhérents à S est appelé l'adhérence de S et noté $\bar{S}$.

Proposition. Soit (M, d) un espace métrique et $S \subset M$.

1. $\overset{\circ}{S} \subset S \subset \bar{S}$.
2. $\bar{S} = \overset{\circ}{S} \cup \partial S$.
3. S est ouvert si et seulement si $S = \overset{\circ}{S}$.
4. S est fermé si et seulement si $S = \bar{S}$.

Exemples.

1. $\bar{B_1} = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}|| \leq 1\}$. Plus généralement, le sous-ensemble

   $$\bar{B}(\mathbf{a}, r) = \{\mathbf{x} \in M : d(\mathbf{x}, \mathbf{a}) \leq r\}$$

   est appelé la boule fermée de centre $\mathbf{a}$ et de rayon r.
2. Soit $f : \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue. Le graphe $G_f = \{(x, f(x)) \in \mathbb{R}^2 : x \in \mathbb{R}\}$ représente une courbe dans $\mathbb{R}^2$. On a $\overset{\circ}{G}_f = \emptyset$. Donc $G_f = \partial G_f$. Le graphe d'une fonction continue est un ensemble fermé dans $\mathbb{R}^2$.
3. Soit $B = \{\mathbf{x} \in \mathbb{R}^2 : ||\mathbf{x}||_2 < 1\}$ et $I = [0, 5]$. L'ensemble S défini par

   $$S = B \times I = \{\mathbf{x} \in \mathbb{R}^3 : x_1^2 + x_2^2 < 1 \text{ et } 0 \leq x_3 \leq 5\}$$

   est un cylindre. l'ensemble S n'est ni ouvert ni fermé. Le bord de S est donné par

   $$\partial S = \partial B \times I \cup B \times \partial I$$

1.4 Suites dans $\mathbb{R}^n$

Nous introdusions les notions de convergence d'une suite dans un espace métrique et de l'espace métrique complet.

Suites. Soit (M, d) un espace métrique. Une suite d'éléments de M est une application $f : \mathbb{N} \longrightarrow M$, qui , à tout $k \in \mathbb{N}$, associe un élément $\mathbf{x}_k = f(k) \in M$. Les suites sont notées $(\mathbf{x}_k)_{k \in \mathbb{N}}$ ou $(\mathbf{x}_k)$.

Suite convergente. Une suite $(\mathbf{x}_k)$ *converge vers* $\mathbf{x} \in M$, si à tout $\epsilon > 0$, on peut associer un entier naturel N_ϵ tel que pour tout $n \geq N_\epsilon$ on a $d(\mathbf{x}_n, \mathbf{x}) < \epsilon$. On écrit alors

$$\lim_{n \to +\infty} \mathbf{x}_n = \mathbf{x}.$$

On dit aussi que la suite $(\mathbf{x}_n)_n$ est *convergente* et admet pour *limite* $\mathbf{x} \in M$. Lorsque la limite existe, elle est unique. Une suite non convergente est dite *divergente*.

Suite de Cauchy. Une suite est dite *suite de Cauchy* si à tout $\epsilon > 0$, on peut associer un $N = N_\epsilon \in \mathbb{N}$ tel que $m, n \geq N$ impliquent $d(\mathbf{x}_n, \mathbf{x}_m) < \epsilon$.

Proposition. Toute suite convergente $(x_n)_n$ est une suite de Cauchy.

Démonstration. voir Analyse 1.

Espace métrique complet. Dans ce cadre général on ne peut pas s'attendre que (comme c'est le cas pour les suites réelles, voir Analyse I, ch. 2.5) toute suite de Cauchy est convergente. On dit q'un espace métrique (M, d) est complet si toute suite de Cauchy converge dans M.

Espace normé complet. Nous rappelons que la métrique sur un espace normé $(E, ||\cdot||)$ est donnée par $d(\mathbf{x}, \mathbf{y}) = ||\mathbf{x} - \mathbf{y}||$. Un espace normé est dit complet si toute suite de Cauchy converge par rapport à cette métrique. Un espace normé complet est appelé espace de Banach.

Exemple - L'espace de Banach $\mathbb{R}$. Nous avons démontré au chapitre 2.5., Analyse I, que l'ensemble $\mathbb{R}$ muni de la métrique $d(x, y) = |x-y|$ est complet. Ce résultat est à la base du résultat correspondent sur l'espace normé $(\mathbb{R}^n, ||\cdot||_2)$.

Proposition. Une suite $(\mathbf{x}_k)_k$ converge dans l'espace normé $(\mathbb{R}^n, ||\cdot||_2)$ si est seulement si les n suites numériques $(x_{1,k})_k, \ldots, (x_{n,k})_k$ convergent. Il en suit le théorème suivant :

Théorème. Toute suite de Cauchy dans $\mathbb{R}^n$ converge. Par conséquent, l'espace normé $(\mathbb{R}^n, ||\cdot||_2)$ est complet.

Suites bornées. Une suite $(\mathbf{x}_k)$ est dite bornée dans un espace métrique (M, d) s'il existe $\mathbf{a} \in M$ et une constante $R > 0$ telle que $d(\mathbf{x}_k, a) \leq R$ pour tout $k \in \mathbb{N}$. Autrement dit, une suite $(\mathbf{x}_k)$ est dite bornée s'il existe une boule $B(\mathbf{a}, R) \subset M$ telle que $\mathbf{x}_k \in B(\mathbf{a}, R)$ pour tout $k \in \mathbb{N}$.

Suites bornées dans $\mathbb{R}^n$. Une suite $(\mathbf{x}_k)$ est bornée dans $(\mathbb{R}^n, ||\cdot||)$ s'il existe une constante $C > 0$ telle que $||\mathbf{x}_k|| \leq C$ pour tout $k \in \mathbb{N}$. Le théorème de Bolzano-Weierstrass s'applique aussi aux suites bornées dans $\mathbb{R}^n$:

Théorème de Bolzano-Weierstrass dans $\mathbb{R}^n$. De toute suite bornée $(\mathbf{x}_k)_k$ dans $\mathbb{R}^n$ on peut extraire une sous-suite convergente $(\mathbf{x}_{n_k})$.

Suites et ensembles dans $\mathbb{R}^n$. Soit $S \subset \mathbb{R}^n$. Alors, $\mathbf{x} \in \bar{S}$ si et seulement si $\mathbf{x}$ est la limite d'une suite d'éléments de S. Si S est fermé et borné, alors de toute suite d'éléments de S peut extraire une sous-suite convergente dans S. En particulier, si la suite est convergente sa limite est dans S. Il en suit

Exemple - Un espace métrique complet. Soit $M \subset \mathbb{R}^n$ fermé et muni de la distance euclidienne. Alors (M, d_2) est un espace métrique complet.

Exemple - Un espace de Banach sans théorème de Bolzano-Weierstrass. On considère $(C([a,b]), ||\cdot||_\infty)$. C'est un espace de Banach puisque si $(f_n)_n$ est une suite de Cauchy par rapport à la norme du maximum $||\cdot||_\infty$, alors par le théorème 4.5 de l'Analyse 1 sur la limite uniforme des fonctions continues la suite $(f_n)_n$ converge vers une fonction $f \in C([a,b])$. Il n'y a pas un théorème de Bolzano-Weierstrass pour la norme $||\cdot||_\infty$. En fait, considérer les fonctions

$$f_n(x) = (b-a)^{-n}(x-a)^n.$$

Alors, $||f_n||_\infty = 1$ pour tout $n \in \mathbb{N}$ et les f_n convergent ponctuellement (simplement) vers la fonction

$$f(x) = \begin{cases} 0, & \text{si } a \leq x < b\,; \\ 1, & \text{si } x = b. \end{cases}$$

Il en suit qu'il n'y a aucune sous-suite qui converge uniforément puisque la limite uniforme est également la limite ponctuelle qui est dans notre cas une fonction discontinue. La question pour quels sous-ensembles d'un espace normé (ou plus généralement d'un espace métrique) on a un théorème de Bolzano-Weierstrass nous amène à la notion d'un ensemble compact que nous n'utilisons pas dans ce cours. Pour l'étudiant intéressé nous donnons seulement sa définition et décrivons sa situation dans ce contexte mathématique. La notion de compacité est donnée par des recouvrements : Un ensemble E dans un espace métrique est compact si de tout recouvrement de E par des ensembles ouverts on peut extraire un sous-recouvrement fini.

Ensemble compact. Pour des ensembles compacts le théorème de Bolzano-Weierstrass est vérifié : De toute suite dans un ensemble compact on peut extraire une sous-suite convergente. Les fonctions continues préservent des ensembles compacts : l'image d'un compact est compacte. Dans un espace normé de dimension finie un ensemble est compact si et seulement si l'ensemble est borné et fermé. C'est le théorème de Heine-Borel (et pour cette raison nous travaillons seulement avec la notion "borné et fermé").

1.5 Fonctions continues

Le concept d'une fonction continue s'étend aux applications entre deux espaces métriques. Cette définition s'applique aux fonctions $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ que nous étudions dans les chapitres suivants.

Fonction continue. Soient (X, d_X) et (Y, d_Y) deux espaces métriques. Une fonction $\mathbf{f} : X \longrightarrow Y$ est dite continue en $\mathbf{a} \in X$ si

$$\lim_{\mathbf{x}\to\mathbf{a}} \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}),$$

c'est-à-dire, si pour toute suite $(\mathbf{x}_n)$ dans X qui converge vers $\mathbf{a}$ on a

$$\lim_{n\to\infty} \mathbf{f}(\mathbf{x_n}) = \mathbf{f}(\mathbf{a}).$$

Proposition. La fonction $\mathbf{f} : X \longrightarrow Y$ est continue en $\mathbf{a} \in X$ si pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que

$$d_X(\mathbf{x}, \mathbf{a}) < \delta \quad \Rightarrow \quad d_Y(\mathbf{f}(\mathbf{x}), \mathbf{f}(\mathbf{a})) < \epsilon.$$

Pour $X = Y = \mathbb{R}$ et $d_X(x, y) = d_Y(x, y) = |x - y|$ nous récuperons le résultat sur la continuité d'une fonction $f : \mathbb{R} \longrightarrow \mathbb{R}$.

Exemple 1. La norme euclidienne $||\cdot||_2 : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ est continue sur l'espace normé $(\mathbb{R}^n, ||\cdot||_2)$. (Idée : Montrer l'inégalité $\big|\, ||\mathbf{x}||_2 - ||\mathbf{y}||_2 \big| \leq ||\mathbf{x} - \mathbf{y}||_2$.)

Exemple 2. Toute norme $||\cdot|| : \mathbb{R}^n \longrightarrow \mathbb{R}_+$ est continue sur l'espace normé $(\mathbb{R}^n, ||\cdot||_2)$. En fait comme ci-dessus

$$\big|\, ||\mathbf{x}|| - ||\mathbf{y}|| \big| \leq ||\mathbf{x} - \mathbf{y}||$$

et par l'inégalité (1.7) : $||\mathbf{x} - \mathbf{y}|| \leq C||\mathbf{x} - \mathbf{y}||_2$ d'où l'affirmation.

Exemple 3. Soit (X, d_X) un espace métrique. Alors, pour tout $\mathbf{x}_0 \in X$ la fonction $f(\mathbf{x}) := d(\mathbf{x}, \mathbf{x}_0)$ est continue. En fait, grâce à l'inégalité triangulaire $|d(\mathbf{x}, \mathbf{x}_0) - d(\mathbf{y}, \mathbf{x}_0)| \leq d(\mathbf{x}, \mathbf{y})$ et on peut choisir $\delta = \epsilon$.

Comme nous avons vu en Analyse I le théorème de Bolzano-Weierstrass implique le résultat suivant sur les fonctions rélles continues.

Théorème. Soit $M \subset (\mathbb{R}^n, ||\cdot||_2)$ borné et fermé. Soit $f : \mathbb{R}^n \to \mathbb{R}$ continue. Alors, f atteint son minimum et son maximum sur M.

Corollaire. Soit $||\cdot||$ une norme sur $\mathbb{R}^n$. Alors, $||\cdot||$ est equivalente à $||\cdot||_2$.

Démonstration. $M = \{\mathbf{x} \in \mathbb{R}^n : ||\mathbf{x}||_2 = 1\}$ est un sous-ensemble borné et fermé de $(\mathbb{R}^n, ||\cdot||_2)$. La fonction continue $||\cdot||$ atteint son minimum et son maximum sur M. Par l'homogénéité des normes ils existent $C_1, C_2 > 0$ telles que

$$C_1||\mathbf{x}||_2 \leq ||\mathbf{x}|| \leq C_2||\mathbf{x}||_2 \tag{1.10}$$

pour tout $\mathbf{x} \in \mathbb{R}^n$.

Théorème du point fixe de Banach. Soit (M, d) un espace métrique complet. Soit $f : M \longrightarrow M$ une contraction, c'est-à-dire il existe $0 < q < 1$ tel que pour tout $x, y \in M$

$$d(f(x), f(y)) \leq q\, d(x, y). \tag{1.11}$$

Alors, il existe un unique $\bar{x} \in M$ tel que $f(\bar{x}) = \bar{x}$.

Convergence uniforme. Soient (X, d_X) et (Y, d_Y) deux espaces métriques. Soient $\mathbf{f}_n, \mathbf{f} : X \longrightarrow Y$. Les $\mathbf{f}_n$ convergent uniformément vers $\mathbf{f}$ si, pour tout $\epsilon > 0$, il existe un entier naturel $N = N_\epsilon$ tel que $d(\mathbf{f}_n(\bar{x}), \mathbf{f}(\bar{x})) < \epsilon$ pour tout $\mathbf{x} \in X$ et tout $n \geq N_\epsilon$.

En Analyse I nous avons démontré que la limite uniforme de fonctions continues $f_n : \mathbb{R} \to \mathbb{R}$ est continue. C'est résultat reste vrai pour des applications entre des espaces métriques.

Théorème. Soient (X, d_X) et (Y, d_Y) deux espaces métriques et $\mathbf{f}_n : X \longrightarrow Y$ des fonctions continues qui convergent uniformément vers $\mathbf{f} : X \longrightarrow Y$. Alors, $\mathbf{f}$ est continue.

1.6 Applications linéaires

1.6.1 Applications linéaires $\mathbb{R}^n \to \mathbb{R}^m$

À toute application linéaire $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ on peut associer une matrice $A \in M_{m,n}$ telle que

$$\mathbf{f}(\mathbf{x}) = A\mathbf{x} \tag{1.12}$$

pour tout $\mathbf{x} \in \mathbb{R}^n$. Plus explicitement, si $A = (a_{ij})_{ij}$, $i = 1 \dots m$, $j = 1 \dots n$, alors

$$A\mathbf{x} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_j \mathbf{e}_i$$

où les $\mathbf{e}_i \in \mathbb{R}^m$ sont les vecteurs de la base canonique. Par l'inégalité de Cauchy-Schwarz :

$$||A\mathbf{x}||_2^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{l=1}^{n} a_{ij} a_{il} x_j x_l = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} x_j \right)^2 \leq \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right) ||\mathbf{x}||_2^2.$$

On dit que l'application définie par A est bornée puisque on a trouvé une constante $C > 0$ telle que

$$||A\mathbf{x}||_2 \leq C||\mathbf{x}||_2 \tag{1.13}$$

pour tout $\mathbf{x} \in \mathbb{R}^n$. Ici $C = ||A||_2 := \left(\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2 \right)^{\frac{1}{2}}$. Il en suit la continuité de l'application linéaire $\mathbf{f}(\mathbf{x}) = A\mathbf{x}$.

Proposition. Toute application linéaire $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ est bornée et donc continue.

Démonstration. Pour tout $\epsilon > 0$ choisir $\delta = C\epsilon^{-1}$ avec C de l'équation (1.13).

Forme bilinéaire. On peut identifier les matrices $A \in M_{m,n}$ avec des formes bilinéaires $b : \mathbb{R}^m \times \mathbb{R}^n \to \mathbb{R}$ par

$$b(\mathbf{y}, \mathbf{x}) = \langle \mathbf{y}, A\mathbf{x} \rangle_{\mathbb{R}^m} = \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} x_j y_i \tag{1.14}$$

Noter que

$$b(\mathbf{y}, \mathbf{x}) = \langle A^T \mathbf{y}, \mathbf{x} \rangle_{\mathbb{R}^n}$$

où $A^T \in M_{n,m}$ est la matrice transposée de A.

1.6.2 Applications linéaires entre des espaces normés

Soient $(X, ||\cdot||_X)$ et $(Y, ||\cdot||_Y)$ deux espaces normés et $A : X \longrightarrow Y$ une application linéaire, c'est-à-dire $A(t_1\mathbf{x}_1 + t_2\mathbf{x}_2) = t_1A(\mathbf{x}_1) + t_2A(\mathbf{x}_2)$ pour tous scalaires t_1, t_2 et $\mathbf{x}_1, \mathbf{x}_2 \in X$. On dit aussi que A définit un opérateur linéaire entre X et Y.

Application linéaire bornée. Une application linéaire $A : X \longrightarrow Y$ entre deux espaces normés X, Y est dite bornée s'il existe $C > 0$ tel que

$$||A(\mathbf{x})||_Y \leq C||\mathbf{x}||_X \quad \text{pour tout } \mathbf{x} \in X. \tag{1.15}$$

Noter q'une application linéaire bornée est lipschitzienne.

Théorème - Critère de continuité pour des applications linéaires. Une application linéaire $A : X \longrightarrow Y$ entre deux espaces normés X, Y est continue si et seulement si elle est bornée.

Démonstration. D'abord notons qu'une application linéaire $A : X \longrightarrow Y$ est continue si et seulement si elle est continue en $\mathbf{x} = \mathbf{0}$. Si A est borné, alors elle est continue (voir la proposition ci-dessus). Si A est continu, alors pour tout $\epsilon > 0$ il existe $\delta > 0$ tel que $||\mathbf{x}||_X < \delta$ implique $||A(\mathbf{x})||_Y < \epsilon$. Fixons un tel couple (ϵ, δ) et supposons que A n'est pas borné. Si A n'est pas borné, alors pour toute constante $C > 0$ il existe $\mathbf{x}_c \in X$ tel que $||A(\mathbf{x}_c)||_Y > C||\mathbf{x}_c||_X$. Choisir $C = \dfrac{2\epsilon}{\delta}$ et prendre un tel $\mathbf{x}_c$. Définir $\mathbf{x} = \dfrac{\mathbf{x}_c\delta}{2||\mathbf{x}_c||_X}$. Alors, $||\mathbf{x}||_X < \delta$ et

$$||A(\mathbf{x})||_Y = \frac{\delta||A(\mathbf{x})||_Y}{2||\mathbf{x}_c||_X} > \frac{\delta C||\mathbf{x}_c||_X}{2||\mathbf{x}_c||_X} = \epsilon$$

d'où la contradiction.

Norme d'une application linéaire continue. Soit $A : X \longrightarrow Y$ application linéaire continue entre deux espaces normés X, Y. On définit la norme $||A||$ de l'application par

$$||A|| = \sup_{||\mathbf{x}||_X \leq 1} ||A(\mathbf{x})||_Y < \infty. \tag{1.16}$$

Il en suit que $||A(\mathbf{x}/||\mathbf{x}||_X)||_Y \leq ||A||$ pour tout $\mathbf{x} \neq \mathbf{0}$ d'où

$$||A(\mathbf{x})||_Y \leq ||A|| \cdot ||\mathbf{x}||_X$$

et $||A||$ donne la meilleure constante (c'est à-dire la plus petite) dans cette inégalité.

1.6.3 Applications linéaires dans un espace euclidien

Soit $(E, \langle\cdot,\cdot\rangle)$ un espace euclidien. Soit $A : E \to E$ un opérateur linéaire. Si $\dim E = n$ on peut identifier E avec $\mathbb{R}^n$. On peut l'identifier avec la forme bilinéaire $b : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ par

$$b(\mathbf{y}, \mathbf{x}) = \langle \mathbf{y}, A\mathbf{x} \rangle = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_jy_i \tag{1.17}$$

L'opérateur adjoint $A^T : E \to E$ est défini par

$$b(\mathbf{y}, \mathbf{x}) = \langle A^T \mathbf{y}, \mathbf{x} \rangle_{\mathbb{R}^n}$$

et A^T est la matrice transposée de A (si $\dim E$ est finie). L'opérateur est dit symétrique si $A = A^T$. Un opérateur linéaire $P : E \to E$ est un projecteur si $P^2 = P$, c'est-à-dire $P(P\mathbf{x}) = P\mathbf{x}$ pour tout $\mathbf{x} \in E$. Si, de plus, P est symétrique on appelle P un projecteur orthogonal. Si P est un projecteur orthogonal, le théorème de Pythagore est vérifié : Pour tout $\mathbf{x} \in E$

$$||\mathbf{x}||^2 = ||P\mathbf{x}||^2 + ||(1-P)\mathbf{x}||^2 \tag{1.18}$$

où $||\mathbf{x}||^2 = \langle \mathbf{x}, \mathbf{x} \rangle$. Il en suit $||P\mathbf{x}|| \leq ||\mathbf{x}||$. Soit $\mathbf{v} \in E$, $\langle \mathbf{v}, \mathbf{v} \rangle = 1$. Alors

$$P\mathbf{x} = \langle \mathbf{v}, \mathbf{x} \rangle \mathbf{v} \tag{1.19}$$

définit un projecteur orthogonal (c'est la projection orthogonale sur $\mathbf{v}$). Sa représentation matricielle est donnée par la matrice $\mathbf{v} \cdot \mathbf{v}^T$ qu'on écrit en physique aussi dans la notation "bra-ket" comme $\mathbf{v}\rangle\langle\mathbf{v}$ puisque la forme bilinéaire y associée est

$$b(\mathbf{y}, \mathbf{x}) = \langle \mathbf{y}, \mathbf{v} \rangle \langle \mathbf{v}, \mathbf{x} \rangle.$$

1.7 Addendum - Les matrices rechargées

1.7.1 Matrices carrées réelles

Soit $A \in M_{n,n}(\mathbb{R})$. Alors, $A\mathbf{x} = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_j \mathbf{e}_i$ et $a_{ij} = \langle A\mathbf{e}_j, \mathbf{e}_i \rangle = \langle \mathbf{e}_i, A\mathbf{e}_j \rangle$.
On peut également représenter par les vecteurs des colonnes ou des lignes.

Matrice transposée. A^T. $(A^T)_{ij} = a_{ji}$. Si $A = (\mathbf{a}_1, \ldots, \mathbf{a}_n)$, $\mathbf{a}_j \in \mathbb{R}^n$, alors

$$A^T = \begin{pmatrix} \mathbf{a}_1^T \\ \cdots \\ \mathbf{a}_n^T \end{pmatrix}.$$

Matrice élémentaire. $E_{\alpha\beta}$. $(E_{\alpha\beta})_{ij} = \begin{cases} 1, & \text{si } i = \alpha \text{ et } j = \beta\,; \\ 0, & \text{sinon.} \end{cases}$ Pour tout $A \in M_{n,n}(\mathbb{R})$: $A = \sum_{\alpha=1}^{n} \sum_{\beta=1}^{n} a_{\alpha\beta} E_{\alpha\beta}$.

Matrice unité. I_n (ou 1), $(I_n)_{ij} = \delta_{ij} = \begin{cases} 1, & \text{si } i = j\,; \\ 0, & \text{si } i \neq j. \end{cases}$ $I_n = \sum_{\alpha=1}^{n} E_{\alpha\alpha}$.

Matrice symétrique. $A = A^T$ ou $a_{ij} = a_{ji}$.

Matrice antisymétrique. $A = -A^T$ ou $a_{ij} = -a_{ji}$. En particulier, $a_{ii} = 0$.

Décomposition d'une matrice dans sa partie symétrique et antisymétrique.

$$A = \frac{A + A^T}{2} + \frac{A - A^T}{2}.$$

Matrice orthogonale. $A^T = A^{-1}$. Alors, $AA^T = A^T A = 1$.

Projecteur. $P \in M_{n,n}(\mathbb{R})$ définit un projecteur si $P^2 = P$. Si, de plus $P = P^T$, la matrice P définit un projecteur orthogonal. Par exemple,

$$P = \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix}$$

est un projecteur et

$$P = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

est un projecteur orthogonal. La matrice de la projection orthogonale $P\mathbf{x} = \langle \mathbf{v}, \mathbf{x} \rangle \mathbf{v}$ (voir éq. (1.19)), $\langle \mathbf{v}, \mathbf{v} \rangle = 1$ est donnée par $P_{ij} = v_i v_j$.

Matrice de Householder. $H_{\mathbf{v}} = 1 - 2\dfrac{\mathbf{v}\rangle\langle\mathbf{v}}{\langle \mathbf{v}, \mathbf{v} \rangle}$ pour $\mathbf{v} \neq \mathbf{0}$. Autrement dit, si $\langle \mathbf{v}, \mathbf{v} \rangle = 1$, alors $H = 1 - 2P$ où P est la matrice de la projection orthogonale sur $\mathbf{v}$. H est symétrique et orthogonal : $H_{\mathbf{v}}^T = H_{\mathbf{v}}$,$H_{\mathbf{v}}^2 = 1$. On utilise une matrice de Householder pour démontrer le résultat suivant : Soient $\mathbf{a} \neq \mathbf{b}$ deux vecteurs tels que $||\mathbf{a}||_2 = ||\mathbf{b}||_2$. Alors, il existe une transformation orthogonale O telle que $O\mathbf{a} = \mathbf{b}$. En fait, on prend une matrice de Householder avec $\mathbf{v} = \mathbf{a} - \mathbf{b}$.

1.7.2 Matrices carrées complexes

Soit $A \in M_{n,n}(\mathbb{C})$, $(A)_{ij} = a_{ij}$.

Matrice conjuguée. $\bar{A}$, $(\bar{A})_{ij} = \overline{a_{ij}}$.

Matrice adjointe. A^* (ou $A^\dagger$), $(A^*)_{ij} = \overline{a_{ji}}$. Donc $A^* = \bar{A}^T$.

Matrice hermitienne. $A = A^*$, c'est-à-dire $a_{ij} = \overline{a_{ji}}$. En particulier, si $A \in M_{n,n}(\mathbb{R})$ et $A = A^T$, alors A est hermitien.

Matrice anti-hermitienne. $A = -A^*$, c'est-à-dire $a_{ij} = -\overline{a_{ji}}$. Alors, iA est une matrice hermitienne.

Décomposition d'une matrice. Pour tout $A \in M_{n,n}(\mathbb{C})$:

$$A = \frac{A + A^*}{2} + i\,\frac{A - A^*}{2i} = B + iC$$

où B, C sont des matrices hermitiennes.

Matrice unitaire. $A^* = A^{-1}$. Donc $A^*A = AA^* = I_n$. En particulier, si $A \in M_{n,n}(\mathbb{R})$ est une matrice orthogonale, alors A est une matrice unitaire.

Matrices de Pauli. Les matrices de Pauli sont des matrices $M_{2,2}(\mathbb{C})$ qui sont unitaires et hermitiennes décrivant en mécanique quantique l'interaction du spin avec les champs éléctromagnétiques :

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (1.20)$$

Il en suit $\sigma_j^2 = 1 = -i\sigma_1\sigma_2\sigma_3$ pour tout $j = 1, 2, 3$. Pour tout $\vec{\mathbf{a}} \in \mathbb{R}^3$ on peut définir

$$\vec{\mathbf{a}}\vec{\sigma} = \sum_{j=1}^{3} a_j\sigma_j = \begin{pmatrix} a_3 & a_1 - ia_2 \\ a_1 + ia_2 & -a_3 \end{pmatrix}.$$

Alors, $\vec{\mathbf{a}}, \vec{\mathbf{b}} \in \mathbb{R}^3$:

$$(\vec{\mathbf{a}}\vec{\sigma})(\vec{\mathbf{b}}\vec{\sigma}) = \langle \vec{\mathbf{a}}, \vec{\mathbf{b}} \rangle + i(\vec{\mathbf{a}} \times \vec{\mathbf{b}})\vec{\sigma}$$

où $\times$ dénote le produit vectoriel entre deux vecteurs.

1.7.3 Notions générales et théorie spectrale

Trace et déterminant. $\det(A) = \sum_{\sigma} (-1)^{|\sigma|} a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}$ où $|\sigma|$ dénote le nombre d'inversions dans σ. $\mathrm{Tr}\,(A) = \sum_{i=1}^{n} a_{ii}$.

$$\det(AB) = \det A \det B, \quad \det A^T = \det A. \quad (1.21)$$

$$\mathrm{Tr}\,(A + B) = \mathrm{Tr}\,(A) + \mathrm{Tr}\,(B), \quad \mathrm{Tr}\,(AB) = \mathrm{Tr}\,(BA). \quad (1.22)$$

Commutateurs et Anti-commutateurs. $[A, B] = AB - BA$ (commutateur), $\{A, B\} = AB + BA$ (anti-commutateur).

Valeurs propres et vecteurs propres. $A\mathbf{u}_j = \lambda_j \mathbf{u}_j$, $\lambda_j \in \mathbb{C}$ (ou $\lambda_j \in \mathbb{C}$).

Résolvente. Si $(A - \lambda)^{-1}$ existe et définit un opérateur linéaire bornée (donc continu) on définit la résolvante de A par $R_\lambda(A) = (A - \lambda)^{-1}$. Dans un espace à dimension finie $R_\lambda(A)$ existe si $A - \lambda$ est inversible. L'ensemble résolvant est l'ensemble des valeurs de λ pour lesquelles la résovante $R_\lambda(A)$ existe et l'on note $\rho(A)$.

Spectre. Le spectre $\sigma(A)$ est l'ensemble de λ pour lequels $R_A(\lambda)$ n'existe pas ; ce qui signifie dans le cas d'un espace de dimension finie que $(A - \lambda)$ n'est pas inversible. Le spectre est le complémentaire de l'ensemble résolvant : $\sigma(A) = \mathbb{C} \setminus \rho(A)$. Pour une matrice le spectre est l'ensemble de ses valeurs propres (réelles et complexes) donné par les zéros du polynôme caractéristique $p(\lambda) = \det(A - \lambda)$. Il en suit les identités suivantes pour les valeurs propres λ_j d'une matrice A :

$$\mathrm{Tr}\,(A) = \sum_{j=1}^{n} \lambda_j \quad (1.23)$$

et

$$\det(A) = \prod_{j=1}^{n} \lambda_j. \quad (1.24)$$

1.7.4 Théorie spectrale des matrices hermitiennes

Soit A hermitien (ou symétrique). Le spectre est réel et en cas de dimension finie il consiste en les valeurs propres réelles :

$$A\mathbf{u}_j = \lambda_j \mathbf{u}_j, \quad j = 1 \dots n, \quad \langle \mathbf{u}_j, \mathbf{u}_k \rangle = \delta_{jk}$$

Alors, $Q = (\mathbf{u}_1 \dots \mathbf{u}_n)$ est unitaire (orthogonale si A est symétrique) et

$$Q^* A Q = \sum_{j=1}^{n} \lambda_j E_{jj} \tag{1.25}$$

c'est-à-dire A est diagonalisable. Alternativement, on peut écrire

$$A = \sum_{j=1}^{n} \lambda_j \mathbf{u}_j \rangle\langle \mathbf{u}_j \tag{1.26}$$

ce qui signifie que

$$A\mathbf{x} = \sum_{j=1}^{n} \lambda_j \langle \mathbf{u}_j, \mathbf{x} \rangle \mathbf{u}_j.$$

Si deux matrices hermitiennes A, B commutent, c'est-à-dire $[A, B] = 0$, on peut construire des vecteurs propres communs.

Chapitre 2

Courbes dans $\mathbb{R}^n$

2.1 Courbes dérivables

Définition. Soit $I \subset \mathbb{R}$ un intervalle. Une courbe (ou un chemin) est une application continue

$$\mathbf{f} : I \longrightarrow \mathbb{R}^n.$$

Une courbe est donnée par un n-tuplet de fonctions continues f_i :

$$\mathbf{f} = \begin{pmatrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{pmatrix}.$$

Une courbe est dite dérivable si les n fonctions f_i sont dérivables. Une courbe est dite de classe $C^k(I)$ si les n fonctions f_i sont de classe $C^k(I)$. Pour tout $t \in I$ on appelle

$$\mathbf{f}'(t) = \begin{pmatrix} f_1'(t) \\ \vdots \\ f_n'(t) \end{pmatrix}$$

le vecteur tangent au point $\mathbf{f}(t)$.

Pour une courbe dérivable on a par définition de la dérivée

$$\mathbf{f}'(t) = \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{\mathbf{f}(t+h) - \mathbf{f}(t)}{h}$$

Autrement dit,

$$\mathbf{f}(t+h) = \mathbf{f}(t) + \mathbf{f}'(t)h + \mathbf{o}(h)$$

où $\mathbf{o}(h)$ désigne une courbe continue telle que

$$\lim_{h \to 0} \frac{\mathbf{o}(h)}{|h|} = \mathbf{0}.$$

Application en Mécanique. Une courbe décrit le mouvement d'un point si on intérprète la variable t comme le temps et $\mathbf{f}(t)$ comme la position du point au temps t (en mécanique les particules physiques sont représentées par des particules ponctuelles c'est-à-dire sans extension spatiale). Alors l'image de $\mathbf{f}$ décrit l'orbite, le graphe de $\mathbf{f}$ le mouvement dans l'espace-temps. Le vecteur tangent au temps t représente le vecteur de vitesse et

$$||\mathbf{f}'(t)||_2 = \sqrt{f_1'(t)^2 + \ldots + f_n'(t)^2}$$

est la vitesse au temps t. Le vecteur

$$\mathbf{f}''(t) = \begin{pmatrix} f_1''(t) \\ \vdots \\ f_n''(t) \end{pmatrix}$$

représente l'accélération au temps t. On écrit typiquement $\mathbf{r}(t), \mathbf{v}(t), \mathbf{a}(t)$ pour la position, la vitesse et l'accél ération et les dérivées temporelles sont souvent notées $\dot{\mathbf{r}}(t), \ddot{\mathbf{r}}(t), \ldots$. L'équation du mouvement pour une particule de masse m est donnée par la loi de Newton

$$m\mathbf{a}(t) = \mathbf{F}(t, \mathbf{r}(t), \mathbf{v}(t)) \tag{2.1}$$

où l'application F signifie la force aggissant sur la particule. C'est un système d'équations différentielles du second ordre auxquelles il faut encore ajouter des conditions initiales :

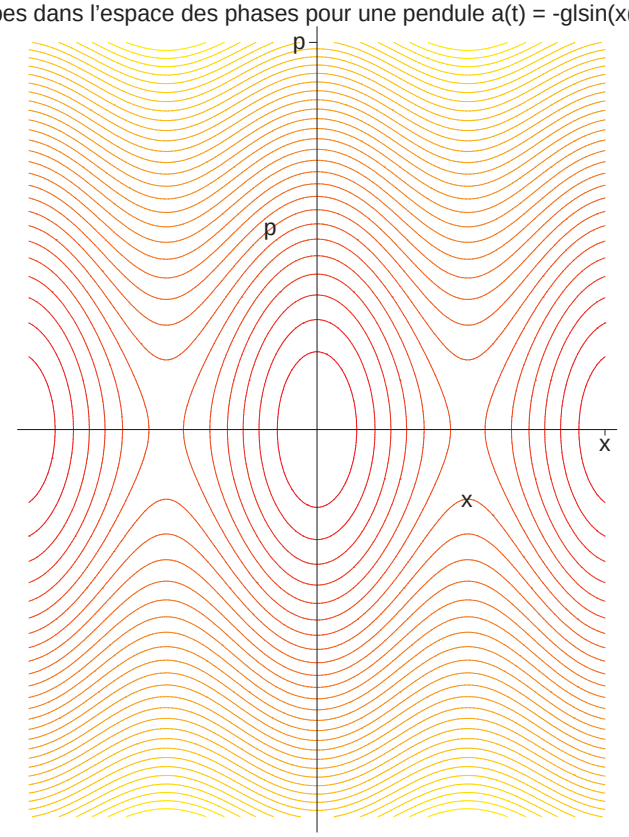
$$m\ddot{\mathbf{r}}(t) = \mathbf{F}(t, \mathbf{r}(t), \dot{\mathbf{r}}(t)), \quad \mathbf{r}(0) = \mathbf{r}_0, \dot{\mathbf{r}}(0) = \mathbf{v}_0. \tag{2.2}$$

Un système de N particules dans $\mathbb{R}^3$ peut être représenté par un vecteur dans $\mathbb{R}^{3N}$. Il est parfois commode de combiner le vecteur de position et le vecteur de vitesse d'une particule de masse m dans un vecteur

$$\begin{pmatrix} \mathbf{r}(t) \\ \mathbf{v}(t) \end{pmatrix} \quad \text{ou encore} \quad \begin{pmatrix} \mathbf{r}(t) \\ \mathbf{p}(t) \end{pmatrix}$$

où $\mathbf{p}(t) = m\mathbf{v}(t)$ désigne sa quantité de mouvement. L'espace de ces vecteurs est appelé l'espace des phases et joue un rôle important en physique.

Des courbes dans l'espace des phases pour une pendule a(t) = -glsin(x(t)).



Définition. Une courbe est dite régulière si $\mathbf{f}'(t) \neq \mathbf{0}$ pour tout $t \in I$. Un $t \in I$ tel que $\mathbf{f}'(t) = \mathbf{0}$ est appelé singulier (ou stationnaire).

Exemples.

1. Cercles. Soit $r > 0$ et $\mathbf{f} : [0, 2\pi] \longrightarrow \mathbb{R}^2$ donnée par

$$\mathbf{f}(t) = \begin{pmatrix} r\cos t \\ r\sin t \end{pmatrix},$$

ou dans $\mathbb{R}^3$ par la fonction

$$\mathbf{f}(t) = \begin{pmatrix} r\cos t \\ r\sin t \\ 0 \end{pmatrix}$$

Noter que dans $\mathbb{R}^2$:

$$\mathrm{Im}(\mathbf{f}) = \mathbf{f}([0, 2\pi]) = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = r^2\}.$$

Noter que dans $\mathbb{R}^3$:

$$\mathrm{Im}(\mathbf{f}) = \mathbf{f}([0, 2\pi]) = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = r^2, z = 0\}.$$

Dans $\mathbb{R}^2$ on a

$$\mathbf{f}'(t) = \begin{pmatrix} -r\sin t \\ r\cos t \end{pmatrix} \quad \text{et} \quad ||\mathbf{f}'(t)||_2 = r$$

ou dans $\mathbb{R}^3$ par la fonction

$$\mathbf{f}'(t) = \begin{pmatrix} -r\sin t \\ r\cos t \\ 0 \end{pmatrix} \quad \text{et} \quad ||\mathbf{f}'(t)||_2 = r$$

Cette courbe est régulière. Noter aussi que $\langle \mathbf{f}'(t), \mathbf{f}(t) \rangle = 0$ et $\mathbf{f}''(t) = -\mathbf{f}(t)$.

2. Une droite dans $\mathbb{R}^n$. Soit $\mathbf{r}_0, \mathbf{v} \in \mathbb{R}^n$, $t_0 \in \mathbb{R}$ $\mathbf{v} \neq \mathbf{0}$ et $\mathbf{r} : \mathbb{R} \longrightarrow \mathbb{R}^n$ donnée par

$$\mathbf{r}(t) = \mathbf{r_0} + \mathbf{v}(t - t_0).$$

Cette paramétrisation d'une droite décrit le mouvement d'une particule libre $\mathbf{a}(t) = 0$ avec des conditions initiales $\mathbf{r}(t_0) = \mathbf{r}_0, \mathbf{r}(t_0) = \mathbf{v}$. La vitesse est constante :

$$\dot{\mathbf{r}}(t) = \mathbf{v} \quad \text{et} \quad ||\dot{\mathbf{r}}(t)||_2 = ||\mathbf{v}||_2$$

Pour décrire son image dans $\mathbb{R}^n$ par un système d'équations il faut trouver $(n-1)$ vecteurs $\mathbf{w}_j$ orthogonaux et orthogonaux à $\mathbf{v}$. On a

$$\mathrm{Im}(\mathbf{r}) = \mathbf{r}(\mathbb{R}) = \{\mathbf{x} \in \mathbb{R}^n : \langle \mathbf{r} - \mathbf{r_0}, \mathbf{w}_j \rangle = 0\}$$

pour tout $j = 1, \ldots, n-1$. Une droite est décrite par un système linéaire de $(n-1)$ équations indépendantes. La droite $\mathbf{f}(t) = \mathbf{r}_0 + \mathbf{v}(t - t_0)$ est une courbe régulière.

3. Tout graphe G_f d'une fonction réelle continue peut être considéré comme courbe dans $\mathbb{R}^2$:

$$\mathbf{f}(t) = \begin{pmatrix} t \\ f(t) \end{pmatrix}$$

Évidemment $\mathrm{Im}(\mathbf{f}) = G_f$. Pour le vecteur tangent on trouve

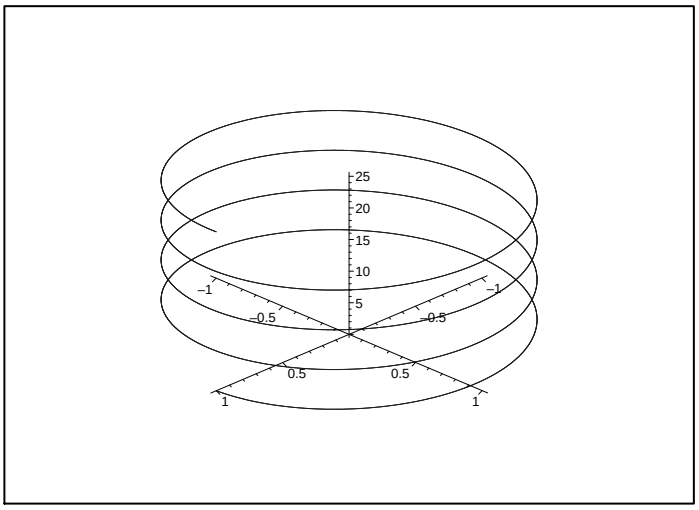
$$\mathbf{f}'(t) = \begin{pmatrix} 1 \\ f'(t) \end{pmatrix} \quad \text{et} \quad ||\mathbf{f}'(t)||_2 = \sqrt{1 + f'(t)^2}$$

4. Hélice. Pour $r > 0$ et $c \in \mathbb{R}$ soit $\mathbf{f} : \mathbb{R} \longrightarrow \mathbb{R}^3$ donnée par

$$\mathbf{f}(t) = \begin{pmatrix} r\cos t \\ r\sin t \\ ct \end{pmatrix}$$

Le vecteur tangent est donné par

$$\mathbf{f}'(t) = \begin{pmatrix} -r\sin t \\ r\cos t \\ c \end{pmatrix} \quad \text{et} \quad ||\mathbf{f}'(t)||_2 = \sqrt{r^2 + c^2}$$



5. Une courbe non injective. Soit $\mathbf{f} : \mathbb{R} \longrightarrow \mathbb{R}^2$ donnée par

$$\mathbf{f}(t) = \begin{pmatrix} t^2 - 1 \\ t^3 - t \end{pmatrix}$$

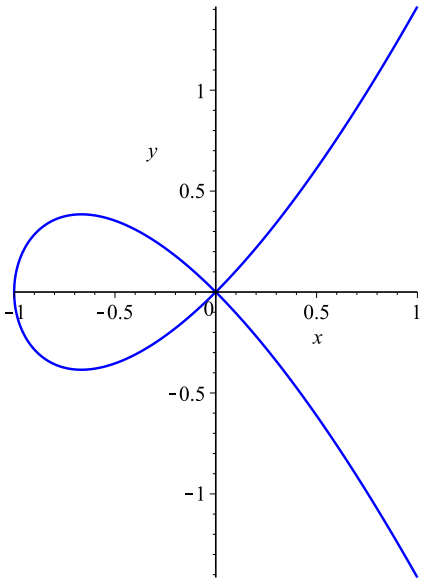
Nous avons $\mathbf{f}(-1) = \mathbf{f}(1) = \mathbf{0}$ et

$$\mathrm{Im}(\mathbf{f}) = \mathbf{f}(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : x^2 + x^3 = y^2\}$$

Nous calculons les vecteurs tangents :

$$\mathbf{f}'(t) = \begin{pmatrix} 2t \\ 3t^2 - 1 \end{pmatrix} \quad \text{et} \quad ||\mathbf{f}'(t)||_2 = \sqrt{9t^4 - 2t^2 + 1}$$

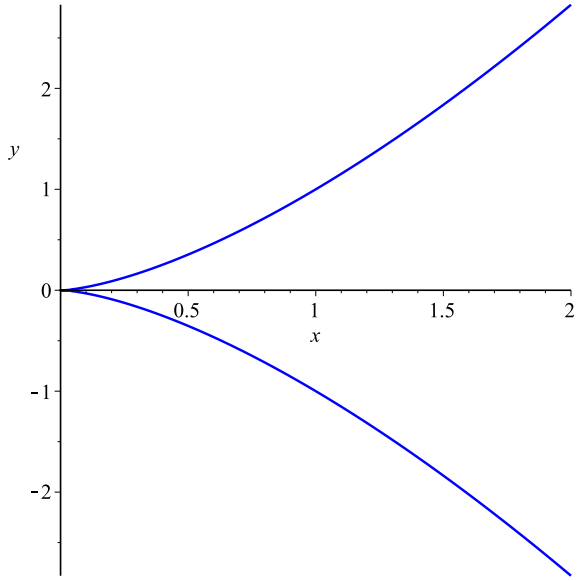
Donc $\mathbf{f}'(-1) = (-2, 2)$ et $\mathbf{f}'(1) = (2, 2)$.



6. Parabole de Neil (ou parabole semi-cubique). On considère la courbe $\mathbf{f} : \mathbb{R} \longrightarrow \mathbb{R}^2$ donnée par

$$\mathbf{f}(t) = \begin{pmatrix} t^2 \\ t^3 \end{pmatrix}.$$

Donc $\mathbf{f}'(t) = \begin{pmatrix} 2t \\ 3t^2 \end{pmatrix}$. Le point $(0,0)$ est un point singulier. L'image de $\mathbf{f}$ est donnée par l'équation $x^3 = y^2$.



Intersection de courbes regulières. Soient $\mathbf{f} : I_1 \longrightarrow \mathbb{R}^n$, $\mathbf{g} : I_2 \longrightarrow \mathbb{R}^n$ deux courbes regulières telles que $\mathbf{f}(t_1) = \mathbf{g}(t_2)$. Autrement dit, $\mathrm{Im}(\mathbf{f}) \cap \mathrm{Im}(\mathbf{g})$ n'est pas vide. L'angle d'intersection ϑ est défini comme l'angle entre les deux vecteurs tangents. On determine l'angle $\vartheta \in [0, \pi]$ par

$$\cos\vartheta = \frac{\langle \mathbf{f}'(t_1), \mathbf{g}'(t_2) \rangle}{||\mathbf{f}'(t_1)||_2 ||\mathbf{g}'(t_2)||_2}$$

Exemple. Considérons l'exemple 5. Nous avons $\mathbf{f}(-1) = \mathbf{f}(1) = \mathbf{0}$ et

$$\cos\vartheta = 0$$

i.e $\vartheta = \pi/2$.

2.2 Longueurs de courbes

L'idée est d'approcher une courbe par une suite de segments et définir ainsi le concept de longueur d'une courbe.

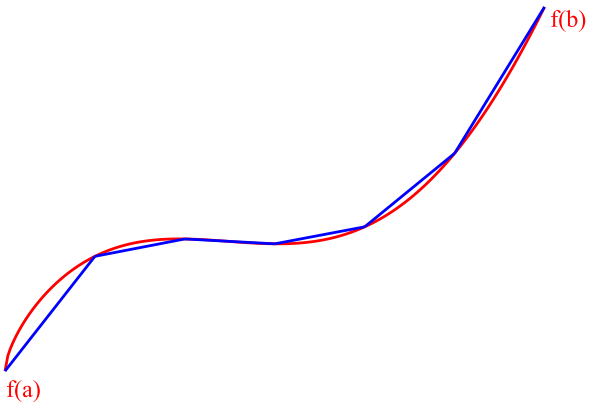
Longueur d'un segement. Soit $\mathbf{f} : I \longrightarrow \mathbb{R}^n$ une courbe et $t_1, t_2 \in I$. Le segment entre les points $\mathbf{f}(t_1)$ et $\mathbf{f}(t_2)$ est la ligne droite $\mathbf{s} : [0,1] \longrightarrow \mathbb{R}^n$ donnée par

$$\mathbf{s}(\tau) = \mathbf{f}(t_1)(1-\tau) + \mathbf{f}(t_2)\tau$$

Le longueur du segment est

$$L(\mathbf{s}) = ||\mathbf{f}(t_2) - \mathbf{f}(t_1)||_2.$$

Définition - courbe rectifiable. Une courbe $\mathbf{f} : [a,b] \longrightarrow \mathbb{R}^n$ est dite rectifiable si pour toute subdivision de $[a,b]$ dont le pas tend vers zéro, les longueurs des suites de segments convergent vers un nombre fini $L > 0$. On appelle L la longueur de la courbe $\mathbf{f}$.



Pour une courbe $\mathbf{f}$ dérivable l'idée est de construire une somme de Riemann pour $v(t) = ||\mathbf{f}'(t)||_2$.

Théorème. Soit $\mathbf{f}$ une courbe de classe $C^1([a,b])$. Alors $\mathbf{f}$ est rectifiable et

$$L = \int_a^b ||\mathbf{f}'(t)||_2 \, dt \tag{2.3}$$

On appelle L aussi la longueur d'arc de $\mathbf{f}$.

Démonstration. Soit $N \in \mathbb{N}$. Pour tout $k = 0, \ldots, N-1$ soit $a_k = \frac{k}{N}(b-a)$ on a par le théorème de la valeur moyenne (à appliquer à chaque f_i') :

$$f_i(a_{k+1}) - f_i(a_k) = \int_{a_k}^{a_{k+1}} f_i'(t)\, dt = \frac{b-a}{n} f_i'(c_{k,i}), \quad c_{k,i} \in]a_k, a_{k+1}[.$$

En calculant la norme euclidienne et ensuite la somme sur k on obtient :

$$\sum_{k=1}^{N-1} ||\mathbf{f}(a_{k+1}) - \mathbf{f}(a_k)||_2 = \frac{b-a}{N} \sum_{k=1}^{N-1} \sqrt{\sum_{i=1}^{n} f_i'(c_{k,i})^2}.$$

Par la continuité de $\mathbf{f}'$ et de la norme euclidienne la dernière somme converge vers l'intégrale de $||\mathbf{f}'(t)||_2$.

Théorème. La longueur d'arc donné par (2.3) ne dépend pas de la paramétrisation choisie. (voir l'analyse 3)

Une inégalité pour la longueur. Soit $\mathbf{f}$ une courbe de classe $C^1([a,b])$. Alors,

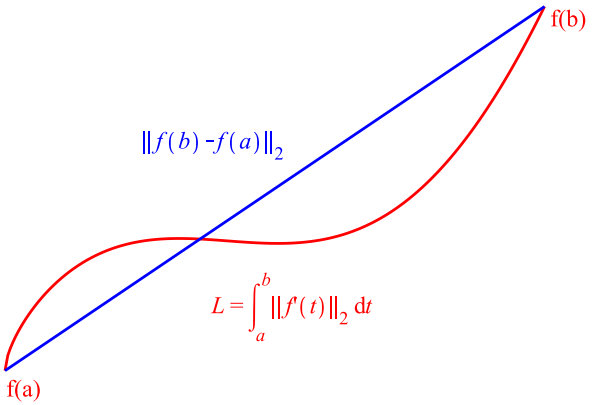
$$|| \int_a^b \mathbf{f}'(t)\, dt||_2 \leq \int_a^b ||\mathbf{f}'(t)||_2\, dt,$$

ce qui est équivalent à

$$||\mathbf{f}(b) - \mathbf{f}(a)||_2 \leq \int_a^b ||\mathbf{f}'(t)||_2\, dt.$$

Démonstration. On utilise la linéarité de l'intégrale et l'inégalité de Cauchy-Schwarz pour le produit scalaire dans $\mathbb{R}^n$:

$$\begin{aligned}
||\mathbf{f}(b) - \mathbf{f}(a)||_2^2 &= \langle \mathbf{f}(b) - \mathbf{f}(a), \mathbf{f}(b) - \mathbf{f}(a) \rangle \\
&= \langle \mathbf{f}(b) - \mathbf{f}(a), \int_a^b \mathbf{f}'(t)\, dt \rangle \\
&= \int_a^b \langle \mathbf{f}(b) - \mathbf{f}(a), \mathbf{f}'(t) \rangle\, dt \\
&\leq \int_a^b ||\mathbf{f}(b) - \mathbf{f}(a)||_2 ||\mathbf{f}'(t)||_2\, dt = ||\mathbf{f}(b) - \mathbf{f}(a)||_2 \int_a^b ||\mathbf{f}'(t)||_2\, dt.
\end{aligned}$$



Exemples.

1. Longueur d'arc d'un cercle.

$$L(\alpha) = \int_0^\alpha ||\mathbf{f}'(t)||_2 \, dt = r\alpha$$

2. Longueur d'arc d'une ligne droite.

$$L = \int_a^b ||\mathbf{v}||_2 \, dt = ||\mathbf{v}||_2(b-a)$$

3. La longueur d'arc d'une fonction $f \in C^1([a,b])$ est donnée par

$$L = \int_a^b \sqrt{1+f'(t)^2}\, dt$$

Calculons la longueur d'arc de la fonction $\cosh x$ entre $a < b$. Evidemment

$$L = \int_a^b \sqrt{1+\sinh(t)^2}\, dt = \int_a^b \cosh t \, dt = \sinh b - \sinh a.$$

4. Longueur d'arc de l'hélice.

$$L = \int_0^b \sqrt{r^2+c^2}\, dt = b\sqrt{r^2+c^2}$$

Intégrale curviligne. Soit $\mathbf{f}$ une courbe de classe $C^1([a,b])$. Pour toute fonction continue $\phi : \mathbb{R}^n \to \mathbb{R}$ on peut définir

$$\int_{\mathrm{Im}(\mathbf{f})} \phi \, ds = \int_a^b \phi(\mathbf{f}(t))||\mathbf{f}'(t)||_2 \, dt. \tag{2.4}$$

C'est un sujet d'Analyse 3.

Chapitre 3

Fonctions réelles sur $\mathbb{R}^n$

3.1 Introduction

Définition. Une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est appelée fonction réelle sur $\mathbb{R}^n$. Pour un réel $c \in \mathrm{Im}(f)$ on appelle l'ensemble

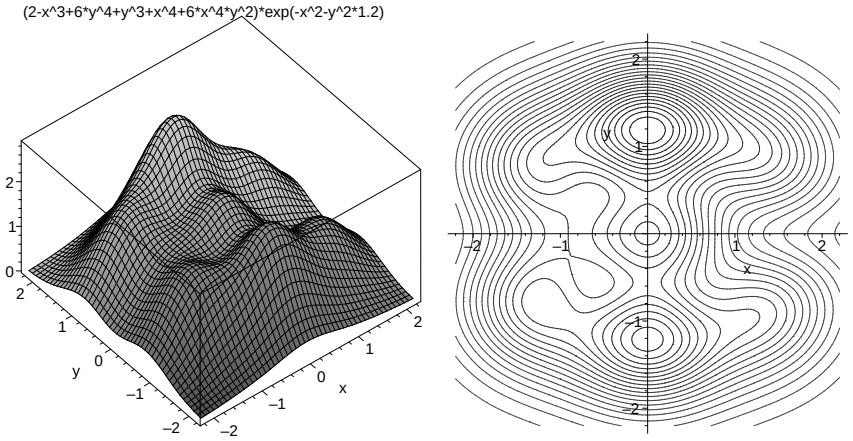
$$N_f(c) = \{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = c\}$$

l'ensemble de niveau c de f. Le graphe de f est donné par

$$G_f = \{\begin{pmatrix} \mathbf{x} \\ f(\mathbf{x}) \end{pmatrix} : \mathbf{x} \in \mathbb{R}^n\} \subset \mathbb{R}^{n+1}.$$

Le graphe décrit une hypersurface d'équation $x_{n+1} = f(\mathbf{x})$ dans $\mathbb{R}^{n+1}$.

Le cas $n = 2$. Pour des fonctions $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ on appelle l'ensemble $N_f(c)$ aussi ligne de niveau. On peut considerer la valeur $f(x_1, x_2)$ comme altitude du point (x_1, x_2). Donc $N_f(c)$ correspond à la ligne de niveau c sur une carte géographique. On peut donc représenter graphiquement une fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ par son graphe dans $\mathbb{R}^3$ ou par la projection de ses lignes de niveaux sur le plan $\mathbb{R}^2$.



Exemples. En physique les fonctions $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ sont souvent appelées champs scalaires. Le potentiel gravitationnel d'une masse ou le potentiel d'une charge éléctrique sont des exemples de champs scalaires :

$$\phi : \mathbb{R}^3 \setminus \{\mathbf{0}\} \longrightarrow \mathbb{R}, \quad \phi(\mathbf{x}) = \frac{k}{||\mathbf{x}||_2}$$

pour une constante réelle k. En mécanique on considère souvent des systèmes dont l'energie est conservée (systèmes hamiltoniens). Pour le mouvement d'une particule de masse m dans l'espace soumise au potentiel $V(\mathbf{x})$ son energie est une fonction réelle de sa quantité du mouvement $\mathbf{p} = m\mathbf{v}$ où $\mathbf{v}$ dénote sa vitesse et $\mathbf{x}$ sa position dans l'espace :

$$E : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}, \quad E(\mathbf{p}, \mathbf{x}) = \frac{||\mathbf{p}||_2^2}{2m} + V(\mathbf{x}).$$

Le mouvement suit les lignes de niveau de l'energie E.

3.2 Limites et continuité d'une fonction réelle

Définition - limite d'une fonction. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction réelle sur $\mathbb{R}^n$. On dit que f admet pour limite $L \in \mathbb{R}$ lorsque $\mathbf{x}$ tend vers $\mathbf{a}$ si pour tout $\epsilon > 0$ il existe un $\delta > 0$ tel que $||\mathbf{x} - \mathbf{a}||_2 < \delta$ implique $|f(\mathbf{x}) - L| < \epsilon$.

Fonction continue. Une fonction f est continue en $\mathbf{a}$ si pour tout $\epsilon > 0$ il existe un $\delta > 0$ tel que $||\mathbf{x} - \mathbf{a}||_2 < \delta)$ implique $|f(\mathbf{x}) - f(\mathbf{a})| < \epsilon$.

Exemple. La norme $|| \cdot ||_2 : \mathbb{R}^n \longrightarrow \mathbb{R}$ est une fonction continue en tout $\mathbf{a} \in \mathbb{R}^n$: Pour tout $\epsilon > 0$ on choisit $\delta = \epsilon$ car

$$|||\mathbf{x}||_2 - ||\mathbf{a}||_2| \leq ||\mathbf{x} - \mathbf{a}||_2.$$

Inexistence d'une limite I. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par

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

$f(x, y)$ n'est pas continue en $(0, 0)$ et, de plus

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

n'existe pas. Une métode simple pour démontrer que la limite n'existe pas est d'étudier le comportement de la fonction f en s'approchant du point $\mathbf{a} = (0, 0)$ sur des lignes droites passant par $(0, 0)$. Par exemple, on considère la ligne droite donnée par $C = \{(x, y) : y = x\}$. Pour tout point $(x, y) \in C \setminus \{(0, 0)\}$ on a

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

Donc

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

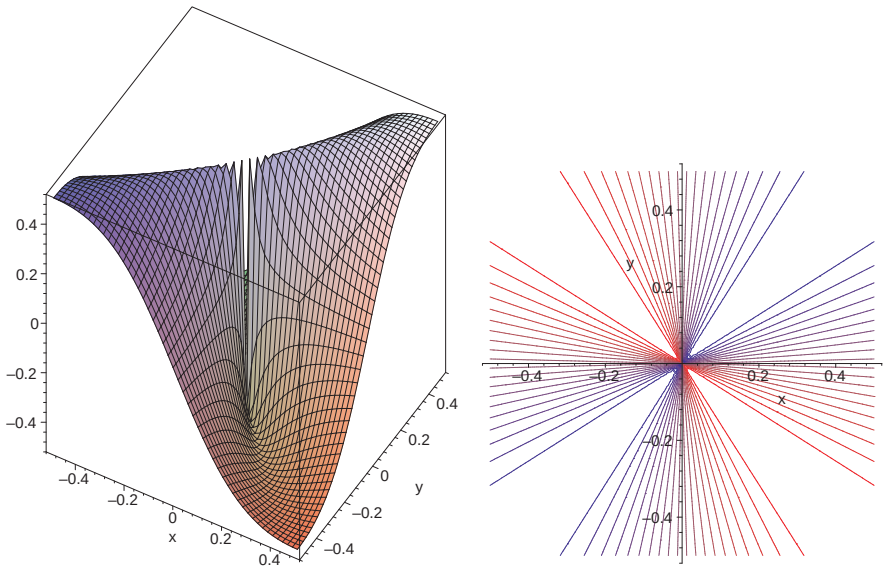
Sur la ligne droite $D = \{(x, y) : y = 0\}$ on a pour tout point $(x, y) \in D\backslash\{(0, 0)\}$:

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

et donc

$$\lim_{\substack{(x,y) \to (0,0) \\ (x,y) \in D}} f(x, y) = \lim_{x \to 0} f(x, 0) = 0.$$

On conclut que $\lim_{(x,y)\to(0,0)} f(x, y)$ n'existe pas.



Le graphe et les ensembles de niveau de la fonction

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

Noter qu'on ne voit pas la discontinuité de la fonction car le programme dessine toujours des segments entre des points calculés. Une conséquence est l'apparence des sommets "bizarres".

Version paramétrée. Plus généralement, pour une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ on considère les droites données par $\mathbf{v}t + \mathbf{a}$, $\mathbf{v} \in S_{n-1} = \{\mathbf{v} \in \mathbb{R}^n : ||\mathbf{v}||_2 = 1\}$ et on étudie les fonctions à une variable t données par

$$g_{\mathbf{v}}(t) = f(\mathbf{v}t + \mathbf{a})$$

lorsque t tend vers 0. Si on trouve deux vecteurs $\mathbf{v}, \mathbf{w} \in S_{n-1}$ tels que $g_{\mathbf{v}}(t)$ et $g_{\mathbf{w}}(t)$ n'ont pas le même comportement lorsque t tend vers 0, alors la limite de f lorsque $\mathbf{x}$ tend vers $\mathbf{a}$ n'existe pas. La notion de la limite d'une fonction décrit son comportement dans un voisinage ; c'est pourquoi il ne suffit pas d'étudier les fontions $g_{\mathbf{v}}(t)$ pour tout $\mathbf{v} \in S_{n-1}$ comme le démontre l'exemple suivant :

Inexistence d'une limite II. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ donné par

$$f(x, y) = \begin{cases} 1, & \text{si } y = x^2 \text{ et } x > 0\,; \\ 0, & \text{sinon.} \end{cases} \tag{3.1}$$

Évidemment f n'est pas continue en $(0,0)$ puisque pour tout voisinage de $(0,0)$: $f(x,x^2)-f(0,0)=1$ et $x>0$ suffisamment petit. D'autre part $g_{\mathbf{v}}(t)=0$ pour tout $\mathbf{v}\in S_1$ et t suffisamment petit.

3.3 Dérivées partielles, différentielle et fonctions de la classe C^1

3.3.1 Dérivées partielles

Définition. Une fonction $f:\mathbb{R}^n\longrightarrow\mathbb{R}$ est dite partiellement différentiable par rapport à x_k au point $\mathbf{a}\in\mathbb{R}^n$ si la limite

$$\frac{\partial f}{\partial x_k}(\mathbf{a})=\lim_{\substack{h\to 0\\ h\neq 0}}\frac{f(\mathbf{a}+h\mathbf{e}_k)-f(\mathbf{a})}{h}$$

existe. La limite $\frac{\partial f}{\partial x_k}(\mathbf{a})$ est appelée dérivée partielle de f par rapport à x_k au point $\mathbf{a}\in\mathbb{R}^n$.

Remarque. On note aussi

$$D_kf(\mathbf{a})=\frac{\partial f}{\partial x_k}(\mathbf{a}).$$

Remarque. La limite $\frac{\partial f}{\partial x_k}(\mathbf{a})$ existe si et seulement si la fonction $t\mapsto f(\mathbf{a}+t\mathbf{e}_k)$ est dérivable en $t=0$. C'est pourquoi nous avons pour les dérivées partielles comme pour la dérivée des fonctions à une variable (voir Analyse I) les propriétés de linéarité et les règles du produit et du quotient :

$$\frac{\partial(\alpha f+\beta g)}{\partial x_k}(\mathbf{a})=\alpha\frac{\partial f}{\partial x_k}(\mathbf{a})+\beta\frac{\partial g}{\partial x_k}(\mathbf{a}) \tag{3.2}$$

$$\frac{\partial(f\cdot g)}{\partial x_k}(\mathbf{a})=g(\mathbf{a})\frac{\partial f}{\partial x_k}(\mathbf{a})+f(\mathbf{a})\frac{\partial g}{\partial x_k}(\mathbf{a}) \tag{3.3}$$

$$\frac{\partial(f/g)}{\partial x_k}(\mathbf{a})=\big(g(\mathbf{a})\frac{\partial f}{\partial x_k}(\mathbf{a})-f(\mathbf{a})\frac{\partial g}{\partial x_k}(\mathbf{a})\big)/g(\mathbf{a})^2. \tag{3.4}$$

Pour la règle de composition voir ch. 3.5.

Définition. Soit $U\subset\mathbb{R}^n$. Une fonction $f:U\longrightarrow\mathbb{R}$ est dite partiellement différentiable dans U si les n dérivées partielles $D_kf(\mathbf{x})$, $k=1,\dots,n$ existent pour tout $\mathbf{x}\in\mathbb{R}^n$. Le vecteur

$$\nabla f(\mathbf{x})=\begin{pmatrix}D_1f(\mathbf{x})\\ \cdot\\ \cdot\\ \cdot\\ D_nf(\mathbf{x})\end{pmatrix}=\begin{pmatrix}\frac{\partial f}{\partial x_1}(\mathbf{x})\\ \cdot\\ \cdot\\ \cdot\\ \frac{\partial f}{\partial x_n}(\mathbf{x})\end{pmatrix}$$

est appelé le gradient de f (au point $\mathbf{x}$).

Remarque. Le gradient est aussi noté par $\mathbf{grad} f(\mathbf{x})$. Noter aussi l'identité

$$\nabla f(\mathbf{x}) = \sum_{k=1}^{n} D_k f(\mathbf{x})\ \mathbf{e_k}$$

donc $D_k f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{e_k} \rangle$.

Exemples.

1. Soit $\mathbf{v} \in \mathbb{R}^n$ fixe et $l : \mathbb{R}^n \longrightarrow \mathbb{R}$ la forme linéaire définie par

   $$l(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle.$$

   La fonction $l(\mathbf{x})$ est partiellement différentiable et

   $$\frac{\partial l}{\partial x_k}(\mathbf{x}) = \lim_{h \to 0} \frac{l(\mathbf{x} + h\mathbf{e}_k) - l(\mathbf{x})}{h} = \langle \mathbf{v}, \mathbf{e_k} \rangle = v_k.$$

   Donc

   $$\nabla l(\mathbf{x}) = \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{e_k} \rangle\ \mathbf{e_k} = \mathbf{v}$$

2. Soit $A \in M_{n,n}(\mathbb{R})$ une matrice symétrique et $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ la forme quadratique définie par

   $$q(\mathbf{x}) = \frac{1}{2} \langle A\mathbf{x}, \mathbf{x} \rangle.$$

   La fonction $q(\mathbf{x})$ est partiellement différentiable et

   $$\frac{\partial q}{\partial x_k}(\mathbf{x}) = \lim_{h \to 0} \frac{q(\mathbf{x} + h\mathbf{e}_k) - q(\mathbf{x})}{h} = \langle A\mathbf{x}, \mathbf{e_k} \rangle.$$

   Donc

   $$\nabla q(\mathbf{x}) = \sum_{k=1}^{n} \langle A\mathbf{x}, \mathbf{e_k} \rangle\ \mathbf{e_k} = A\mathbf{x}$$

3. Considérons la fonction $r : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par $r(\mathbf{x}) = ||\mathbf{x}||_2$. Elle est partiellement différentiable sur l'ensemble $\mathbb{R}^n \setminus \{\mathbf{0}\}$ avec

   $$\frac{\partial r}{\partial x_k}(\mathbf{x}) = \frac{x_k}{r(\mathbf{x})}.$$

   Par conséquent, son gradient est

   $$\nabla r(\mathbf{x}) = \frac{\mathbf{x}}{r(\mathbf{x})} = \frac{\mathbf{x}}{r}.$$

   Si $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$ est dérivable, alors la fonction composeé $f(r) = f(r(\mathbf{x}))$ est partiellement différentiable sur $\mathbb{R}^n \setminus \{\mathbf{0}\}$ et

   $$\nabla f(r) = \frac{f'(r)\mathbf{x}}{r}$$

   où f' dénote la dérivée de f par rapport à la variable r.

4. On va montrer que la fonction $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par

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

est partiellement différentiable sur $\mathbb{R}^n$. D'abord notons que pour tout $(x,y) \neq (0,0)$ on a

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

et par symétrie

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

f est aussi partiellement différentiable en $(0,0)$ car

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

et

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

Noter que f n'est pas continue en $(0,0)$ (voir 3.2). Donc contrairement au cas $n = 1$ l'existence des dérivées partielles n'implique pas la continutité d'une fonction. Il faut, en plus, la continuité des dérivées partielles.

Proposition - la continuité des dérivées partielles implique la continuité de la fonction. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction dont les n dérivées partielles existent et sont continues en $\mathbf{a} \in \mathbb{R}^n$. Alors, la fonction f est continue en $\mathbf{a}$.

Démonstration. Pour $\mathbf{h} \in \mathbb{R}^n$ soit $\mathbf{a}_k = \mathbf{a} + \sum_{j=1}^{k} h_j \mathbf{e}_j$, $k = 0, \ldots, n$. Alors, $\mathbf{a}_0 = \mathbf{a}$, $\mathbf{a}_n = \mathbf{a} + \mathbf{h}$ et $\mathbf{a}_k - \mathbf{a}_{k-1} = h_k \mathbf{e}_k$, $k = 1, \ldots, n$. On écrit :

$$f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) = \sum_{k=1}^{n} f(\mathbf{a}_k) - f(\mathbf{a}_{k-1})$$

Par le théorème des accroissement finis pour des fonctions à une variable pour tout $k = 1, \ldots, n$ il existe $c_k \in \mathbb{R}$, $|c_k| \leq |h_k|$ tel que

$$f(\mathbf{a}_k) - f(\mathbf{a}_{k-1}) = h_k \frac{\partial f}{\partial x_k}(\mathbf{a}_{k-1} + c_k \mathbf{e}_k). \tag{3.5}$$

Par conséquent,

$$\begin{aligned} f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) &= \sum_{k=1}^{n} h_k \frac{\partial f}{\partial x_k}(\mathbf{a}_{k-1} + c_k \mathbf{e}_k) \\ &= \sum_{k=1}^{n} h_k \big(\frac{\partial f}{\partial x_k}(\mathbf{a}_{k-1} + c_k \mathbf{e}_k) - \frac{\partial f}{\partial x_k}(\mathbf{a}) \big) + \langle \nabla f(\mathbf{a}), \mathbf{h} \rangle \end{aligned} \tag{3.6}$$

La continuité des dérivées partielles et de la forme linéaire $\langle \nabla f(\mathbf{a}), \mathbf{h} \rangle$ implique $f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) \to 0$ lorsque $\mathbf{h} \to \mathbf{0}$.

3.3.2 Fonctions différentiables et différentielle

Définition. Une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est dite différentiable en un point $\mathbf{a} \in \mathbb{R}^n$ si et seulement s'il existe une forme linéaire $d_{\mathbf{a}}f(\mathbf{x}) := \langle d_{\mathbf{a}}f, \mathbf{x}\rangle$, appelée différentielle de f au point $\mathbf{a}$, telle que

$$\lim_{\substack{\mathbf{h}\to\mathbf{0}\\ \mathbf{h}\neq\mathbf{0}}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - d_{\mathbf{a}}f(\mathbf{h})}{||\mathbf{h}||_2} = \mathbf{0}.$$

La différentielle $d_{\mathbf{a}}f$ est aussi appelée la dérivée total de f en $\mathbf{a}$. f est dite différentiable dans $U \subset \mathbb{R}^n$ ouvert, si f différentiable en tout point $\mathbf{a} \in U$.

Dérivées directionnelles. Soit $\mathbf{a}, \mathbf{v} \in \mathbb{R}^n$, $\mathbf{v} \neq \mathbf{0}$ $\mathbf{k} : \mathbb{R} \longrightarrow \mathbb{R}^n$ la ligne droite donnée par $\mathbf{k}(t) = \mathbf{a} + \mathbf{v}t$ La dérivée

$$\frac{d}{d\,t} f(\mathbf{a}+\mathbf{v}t)\bigg|_{t=0}$$

est appelée la dérivée directionnelle de f en $\mathbf{a}$ suivant le vecteur $\mathbf{v}$. Si f est différentiable en $\mathbf{a}$, alors

$$\frac{d}{d\,t} f(\mathbf{a}+\mathbf{v}t)\bigg|_{t=0} = \langle d_{\mathbf{a}}f, \mathbf{v}\rangle.$$

En prenant les directions $\mathbf{e}_k$ on voit que tout fonction différentiable en $\mathbf{a}$ est partiellement différentiable en $\mathbf{a}$. Il en suit que pour tout fonction différentiable en $\mathbf{a}$:

$$d_{\mathbf{a}}f(\mathbf{v}) = \langle d_{\mathbf{a}}f, \mathbf{v}\rangle = \langle \nabla f(\mathbf{a}), \mathbf{v}\rangle, \tag{3.7}$$

donc

$$\frac{d}{d\,t} f(\mathbf{a}+\mathbf{v}t)\bigg|_{t=0} = \langle \nabla f(\mathbf{a}), \mathbf{v}\rangle.$$

Il en suit également le développement limité d'ordre 1 :

Théorème. Soit $U \subset \mathbb{R}^n$ ouvert et$f : U \longrightarrow \mathbb{R}$ différentiable en $\mathbf{a} \in U$. Alors,

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{h}\rangle + o(\mathbf{h}) \tag{3.8}$$

et

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{o(\mathbf{h})}{||\mathbf{h}||_2} = 0.$$

En particulier, f est continue en $\mathbf{a}$.

Resumé. Une fonction différentiable en $\mathbf{a}$ est partiellement différentiable et continue en ce point. La différentielle $d_{\mathbf{a}}f$ de f en $\mathbf{a}$ est la forme linéaire $\langle \nabla f(\mathbf{a}), \mathbf{h}\rangle$.

3.3.3 Fonctions de la classe C^1

Définition. Soit $U \subset \mathbb{R}^n$ un ensemble ouvert. Une fonction $f : U \longrightarrow \mathbb{R}$ est dite de classe $C^1(U)$ si les n dérivées partielles existent et sont continues en tout $\mathbf{x} \in U$.

Remarque. Comme la continuité des dérivées partielles en un point $\mathbf{a}$ implique la continuité de la fonction en ce point $\mathbf{a}$ elle implique également que f est différentiable en $\mathbf{a}$ (voir éq. (3.6)). Il en suit que toute fonction de classe $C^1(U)$ est différentiable (et donc continue) sur U :

Proposition. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \longrightarrow \mathbb{R}$ de classe $C^1(U)$. Alors, f est différentiable en tout $\mathbf{a} \in U$. En particulier, elle admet le développement du premier ordre éq (3.8) en tout $\mathbf{a} \in U$.

Théorème de la valeur moyenne. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \longrightarrow \mathbb{R}$ de classe $C^1(U)$. Soient $\mathbf{x}, \mathbf{y} \in U$ tels que le segment $\mathbf{k}(t) := (1-t)\mathbf{x}+t\mathbf{y}$, $0 \leq t \leq 1$ est dans U. Alors,

$$f(\mathbf{y}) - f(\mathbf{x}) = \langle \int_0^1 \nabla f((1-t)\mathbf{x} + t\mathbf{y})\, dt, \mathbf{y} - t\mathbf{x} \rangle. \tag{3.9}$$

Démonstration. On définit $g : [0,1] \to \mathbb{R}$ par $g(t) := f((1-t)\mathbf{x} + t\mathbf{y})$. Alors, g est de classe C^1 tel que $g(0) = f(\mathbf{x})$, $g(1) = f(\mathbf{x})$ et $g'(t) = \langle \nabla f((1-t)\mathbf{x} + t\mathbf{y})\, dt, \mathbf{y} - t\mathbf{x} \rangle$. L'éq (3.9) est équivalente à

$$g(1) - g(0) = \int_0^1 g'(t)\, dt.$$

3.4 Hyperplan tangent

Pour des fonctions différentiables (nous prenons toujours des fonctions de classe C^1 pour simplifier la présentation) l'interprétation géométrique du terme linéaire $f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{h} \rangle$ est celle de l'hyperplan tangent en généralisant la notion de la tangente pour des fonctions définies sur $\mathbb{R}$.

Exemple . Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction de classe C^1 donnée par

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

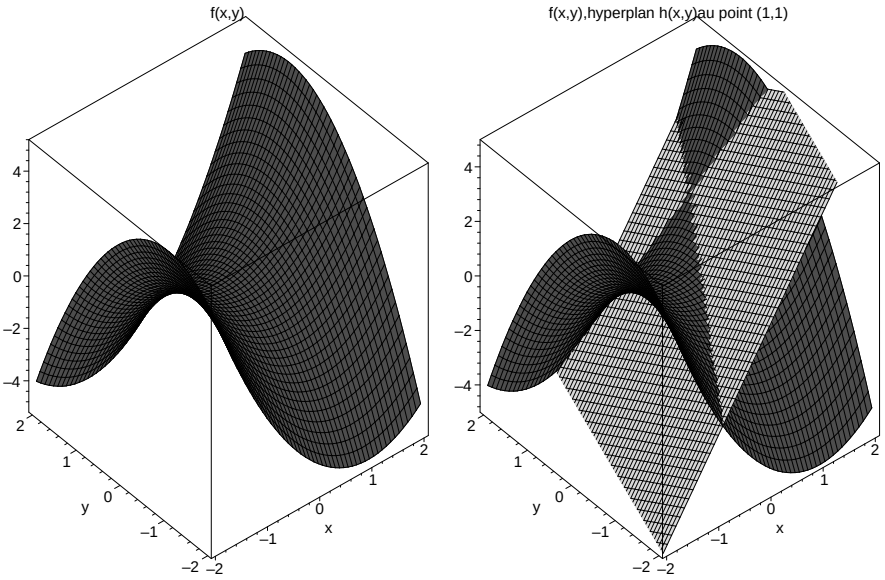
On a

$$\nabla f(x, y) = \begin{pmatrix} 2x + y \\ -2y + x \end{pmatrix}$$

Dans le voisinage du point $(a, b) = (1, 1)$ la fonction se comporte comme

$$\begin{aligned} f(x, y) &\approx f(a, b) + \langle \nabla f(a, b), (x - a, y - b) \rangle \\ &= f(a, b) + (2a + b)(x - a) + (-2b + a)(y - b) \\ &= 1 + 3(x - 1) - (y - 1) \end{aligned}$$

La fonction $h(x, y) = 1 + 3(x-1) - (y-1)$ décrit un plan dans $\mathbb{R}^3$, i.e. le graphe G_h donne un plan dans $\mathbb{R}^3$. On a $(a, b, f(a, b)) = (a, b, h(a, b))$ donc c'est un plan passant par le point $(a, b, f(a, b))$ du graphe de f, tangent au graphe de f. Ce plan constitue une généralisation de la droite tangente pour les fonctions réelles d'une variable. Si l'on généralise au cas de $\mathbb{R}^n$ on l'appelle l'hyperplan tangent.



Le graphe de la fonction $f(x, y) = x^2 + xy - y^2$ et son hyperplan tangent au point $(1, 1, 1) \in \mathbb{R}^3$.

Équation d'un hyperplan. Soit $\mathbf{v} \in \mathbb{R}^n$ et $l : \mathbb{R}^n \longrightarrow \mathbb{R}$ la forme linéaire définie par

$$l(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle.$$

Pour tout $\mathbf{a} \in \mathbb{R}^n$ et $h \in \mathbb{R}$ le graphe de la fonction

$$h_{\mathbf{a}}(\mathbf{x}) = h + l(\mathbf{x} - \mathbf{a}) = h + \langle \mathbf{v}, \mathbf{x} - \mathbf{a} \rangle$$

définit un hyperplan passant par le point $(\mathbf{a}, h) \in \mathbb{R}^{n+1}$. L'équation de cet hyperplan dans $\mathbb{R}^{n+1}$ est alors donné par

$$x_{n+1} = h + \langle \mathbf{v}, \mathbf{x} - \mathbf{a} \rangle = h + \sum_{k=1}^{n} v_k (x_k - a_k).$$

Hyperplan tangent. Si $f : U \longrightarrow \mathbb{R}$ est de classe $C^1(U)$, alors pour tout $\mathbf{a} \in U$ il existe un voisinage $V \subset U$ de $\mathbf{a}$ telle que

$$f(\mathbf{x}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle + o(\mathbf{x} - \mathbf{a}).$$

Par conséquent, l'équation de l'hyperplan tangent est donné par

$$x_{n+1} = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle.$$

Si $n = 1$ on retrouve l'équation de la tangente $x_2 = f(a) + f'(a)(x_1 - a)$ vue en Analyse 1.

3.5 Gradient et lignes de niveau

La règle de composition. Soit I un intervalle $\mathbf{k} : I \longrightarrow \mathbb{R}^n$ une courbe de classe C^1. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction réelle de classe C^1. Alors, la fonction

composée $f \circ k : I \longrightarrow \mathbb{R}$ est de la classe C^1 et

$$\frac{d}{dt} f(\mathbf{k}(t)) = \langle \nabla f(\mathbf{k}(t)), \mathbf{k}'(t) \rangle.$$

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

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

et $\mathbf{k} : \mathbb{R} \longrightarrow \mathbb{R}^2$ la spirale logarithmique donnée par

$$\mathbf{k}(t) = (e^{-ct} \cos t, e^{-ct} \sin t), \quad c > 0.$$

Avec

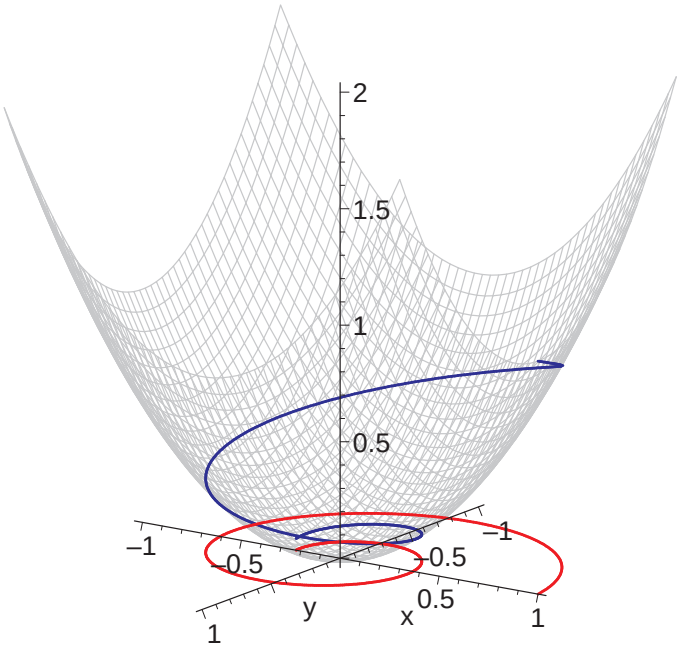
$$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}$$

et

$$\mathbf{k}'(t) = \begin{pmatrix} -ce^{-ct} \cos t - e^{-ct} \sin t \\ -ce^{-ct} \sin t + e^{-ct} \cos t \end{pmatrix}$$

on trouve

$$\begin{aligned} \frac{d}{dt} f(\mathbf{k}(t)) &= 2\mathbf{k}_1(t)\mathbf{k}_1'(t) + 2\mathbf{k}_2(t)\mathbf{k}_2'(t) \\ &= -2ce^{-2ct} \end{aligned}$$



La courbe $\mathbf{k}(t)$ dans $\mathbb{R}^2$ (en rouge) et la courbe $\mathbf{c}(t) = \begin{pmatrix} \mathbf{k}(t) \\ f(\mathbf{k}(t)) \end{pmatrix}$ (en bleu) est une courbe dans le graphe $G_f \subset \mathbb{R}^3$. La dérivée $\frac{d}{dt} f(\mathbf{k}(t))$ signifie la variation de l'altitude avec t ou plus précisement la pente de la fonction f en suivant de la courbe $\mathbf{k}(t)$.

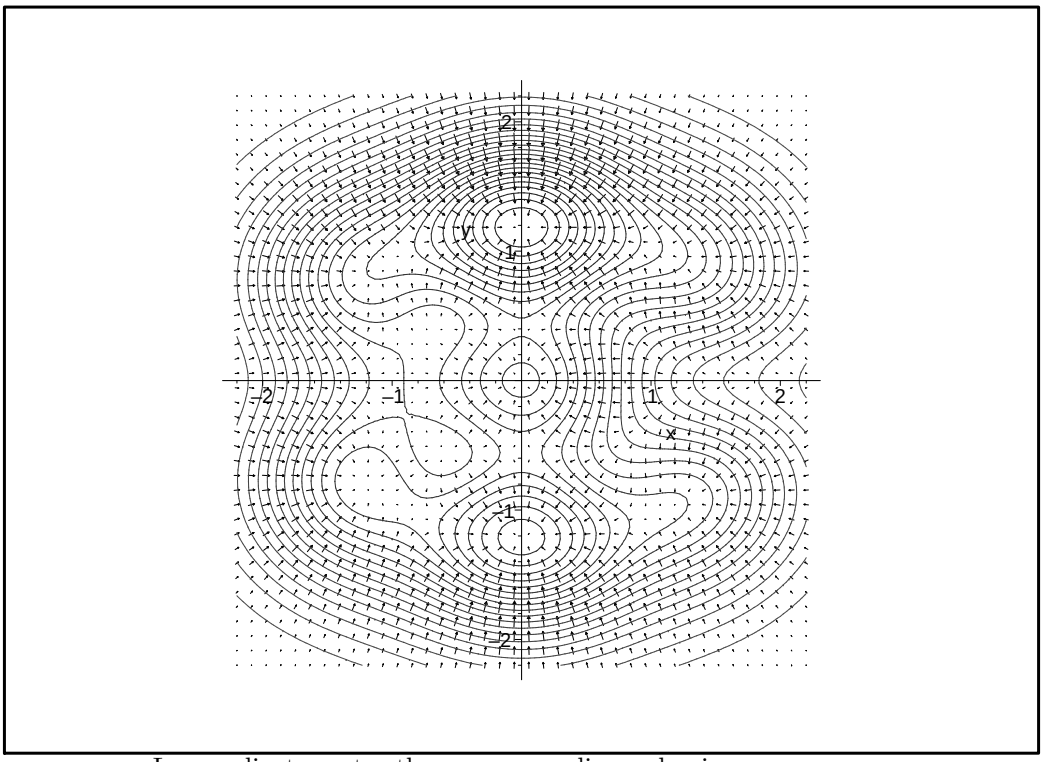
Courbes de niveau. Si $\mathbf{k}$ est une courbe telle que

$$f(\mathbf{k}(t)) = c = \text{const}$$

i.e. $\mathbf{k}$ est une courbe dans l'ensemble de niveau c de f on a

$$\frac{d}{d\,t} f(\mathbf{k}(t)) = \langle \nabla f(\mathbf{k}(t)), \mathbf{k}'(t) \rangle = \frac{d}{d\,t}\, c = 0.$$

On note que le gradient de f est orthogonal au vecteur tangent d'une courbe de niveau, i.e. d'une courbe dans l'ensemble de niveau $N_f(c)$.



Les gradients sont orthogonaux aux lignes de niveau.

Exemple. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction de classe C^1 donnée par

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

On a

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix}.$$

Les ensembles de niveau c de f pour $c > 0$ sont des cercles de centre $(0, 0)$ et de rayon $\sqrt{c}$. On peut les representer sous forme paramétrisée par des courbes $\mathbf{k}_c : [0, 2\pi] \longrightarrow \mathbb{R}^2$ données par

$$\mathbf{k}_c(t) = \begin{pmatrix} \sqrt{c}\, \cos t \\ \sqrt{c}\, \sin t \end{pmatrix}$$

Une interprétation géométrique du gradient. On peut interpreter la dérivée

$$\frac{d}{dt}f(\mathbf{k}(t))$$

comme pente de la fonction f en suivant la courbe $\mathbf{k}(t)$. Par la règle de composition on a

$$\frac{d}{dt}f(\mathbf{k}(t)) = \langle \nabla f(\mathbf{k}(t)), \mathbf{k}'(t)\rangle = ||\nabla f(\mathbf{k}(t))||_2||\mathbf{k}'(t)||_2 \cos\alpha(t)$$

où $\alpha(t)$ signifie l'angle entre le gradient de f et le vecteur tangent de la courbe $\mathbf{k}$ au temps t. On voit que cette valeur est maximale si $\alpha(t) = 0$. Donc le vecteur $\nabla f(\mathbf{a})$ dans un point $\mathbf{a}$ point dans la direction de la pente la plus forte de f.

La règle de composition - 2. Soit $D \subset \mathbb{R}^n$ ouvert. Soient $\rho : D \longrightarrow \mathbb{R}$ une fonction de classe C^1 et $f : \mathbb{R} \longrightarrow \mathbb{R}$ de classe C^1 avec fonction dérivée f'. Alors, la fonction composée $f \circ \rho : D \longrightarrow \mathbb{R}$ est de la classe C^1 et

$$\nabla f(\rho(\mathbf{x})) = f'(\rho(\mathbf{x}))\nabla\rho(\mathbf{x}).$$

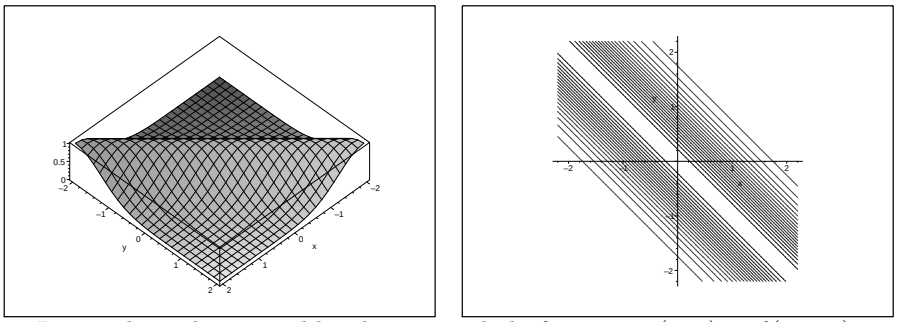
Exemple. Voir l'exemple 3, p.22 : si $f = f(r)$ dépend uniquement de la distance radiale on a pour tout $(\mathbf{x}) \neq (\mathbf{0})$:

$$\nabla f(r) = \frac{f'(r)\mathbf{x}}{r}.$$

Exemple. Pour trouver des solutions de l'équation aux dérivées partielles

$$\frac{\partial u}{\partial x}(x,y) - \frac{\partial u}{\partial y}(x,y) = 0$$

on note que $u(x,y) = f(x+y)$ (donc $\rho(x,y) = x+y$) pour toute f dérivable donne une solution.



Le graphe et les ensembles de niveau de la fonction $u(x,y) = f(x+y)$.

3.6 Dérivées partielles d'ordre supérieur

Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \longrightarrow \mathbb{R}$ une fonction partiellement différentiable. Les dérivées partielles $D_k f : U \longrightarrow \mathbb{R}$ peuvent elles-mêmes être partiellement différentiables. Dans ce cas on dit que f est deux fois partiellement différentiable. Donc les dérivées partielles d'ordre deux $D_j D_k f$ existent.

Exemple. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ la fonction réelle définie par

$$f(x, y) = x^2 \sin y + ye^x$$

Évidemment f est de classe $C^1(\mathbb{R}^2)$ et

$$D_x f(x, y) = 2x \sin y + ye^x, \quad D_y f(x, y) = x^2 \cos y + e^x$$

Les dérivées partielles sont des fonctions réelles de classe $C^1(\mathbb{R}^2)$ et on peut calculer les quatre dérivées partielles

$$D_x(D_x f(x, y)) = D_{xx} f(x, y), \quad D_y(D_x f(x, y)) = D_{yx} f(x, y)$$

$$D_x(D_y f(x, y)) = D_{xy} f(x, y), \quad D_y(D_y f(x, y)) = D_{yy} f(x, y)$$

Nous avons

$$D_{xx} f(x, y) = 2 \sin y + ye^x, \quad D_{yx} f(x, y) = 2x \cos y + e^x$$

$$D_{xy} f(x, y) = 2x \cos y + e^x, \quad D_{yy} f(x, y) = -x^2 \sin y$$

Noter que $D_{xy} f(x, y) = D_{yx} f(x, y)$. C'est une propriété générale de dérivées partielles sous certaines hypothèses.

Définition. Soit $U \subset \mathbb{R}^n$ un ensemble ouvert. Une fonction $f : U \longrightarrow \mathbb{R}$ est dite de classe $C^m(U)$ si toutes ses dérivées partielles d'ordre m existent et sont continues en tout $\mathbf{x} \in U$. La fonction f est dite de classe $C^\infty(U)$ si toutes ses dérivées partielles successives existent et sont continues en tout $\mathbf{x} \in U$.

Théorème. Soit $U \subset \mathbb{R}^n$ un ensemble ouvert et $f : U \longrightarrow \mathbb{R}$ de classe $C^2(U)$. Alors, pour tout $1 \leq j, k \leq n$ et tout $\mathbf{x} \in U$:

$$D_{jk} f(\mathbf{x}) = D_{kj} f(\mathbf{x}).$$

Remarque. Ce résultat se généralise aux dérivées partielles d'ordre ≥ 2. Par exemple, si $f : U \longrightarrow \mathbb{R}$ de classe $C^3(U)$ on a

$$D_{jkl} f(\mathbf{x}) = D_{kjl} f(\mathbf{x}) = D_{jlk} f(\mathbf{x}) = \ldots \text{etc.}$$

Il y a $\binom{n+m-1}{m}$ dérivées partielles d'ordre m (au lieu de n^m).

Définition. La matrice

$$(D_{jk} f(\mathbf{x}))_{j,k} = \begin{pmatrix} D_{11} f(\mathbf{x}) & \cdots & D_{n1} f(\mathbf{x}) \\ \cdots & & \cdots \\ D_{1n} f(\mathbf{x}) & \cdots & D_{nn} f(\mathbf{x}) \end{pmatrix}$$

est appelée la matrice hessienne de f et notée $\mathrm{Hess}\,(f)(\mathbf{x})$.

Remarque. Noter que

$$\text{Hess}\,(f)(\mathbf{x}) = \begin{pmatrix} D_1\nabla f(\mathbf{x}) & \cdots & D_n\nabla f(\mathbf{x}) \end{pmatrix}.$$

Définition. Soit $U \subset \mathbb{R}^n$ un ensemble ouvert et $f : U \longrightarrow \mathbb{R}$ de classe $C^2(U)$. Alors, la fonction $\Delta f : U \longrightarrow \mathbb{R}$ définie par

$$\Delta f(\mathbf{x}) = \sum_{k=1}^{n} D_{kk} f(\mathbf{x})$$

est appelée le laplacien de la fonction f. On appelle le symbole Δ le laplacien.

Remarque. Noter que

$$\Delta f(\mathbf{x}) = \text{tr}\,(\text{Hess}\,(f)(\mathbf{x}))$$

où $\text{tr}\,(A)$ dénote la trace de la matrice A.

Exemples.

1. Pour $\mathbf{v} \in \mathbb{R}^n$ on considère la forme linéaire $l : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par

   $$l(\mathbf{x}) = \langle \mathbf{v}, \mathbf{x} \rangle.$$

   La fonction $l(\mathbf{x})$ est de classe $C^2(\mathbb{R}^n)$ (même de classe $C^\infty(\mathbb{R}^n)$). Par ch. 3.3

   $$\nabla l(\mathbf{x}) = \sum_{k=1}^{n} \langle \mathbf{v}, \mathbf{e_k} \rangle \, \mathbf{e_k} = \mathbf{v}$$

   Donc

   $$\text{Hess}\,(l)(\mathbf{x}) = 0$$

   en tout $\mathbf{x} \in \mathbb{R}^n$.

2. Soit $A \in M_{n,n}(\mathbb{R})$ une matrice symétrique et $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ la forme quadratique définie par

   $$q(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x} \rangle.$$

   est de classe $C^2(\mathbb{R}^n)$ (même de classe $C^\infty(\mathbb{R}^n)$). Par ch. 3.3

   $$\nabla q(\mathbf{x}) = A\mathbf{x}.$$

   Alors

   $$\text{Hess}\,(q)(\mathbf{x}) = A$$

   en tout $\mathbf{x} \in \mathbb{R}^n$.

3. Considérons la fonction $r : \mathbb{R}^n \longrightarrow \mathbb{R}$ définie par $r(\mathbf{x}) = ||\mathbf{x}||_2$. Elle est de classe $C^2(\mathbb{R}^n \setminus \{\mathbf{0}\})$ (même de classe $C^\infty(\mathbb{R}^n \setminus \{\mathbf{0}\})$). En utilisant

   $$\nabla r(\mathbf{x}) = \frac{\mathbf{x}}{r}$$

on trouve $(D_{jk}f(\mathbf{x}))_{j,k} = \frac{1}{r}\delta_{j,k} - \frac{x_j x_k}{r^3}$ où $\delta_{j,k} = 1$ si $j = k$ et $\delta_{j,k} = 0$ sinon, i.e.

$$\text{Hess}\,(r)(\mathbf{x}) = \begin{pmatrix} \frac{1}{r} - \frac{x_1^2}{r^3} & -\frac{x_1 x_2}{r^3} & \cdots & -\frac{x_1 x_n}{r^3} \\ -\frac{x_1 x_2}{r^3} & \frac{1}{r} - \frac{x_2^2}{r^3} & \cdots & -\frac{x_2 x_n}{r^3} \\ \cdots & & & \cdots \\ -\frac{x_1 x_n}{r^3} & -\frac{x_2 x_n}{r^3} & \cdots & \frac{1}{r} - \frac{x_n^2}{r^3} \end{pmatrix}$$

et

$$\Delta r = \Delta r(\mathbf{x}) = \sum_{k=1}^{n} \frac{1}{r} - \frac{x_k^2}{r^3} = \frac{n-1}{r}.$$

Si $f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R}$ est de classe $C^2(\mathbb{R}^n \setminus \{\mathbf{0}\})$, alors la fonction composeé $f(r) = f(r(\mathbf{x}))$ est de classe $C^2(\mathbb{R}^n \setminus \{\mathbf{0}\})$. Son gradient est donné par

$$\nabla f(r) = \frac{f'(r)\mathbf{x}}{r}$$

où f' dénote la dérivée de f par rapport à la variable r. La matrice hessienne de $f(r)$ est donnée par

$$(D_{jk}f(r))_{j,k} = \frac{f'(r)}{r}\delta_{j,k} + \frac{x_j x_k}{r}(\frac{f'(r)}{r})'$$

et en particulier

$$\Delta f(r) = \sum_{k=1}^{n} \frac{f'(r)}{r} + \frac{x_k^2}{r^2}(f''(r) - \frac{f'(r)}{r}) = f''(r) + \frac{n-1}{r}f'(r).$$

Chapitre 4

Champs vectoriels sur $\mathbb{R}^n$

4.1 Dérivées des Champs vectoriels

Définition. Soit $U \subset \mathbb{R}^n$. Une application $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ est appelée champ vectoriel sur U.

Définition. Soit $U \subset \mathbb{R}^n$ ouvert. Une application $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ est dite partiellement différentiable en $\mathbf{a} \in U$ si les dérivées partielles

$$D_j v_i(\mathbf{a}) = \frac{\partial v_i}{\partial x_j}(\mathbf{a}) := \lim_{h \to 0} \frac{v_i(\mathbf{a} + h\mathbf{e}_j) - v_i(\mathbf{a})}{h}, \quad 1 \leq i \leq m, 1 \leq j \leq n \quad (4.1)$$

existent. La matrice $m \times n$ $J_{\mathbf{v}}(\mathbf{a})$ définie par

$$(J_{\mathbf{v}}(\mathbf{a}))_{ij} = (D_j v_i(\mathbf{a}))_{i,j} = \begin{pmatrix} D_1 v_1(\mathbf{a}) & \cdots & D_n v_1(\mathbf{a}) \\ \cdots & & \cdots \\ D_1 v_m(\mathbf{a}) & \cdots & D_n v_m(\mathbf{a}) \end{pmatrix} \quad (4.2)$$

où $1 \leq i \leq m$, $1 \leq j \leq n$ est appelée la matrice jacobienne de $\mathbf{v}$ en $\mathbf{a} \in U$.

Définition. Soit $U \subset \mathbb{R}^n$ ouvert. Une application $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ est dite différentiable en $\mathbf{a} \in U$ s'il existe une application linéaire $D_{\mathbf{a}}\mathbf{v}(\mathbf{x}) := J_{\mathbf{v}}(\mathbf{a})\mathbf{x}$, $J_{\mathbf{v}}(\mathbf{a}) \in M_{m,n}$ telle que

$$\lim_{\substack{\mathbf{h} \to \mathbf{0} \\ \mathbf{h} \neq \mathbf{0}}} \frac{\mathbf{v}(\mathbf{a} + \mathbf{h}) - \mathbf{v}(\mathbf{a}) - D_{\mathbf{a}}\mathbf{v}(\mathbf{h})}{||\mathbf{h}||_2} = \mathbf{0}. \quad (4.3)$$

Resumé. Si différentiable $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ en $\mathbf{a} \in U$ est partiellement différentiable et continue en ce point. Sa matrice jacobienne donne l'application linéaire qui approche $\mathbf{v}$ dans un voisinage de $\mathbf{a} \in U$.

Définition. Soit $U \subset \mathbb{R}^n$ ouvert. Une application $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ est dite de classe $C^1(U)$ si les dérivées partielles $D_j v_i : U \longrightarrow \mathbb{R}$, $1 \leq i \leq m,\ 1 \leq j \leq n$ existent et sont continues sur U.

Théorème. Soit $U \subset \mathbb{R}^n$ ouvert et $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ de classe $C^1(U)$. Alors, pour tout $\mathbf{a} \in U$, $\mathbf{v}$ est différentiable. En particulier,

$$\mathbf{v}(\mathbf{a}+\mathbf{h}) = \mathbf{v}(\mathbf{a}) + J_{\mathbf{v}}(\mathbf{a})\mathbf{h} + \boldsymbol{o}(\mathbf{h}) \tag{4.4}$$

et

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\boldsymbol{o}(\mathbf{h})}{||\mathbf{h}||_2} = \mathbf{0}.$$

Corollaire. Soit $U \subset \mathbb{R}^n$ ouvert. Tout champ vectoriel $\mathbf{v} : U \longrightarrow \mathbb{R}^m$ de classe $C^1(U)$ est continue sur U.

Exemples.

1. Si $f : \mathbb{R} \to \mathbb{R}$ est dérivable en x, alors $J_f(x) = f'(x) \in M_{1,1}(\mathbb{R})$.
2. Soit $\mathbf{k} : \mathbb{R} \to \mathbb{R}^n$ en courbe dérivable en t. Alors, $J_{\mathbf{k}} = \dot{\mathbf{k}}(t) \in M_{n,1}(\mathbb{R})$.
3. Si $f : \mathbb{R}^n \to \mathbb{R}$ est partiellement différentiable en $\mathbf{x}$, alors

$$J_f(\mathbf{x}) = \big(D_1 f(\mathbf{x}), \dots D_n f(\mathbf{x})\big) = \nabla^T f(\mathbf{x}) M_{1,n}(\mathbb{R}).$$

Si f est différentiable en $\mathbf{x}$, alors $d_{\mathbf{x}} f(\mathbf{h}) = J_f(\mathbf{x}) \cdot \mathbf{h} = \langle \nabla f(\mathbf{x}), \mathbf{h} \rangle$.

4.1.1 La règle de composition

Théorème. Soient $U_1 \subset \mathbb{R}^n$, $U_2 \subset \mathbb{R}^m$ ouverts, $\mathbf{v} : U_1 \longrightarrow \mathbb{R}^m$ et $\mathbf{w} : U_2 \longrightarrow \mathbb{R}^k$ des champs vectoriels de classe $C^1(U_1)$ respectivement $C^1(U_2)$ tels que $\mathbf{v}(U_1) \subset U_2$. Alors, le champ vectoriel

$$\mathbf{w} \circ \mathbf{v} : U_1 \longrightarrow \mathbb{R}^k$$

est de classe $C^1(U_1)$ et

$$J_{\mathbf{w}\circ\mathbf{v}}(\mathbf{x}) = J_{\mathbf{w}}(\mathbf{v}(\mathbf{x})) \cdot J_{\mathbf{v}}(\mathbf{x}). \tag{4.5}$$

Calcul pratique. On voit de (4.5) que les éléments $(J_{\mathbf{w}\circ\mathbf{v}}(\mathbf{x}))_{i,j}$ se calcul d'une manière assez intuitive :

$$\frac{\partial\, w_i(\mathbf{v}(\mathbf{x}))}{\partial\, x_j} = \sum_{k=1}^{n} \frac{\partial\, v_k}{\partial\, x_j}(\mathbf{x}) \frac{\partial\, w_i(\mathbf{v})}{\partial\, v_k}(\mathbf{v}(\mathbf{x})) \tag{4.6}$$

Exemples. Au chapitre 3 nous avons déjà vu deux cas particuliers de la règle de composition :

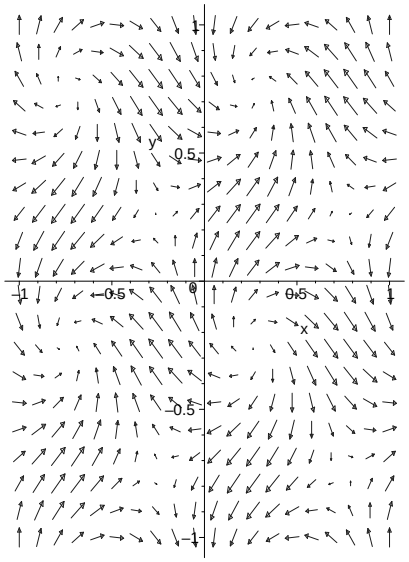
1. Soient $f : \mathbb{R}^n \to \mathbb{R}$, $\mathbf{k} : \mathbb{R} \to \mathbb{R}^n$ de classe C^1. Alors, $f \circ \mathbf{k} : \mathbb{R} \to \mathbb{R}$ est de classe C^1 et $J_{f\circ\mathbf{k}}(t) = \dfrac{d(f\circ\mathbf{k})(t)}{dt} = J_f(\mathbf{k}(t)) \cdot J_{\mathbf{k}}(t) = \langle \nabla f(\mathbf{k}(t)), \dot{\mathbf{k}}(t) \rangle$.
2. Soient $f : \mathbb{R} \to \mathbb{R}$, $\rho : \mathbb{R}^n \to \mathbb{R}$ de classe C^1. Alors, $f \circ \rho : \mathbb{R}^n \to \mathbb{R}$ est de classe C^1 et $J_{f\circ\rho}(\mathbf{x}) = J_f(\rho(\mathbf{x})) \cdot J_\rho(\mathbf{x}) = f'(\rho(\mathbf{x})) \cdot \nabla^T \rho(\mathbf{x})$.

Dans la suite nous étudions notamment les applications de la règle de composition aux changements des coordonnées.

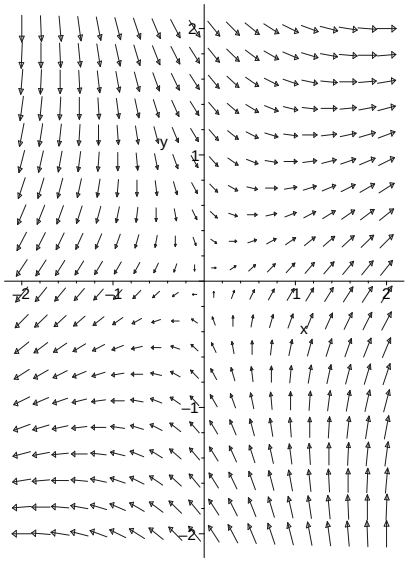
4.2 Champs vectoriels $\mathbb{R}^n \longrightarrow \mathbb{R}^n$

Introduction. Des Champs vectoriels $\mathbb{R}^n \longrightarrow \mathbb{R}^n$ ont des applications importantes : en physique ils décrivent par exemple les champs magnétiques et éléctriques ou les champs de vitesse d'une fluide. Les changements des coordonnées sont également des applications $\mathbb{R}^n \longrightarrow \mathbb{R}^n$.

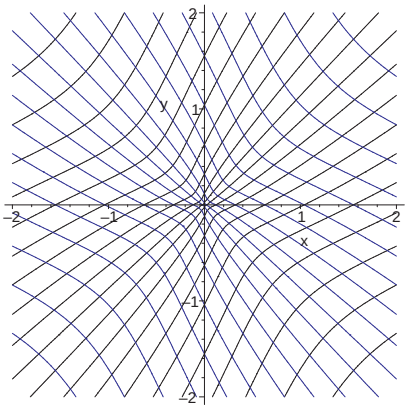
Représentaion graphique. Soit $U \subset \mathbb{R}^n$. Un champs vectoriel $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ est graphiquement représenté en attachant un flêche (i.e. un vecteur) à tout point $\mathbf{x} \in U$. Si l'application $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ donne un changement des coordonnées, alors on représente l'application $\mathbf{v}(\mathbf{x})$ par les ensembles de niveau de n fonctions reélles $v_k : U \longrightarrow \mathbb{R}$.

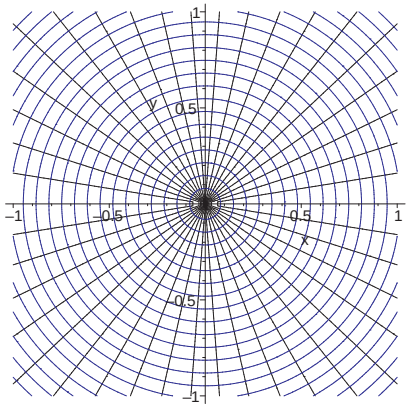


$\sin(\pi(x+y))\mathbf{e}_1 + cos(\pi(x-y))\mathbf{e}_2$



$\frac{x+y}{\sqrt[4]{4(x^2+y^2)}}\mathbf{e}_1 + \frac{x-y}{\sqrt[4]{4(x^2+y^2)}}\mathbf{e}_2$





Lignes de niveau de $\frac{x+y}{\sqrt[4]{4(x^2+y^2)}}$ et $\frac{x-y}{\sqrt[4]{4(x^2+y^2)}}$ respectivement de $\sqrt{x^2+y^2}$ et $\arctan(\frac{y}{x})$.

La divergence et le jacobien. La matrice jacobienne $J_{\mathbf{v}}$ est une matrice carrée. La trace de $J_{\mathbf{v}}$ est appelée la divergence du champ $\mathbf{v}$ et notée $\operatorname{div}\mathbf{v}$:

$$\operatorname{div}\mathbf{v}(\mathbf{x}) = \langle \nabla, \mathbf{v}(\mathbf{x})\rangle = \operatorname{tr}\left(J_{\mathbf{v}}(\mathbf{x})\right) = \sum_{k=1}^{n} D_k v_k(\mathbf{x}).$$

Le déterminant, $\det J_{\mathbf{v}(\mathbf{x})}$ est appelé le jacobien de $\mathbf{v}$ au point $\mathbf{x} \in \mathbb{R}^n$.

Remarque. Les colonnes de la matrice jacobienne d'un champ vectoriel $\mathbf{v}(\mathbf{x})$ sont les dérivées partielles de $\mathbf{v}(\mathbf{x})$, i.e.

$$J_{\mathbf{v}}(\mathbf{x}) = \begin{pmatrix} D_1\mathbf{v}(\mathbf{x}) & \cdots & D_n\mathbf{v}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \frac{\partial}{\partial x_1}\mathbf{v}(\mathbf{x}) & \cdots & \frac{\partial}{\partial x_n}\mathbf{v}(\mathbf{x}) \end{pmatrix}.$$

Alternativement : Les lignes de la matrice jacobienne d'un champ vectoriel $\mathbf{v}(\mathbf{x})$ sont les transposées des gradients des composantes $v_k(\mathbf{x})$, i.e.

$$J_{\mathbf{v}}(\mathbf{x}) = \begin{pmatrix} (\nabla v_1(\mathbf{x}))^T \\ \cdot \\ \cdot \\ \cdot \\ (\nabla v_n(\mathbf{x}))^T \end{pmatrix}.$$

Exemples.

1. Soit $A \in M_{n,n}(\mathbb{R})$ et $\mathbf{v} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ donné par l'application linéaire

   $$\mathbf{v}(\mathbf{x}) = A\mathbf{x}.$$

   Alors, $J_{\mathbf{v}}(\mathbf{x}) = A$ pour tout $\mathbf{x} \in \mathbb{R}$ et donc $\operatorname{div}\mathbf{v}(\mathbf{x}) = \operatorname{tr}(A)$. Son jacobien est $\det A$. Si la matrice A est inversible on peut interpréter l'application donnée par $A\mathbf{x}$ comme un changement linéaire des coordonnées. Par exemple, soit $\mathbf{v} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ donné par

   $$\mathbf{v}(x, y) = \begin{pmatrix} -y \\ x \end{pmatrix}$$

   Alors,

   $$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

2. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \longrightarrow \mathbb{R}$ une fonction réelle de classe $C^2(U)$. Alors, le gradient de f définit un champ vectoriel $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ de classe $C^1(U)$ donné par

   $$\mathbf{v}(\mathbf{x}) = \nabla f(\mathbf{x}).$$

   Alors, $J_{\mathbf{v}}(\mathbf{x}) = J_{\nabla f}(\mathbf{x}) = \operatorname{Hess}(f)$ et $\operatorname{div}\mathbf{v}(\mathbf{x}) = \operatorname{tr}(\operatorname{Hess}(f))$. Autrement dit

   $$\operatorname{div}(\nabla f(\mathbf{x})) = \Delta f(\mathbf{x}) \tag{4.7}$$

3. On considère l'application $\mathbf{v} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ donnée par

   $$\mathbf{v}(r, \phi) = \begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix}.$$

Cette application donne le changement de coordonnées entre les coordonnées polaires et les coordonnés cartésiennes si on pose

$$\begin{aligned} x &= v_1(r,\phi) = r\cos\phi \\ y &= v_2(r,\phi) = r\sin\phi \end{aligned}$$

Sa matrice jacobienne est

$$\begin{aligned} J_{\mathbf{v}}(r,\phi) &= \begin{pmatrix} D_1v_1(r,\phi) & D_2v_1(r,\phi) \\ D_1v_2(r,\phi) & D_2v_2(r,\phi) \end{pmatrix} \\ &= \begin{pmatrix} D_rv_1(r,\phi) & D_\phi v_1(r,\phi) \\ D_rv_2(r,\phi) & D_\phi v_2(r,\phi) \end{pmatrix} \\ &= \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix} \end{aligned}$$

4. Soit l'application $\mathbf{v} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ donné par

$$\mathbf{v}(r,\theta,\phi) = \begin{pmatrix} r\sin\theta\cos\phi \\ r\sin\theta\sin\phi \\ r\cos\theta \end{pmatrix}.$$

Cette application donne le changement de coordonnées entre le coordonnées sphériques et les coordonnées cartésiennes si on pose

$$\begin{aligned} x &= v_1(r,\theta,\phi) = r\sin\theta\cos\phi \\ y &= v_2(r,\theta,\phi) = r\sin\theta\sin\phi \\ z &= v_3(r,\theta,\phi) = r\cos\theta \end{aligned}$$

Sa matrice jacobienne est

$$\begin{aligned} D\mathbf{v}(r,\phi) &= \begin{pmatrix} D_1v_1(r,\theta,\phi) & D_2v_1(r,\theta,\phi) & D_3v_1(r,\theta,\phi) \\ D_1v_2(r,\theta,\phi) & D_2v_2(r,\theta,\phi) & D_3v_2(r,\theta,\phi) \\ D_1v_3(r,\theta,\phi) & D_2v_3(r,\theta,\phi) & D_3v_3(r,\theta,\phi) \end{pmatrix} \\ &= \begin{pmatrix} D_rv_1(r,\theta,\phi) & D_\theta v_1(r,\theta,\phi) & D_\phi v_1(r,\theta,\phi) \\ D_rv_2(r,\theta,\phi) & D_\theta v_2(r,\theta,\phi) & D_\phi v_2(r,\theta,\phi) \\ D_rv_3(r,\theta,\phi) & D_\theta v_3(r,\theta,\phi) & D_\phi v_3(r,\theta,\phi) \end{pmatrix} \\ &= \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix} \end{aligned}$$

5. Soit $U = \mathbb{R}^n \setminus \{(0,0)\}$ et $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ le champs vectoriel donné par

$$\mathbf{v}(\mathbf{x}) = \frac{\mathbf{x}}{r}$$

où $r = ||\mathbf{x}||_2$. Notons, que $\frac{\mathbf{x}}{r} = \nabla r$. Donc $\operatorname{div}\mathbf{v}(\mathbf{x}) = \Delta r = (n-1)/r$.

Une application de la règle du produit. Soit $U \subset \mathbb{R}^n$ ouvert ,$f : U \longrightarrow \mathbb{R}$ une fonction réelle de classe $C^1(U)$ et $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ un champ vectoriel de classe $C^1(U)$. Alors, pour tout $\mathbf{x} \in U$ et tout $k = 1, \ldots, n$

$$D_k(fv_k)(\mathbf{x}) = f(\mathbf{x})D_kv_k(\mathbf{x}) + D_kf(\mathbf{x})v_k(\mathbf{x})$$

et en prenant la somme sur $k = 1, \ldots, n$:

$$\operatorname{div}(f\mathbf{v}(\mathbf{x})) = f(\mathbf{x}) \cdot \operatorname{div}\mathbf{v}(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{v}(\mathbf{x}) \rangle$$

ou sous forme du produit scalaire :

$$\langle \nabla, f\mathbf{v} \rangle = f\langle \nabla, \mathbf{v} \rangle + \langle \nabla f, \mathbf{v} \rangle$$

4.2.1 Le rotationnel

Définition. Soit $U \subset \mathbb{R}^3$ ouvert et $\mathbf{v} : U \longrightarrow \mathbb{R}^3$ un champ vectoriel de classe $C^1(U)$. On appelle rotationnel de $\mathbf{v}$ le champ vectoriel $\mathbf{rot}\,\mathbf{v} : U \longrightarrow \mathbb{R}^3$ défini par

$$\mathbf{rot}\,\mathbf{v}(\mathbf{x}) = \begin{pmatrix} D_2 v_3(\mathbf{x}) - D_3 v_2(\mathbf{x}) \\ D_3 v_1(\mathbf{x}) - D_1 v_3(\mathbf{x}) \\ D_1 v_2(\mathbf{x}) - D_2 v_1(\mathbf{x}) \end{pmatrix}.$$

On peut écrire le rotationnel d'un champ vectoriel formellement comme produit vectoriel de l'opérateur ∇ et $\mathbf{v}$:

$$\mathbf{rot}\,\mathbf{v}(\mathbf{x}) = \nabla \times \mathbf{v}(\mathbf{x})$$

Rotationnel d'un gradient. Soit $f : U \longrightarrow \mathbb{R}$ une fonction réelle de classe $C^2(U)$. Alors,

$$\mathbf{rot}\,\mathbf{grad}\, f(\mathbf{x}) = \mathbf{0}$$

ou en utilisant la représentation par le produit vectoriel :

$$\nabla \times \nabla f(\mathbf{x}) = \mathbf{0}.$$

Ceci nous donne la condition néccessaire pour un champ vectoriel d'être le gradient d'une fonction réelle. Pour les conditions suffisantes voir l'Analyse 3.

Divergence du rotationnel. Soit $\mathbf{v} : U \longrightarrow \mathbb{R}^3$ un champ vectoriel de classe $C^2(U)$. Alors,

$$\operatorname{div}(\mathbf{rot}\,\mathbf{v}(\mathbf{x})) = \mathbf{0}$$

ou en utilisant la représentation par le produit vectoriel et le produit scalaire :

$$\langle \nabla, \nabla \times \mathbf{v}(\mathbf{x}) \rangle = 0.$$

Ceci nous donne la condition néccessaire pour un champ vectoriel d'être le rotaionnel d'un champ vectoriel. Pour les conditions suffisantes voir l'Analyse 3.

Exemple - Un champ magnétique constant. Soit $B > 0$ et $\mathbf{A}$ le "potentiel vectoriel" donné par

$$\mathbf{A}(x, y, z) = \begin{pmatrix} -By/2 \\ Bx/2 \\ 0 \end{pmatrix} = -By/2\mathbf{e}_1 + Bx/2\mathbf{e}_2.$$

On a

$$\nabla \times \mathbf{A}(x, y, z) = \begin{pmatrix} 0 \\ 0 \\ B \end{pmatrix} = B\mathbf{e}_3.$$

Si f est une fonction réelle de classe C^1 le rotationnel de $\mathbf{A} + \nabla f$ est toujours $B\mathbf{e}_3$. Cette propriétéest appelée l'invariance sous gauge

Le rotationnel en deux dimensions. Pour un champ vectoriel $\mathbf{v} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ on peut définir le rotationnel par le scalaire

$$\operatorname{rot} \mathbf{v}(\mathbf{x}) = D_1 v_2(\mathbf{x}) - D_2 v_1(\mathbf{x})$$

Si $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ est une fonction réelle de classe $C^2(U)$ on a de nouveau

$$\operatorname{rot} \mathbf{grad}\, f(\mathbf{x}) = 0$$

ou en utilisant la représentation par le produit vectoriel :

$$\nabla \times \nabla f(\mathbf{x}) = 0.$$

4.2.2 Inversibilité des champs vectoriels

Inversibilité des champs vectoriels Soit $U, V \subset \mathbb{R}^n$ ouverts. Si le champ vectoriel $\mathbf{v} : U \longrightarrow V$ est une application bijective, alors l'application réciproque $\mathbf{w} = \mathbf{v}^{-1} : V \longrightarrow U$ existe et

$$(\mathbf{w} \circ \mathbf{v})(\mathbf{x}) = \mathbf{x}, \quad (\mathbf{v} \circ \mathbf{w})(\mathbf{y}) = \mathbf{y}$$

pour tout $\mathbf{x} \in U$ et $\mathbf{y} \in V$.

Inversibilité - condition nécessaire. Soient $U \subset \mathbb{R}^n$ ouvert, $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ un champ vectoriel de classe $C^1(U)$ Si $\mathbf{v}$ et inversible, avec $\mathbf{w} = \mathbf{v}^{-1}$ l'application réciproque de classe C^1, alors

$$J_{\mathbf{w}\circ\mathbf{v}}(\mathbf{x}) = Id_n = J_{\mathbf{w}}(\mathbf{v}(\mathbf{x})) \cdot J_{\mathbf{v}}(\mathbf{x}) \tag{4.8}$$

et

$$\det J_{\mathbf{w}\circ\mathbf{v}}(\mathbf{x}) = 1 = \det J_{\mathbf{w}}(\mathbf{v}(\mathbf{x})) \cdot \det J_{\mathbf{v}}(\mathbf{x}). \tag{4.9}$$

Par conséquent, si $\mathbf{v}$ est inversible, alors $\det J_{\mathbf{v}}(\mathbf{x}) \neq 0$. La relation (4.9) nous permet également de calculer la matrice jacobienne du champ réciproque en $\mathbf{v}(\mathbf{x})$.

Exemple 1. L'application $\mathbf{v} : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2$ définie par

$$\mathbf{v}(x,y) = \frac{1}{\sqrt{x^2+y^2}} \begin{pmatrix} -y \\ x \end{pmatrix}$$

n'est pas inversible car

$$D\mathbf{v}(x,y) = \begin{pmatrix} D_1 v_1(x,y) & D_2 v_1(x,y) \\ D_1 v_2(x,y) & D_2 v_2(x,y) \end{pmatrix} = \frac{1}{\sqrt{x^2+y^2}^3} \begin{pmatrix} xy & -x^2 \\ y^2 & -xy \end{pmatrix}$$

et son jacobien est toujours zéro.

Exemple 2. Soit $\mathbf{v} : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2$ défini par

$$\mathbf{v}(x,y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

Sa matrice jacobien est donné par

$$J_{\mathbf{v}}(x,y) = \begin{pmatrix} D_1 v_1(x,y) & D_2 v_1(x,y) \\ D_1 v_2(x,y) & D_2 v_2(x,y) \end{pmatrix} = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}.$$

Son jacobien est $4x^2 + 4y^2 > 0$. L'application $\mathbf{v} : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2$ n'est pas bijective car $\mathbf{v}(x,y) = \mathbf{v}(-x,-y)$. Par conséquent, $\mathbf{v}$ n'est pas inversible sur son domaine de définition.

Applications localement inversibles. Le théorème suivante montre q'un champ vectoriel $\mathbf{v}$ de classe C^1 est toujour inversible dans un voisinage d'un point $\mathbf{a}$ si le jacobien est non nul au point $\mathbf{a}$. Une telle application est appelée localement inversible au point $\mathbf{a}$.

Théorème des applications localement inversibles. Soit $U \subset \mathbb{R}^n$ ouvert. Si le jacobien d'un champ vectoriel $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ de classe C^1 est non nul au point $\mathbf{a} \in U$, alors $\mathbf{v}$ est localement inversible autour du point $\mathbf{a} \in U$ avec une inverse de classe C^1.

Remarque. Une application de ce résultat est le théorème de fonctions implicites. On l'appliquera le théorème des applications localement inversibles plus loin au changement de coordonnées. Sa démonstration est basée sur le théorème du point fixe de Banach donné au premier chapitre.

Démonstration. Sans perdre la généralité on peut supposer $\mathbf{a} = \mathbf{0}$, $\mathbf{v}(\mathbf{a}) = \mathbf{0}$ et $J_{\mathbf{v}}(\mathbf{a}) = E_n$ où E_n présente la matrice d'identité. Par conséquent,

$$\mathbf{v}(\mathbf{h}) = \mathbf{h} + \boldsymbol{o}(\mathbf{h}).$$

On définit $\mathbf{g}(\mathbf{h}) = \mathbf{h} - \mathbf{v}(\mathbf{h})$. Par la continuité de $J_{\mathbf{v}}(\mathbf{x})$ et le théorème des accroissements finis (voir Analyse 2) il existe $\delta > 0$ tel que pour tout $\mathbf{x}_j, ||\mathbf{x}_j||_2 \leq \delta, j = 1, 2$:

$$||\mathbf{g}(\mathbf{x}_1) - \mathbf{g}(\mathbf{x}_1)||_2 \leq \frac{1}{2}||\mathbf{x}_1 - \mathbf{x}_2||_2. \tag{4.10}$$

En particulier, $\mathbf{g} : \overline{B}_\delta(\mathbf{0}) \longrightarrow \overline{B}_{\delta/2}(\mathbf{0})$ (prendre $\mathbf{x}_2 = \mathbf{0}$). Pour tout $\mathbf{y} \in \overline{B}_{\delta/2}(\mathbf{0})$ on définit une application $\mathbf{g}_\mathbf{y} : \overline{B}_\delta(\mathbf{0}) \longrightarrow \overline{B}_\delta(\mathbf{0})$ par

$$\mathbf{g}_\mathbf{y}(\mathbf{x}) = \mathbf{y} + \mathbf{g}(\mathbf{x}) \tag{4.11}$$

Par (4.10) l'application $\mathbf{g}_\mathbf{y}$ est une contraction sur l'espace métrique complét $\overline{B}_\delta(\mathbf{0})$. Elle admet donc un unique point fixe $\bar{\mathbf{x}}_\mathbf{y} \in \overline{B}_\delta(\mathbf{0})$. Autrement dit, pour tout $\mathbf{y} \in \overline{B}_{\delta/2}(\mathbf{0})$ il existe un unique $\bar{\mathbf{x}}_\mathbf{y} \in \overline{B}_\delta(\mathbf{0})$ tel que $\mathbf{y} = \mathbf{v}(\bar{\mathbf{x}}_\mathbf{y})$.

Applications linéaires. Pour des applications linéaires la condition $\det J_{\mathbf{v}}(\mathbf{x}) \neq 0$ est nécessaire et suffisante : soit $A \in M_{n,n}(\mathbb{R})$ et $\mathbf{v} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ donné par

$$\mathbf{v}(\mathbf{x}) = A\mathbf{x}.$$

Alors, $D\mathbf{v}(\mathbf{x}) = A$ pour tout $\mathbf{v} \in \mathbb{R}$ et le champ $\mathbf{v}$ est inversible si et seulement si la matrice A est inversible. Dans ce cas

$$\mathbf{w}(\mathbf{x}) = \mathbf{v}^{-1}(\mathbf{x}) = A^{-1}\mathbf{x}.$$

Une condition nécessaire et suffisante pour que la matrice A soit inversible est $\det A \neq 0$.

Le cas $n = 2$. Pour $a, b, c, d \in \mathbb{R}$ soit

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

tel que $\det A = ad - bc \neq 0$. Alors

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Exemple 2 - suite. On considère de nouveau $\mathbf{v} : \mathbb{R}^2 \setminus \{(0,0)\} \longrightarrow \mathbb{R}^2$ défini par

$$\mathbf{v}(x, y) = \begin{pmatrix} x^2 - y^2 \\ 2xy \end{pmatrix}.$$

Par le théorème des applications localement inversibles, $\mathbf{v}$ est localement inversible autour de tout point $(a, b) \in \mathbb{R}^2 \setminus \{(0,0)\}$. Pour calculer son application réciproque on doit résoudre le système

$$x^2 - y^2 = s, 2xy = t \tag{4.12}$$

pour (x, y). On cherche également le domaine ouvert maximal sur lequel $\mathbf{v}$ est inversible. Nous allons démontrer dans la suite que le système admet une solution unique (x, y) si $x > 0$, c'est-à-dire que l'application $\mathbf{v} :]0, \infty[\times\mathbb{R} \longrightarrow \mathbb{R}^2$ est bijective. En fait, la première équation de (4.12) implique $x = \sqrt{y^2 + s}$ puisque $x > 0$ (la solution $x = -\sqrt{y^2 + s}$ n'est pas admissible dans ce cas). Par la deuxième équation $2y\sqrt{y^2 + s} = t$, d'où en prenant le carré

$$4(y^2)^2 + 4sy^2 - t^2 = 0.$$

Si $t = 0$, alors $y = 0$ car $x = \sqrt{y^2 + s} > 0$, donc $x = \sqrt{s}$. . Si $t \neq 0$ seule la racine $y^2 = \dfrac{-s + \sqrt{s^2 + t^2}}{2} > 0$ est admissible (car $\dfrac{-s - \sqrt{s^2 + t^2}}{2} < 0$). La deuxième équation de (4.12) implique que y et t ont le même signe. On a trouvé :

$$\mathbf{v}^{-1}(s, t) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{s + \sqrt{s^2 + t^2}}{2}} \\ \\ \sqrt{\frac{-s + \sqrt{s^2 + t^2}}{2}} \end{pmatrix}$$

si $y > 0$,

$$\mathbf{v}^{-1}(s,t) = \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \sqrt{\frac{s+\sqrt{s^2+t^2}}{2}} \\ \\ -\sqrt{\frac{-s+\sqrt{s^2+t^2}}{2}} \end{pmatrix}$$

si $y < 0$ et si $t = 0$, alors $s > 0$

$$\mathbf{v}^{-1}(s,0) = \begin{pmatrix} x \\ 0 \end{pmatrix} = \begin{pmatrix} \sqrt{s} \\ \\ 0 \end{pmatrix}.$$

Cette application représente la partie réelle et imaginaire de l'application complexe $z = x + iy \mapsto z^2 = x^2 - y^2 + 2ixy$. Nous avons donc construit la partie réelle et imaginaire de $z \mapsto \sqrt{z}$ pour $\operatorname{Re} z > 0$ (voir Analyse 1, chapitre 1)..

Exemples - Calcul des champs réciproques et des matrices jacobiennes. Soit $\mathbf{v}$ un champ vectoriel de classe C^1 localement inversible au point $\mathbf{a}$. Pour calculer le champs réciproque il faut résoudre le système $\mathbf{y} = \mathbf{v}(\mathbf{x})$ pour tout $\mathbf{x}$ dans un voisinage de $\mathbf{a}$ (voir l'exemple ci-dessus). Le théorème affirme qu'il existe une solution unique dans un voisinage de $\mathbf{a}$. On la note $\mathbf{w}(\mathbf{y}) = \mathbf{v}^{-1}(\mathbf{y})$. Souvent on veut seulement déterminer la matrice jacobienne du champs réciproque (voir aussi le chapitre sur les intégrales multiples). Ceci peut se faire par la règle de composition sans passer par la détermination explicite du champs $\mathbf{v}^{-1}(\mathbf{y})$: la matrice jacobienne du champ vectoriel réciproque $\mathbf{w}$ au point $\mathbf{v}(\mathbf{a})$ est la matrice inverse de la matrice jacobienne $J_{\mathbf{v}}(\mathbf{a})$.

1. On considère l'application $\mathbf{v} :]0,\infty[\times \mathbb{R} \longrightarrow \mathbb{R}^2$ donné par

$$\mathbf{v}(r,\phi) = \begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix}.$$

Cette application donne le changement de coordonnées entre le coordonnées polaires et les coordonnées cartésiennes si on pose

$$\begin{aligned} x &= v_1(r,\phi) = r\cos\phi \\ y &= v_2(r,\phi) = r\sin\phi \end{aligned}$$

Sa matrice jacobien est

$$J_{\mathbf{v}}(r,\phi) = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}$$

Son jacobien $\det J_{\mathbf{v}}(r,\phi) = r$ est non nul si $r > 0$. Alors, $\mathbf{v}(r,\phi)$ est localement inversible en tout point $(r,\phi) \in]0,\infty[\times\mathbb{R}$. La matrice inverse de $J_{\mathbf{v}}(r,\phi)$,$(J_{\mathbf{v}}(r,\phi))^{-1}$ est donnée par

$$(J_{\mathbf{v}}(r,\phi))^{-1} = \frac{1}{r}\begin{pmatrix} r\cos\phi & r\sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix}$$

Soit $r > 0$. En utilisant les rélations $r = \sqrt{x^2+y^2}$, $x/r = \cos\phi$ et $y/r = \sin\phi$ on trouve

$$(J_{\mathbf{v}}(r,\phi))^{-1} = \frac{1}{x^2+y^2}\begin{pmatrix} x\sqrt{x^2+y^2} & y\sqrt{x^2+y^2} \\ -y & x \end{pmatrix} = J_{\mathbf{v}^{-1}}(x,y)$$

où $\mathbf{v}^{-1}$ est l'application réciproque de $\mathbf{v}$. On peut donner une réprésentation explicite de l'application réciproque. Soient $V = \{(r,\phi) : r > 0, \phi \in]-\pi,\pi[\}$ et $W = \mathbb{R}^2 \setminus (]-\infty,0] \times \{0\})$. Alors, l'application $\mathbf{v} : V \longrightarrow W$ est bijective et son application réciproque est donnée par

$$\mathbf{v}^{-1}(x,y) = \begin{pmatrix} \sqrt{x^2+y^2} \\ 2\arctan\frac{y}{x+\sqrt{x^2+y^2}} \end{pmatrix}.$$

2. On considère l'application $\mathbf{v} : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ donné par

$$\mathbf{v}(x,y) = \begin{pmatrix} e^y \\ x^2 \end{pmatrix}.$$

Cette application est localement inversible si $x > 0$ (ou $x < 0$) En effet,

$$J_{\mathbf{v}}(x,y) = \begin{pmatrix} 0 & e^y \\ 2x & 0 \end{pmatrix}$$

est inversible et

$$(J_{\mathbf{v}}(x,y))^{-1} = \begin{pmatrix} 0 & \frac{1}{2x} \\ e^{-y} & 0 \end{pmatrix}.$$

La forme explicite du champ réciproque est

$$\mathbf{w}(v_1,v_2)) = \begin{pmatrix} \sqrt{v_2} \\ \ln v_1 \end{pmatrix}.$$

3. On considère l'application $\mathbf{v} : \mathbb{R}^2 \longrightarrow]0,\infty[\times]0,\infty[$ donnée par

$$\mathbf{v}(x,y) = \begin{pmatrix} e^{x+y} \\ e^{x-y} \end{pmatrix}.$$

Sa matrice jacobienne

$$J_{\mathbf{v}}(x,y) = \begin{pmatrix} e^{x+y} & e^{x+y} \\ e^{x-y} & -e^{x-y} \end{pmatrix}$$

est inversible et

$$(J_{\mathbf{v}}(x,y))^{-1} = \frac{1}{2e^{2x}}\begin{pmatrix} e^{x-y} & e^{x+y} \\ e^{x-y} & -e^{x+y} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} e^{-(x+y)} & e^{-(x-y)} \\ e^{-(x+y)} & -e^{-(x-y)} \end{pmatrix} = \frac{1}{2}\begin{pmatrix} \frac{1}{v_1} & \frac{1}{v_2} \\ \frac{1}{v_1} & -\frac{1}{v_2} \end{pmatrix}.$$

Noter que $v_1 v_2 = e^{2x}$ et $v_1/v_2 = e^{2y}$ donc

$$\mathbf{w}(v_1,v_2)) = \frac{1}{2}\begin{pmatrix} \ln(v_1 v_2) \\ \ln(\frac{v_1}{v_2}) \end{pmatrix} = \begin{pmatrix} \ln v_1 + \ln v_2 \\ \ln v_1 - \ln v_2 \end{pmatrix}.$$

4. Changement de coordonnées entre le coordonnées sphériques et les coordonnées cartésiennes dans $\mathbb{R}^3$: On considère l'application

$$\begin{aligned} x &= v_1(r,\theta,\phi) = r\sin\theta\cos\phi \\ y &= v_2(r,\theta,\phi) = r\sin\theta\sin\phi \\ z &= v_3(r,\theta,\phi) = r\cos\theta \end{aligned}$$

Sa matrice jacobien est

$$J_{\mathbf{v}}(r,\phi,\theta) = \begin{pmatrix} \sin\theta\cos\phi & r\cos\theta\cos\phi & -r\sin\theta\sin\phi \\ \sin\theta\sin\phi & r\cos\theta\sin\phi & r\sin\theta\cos\phi \\ \cos\theta & -r\sin\theta & 0 \end{pmatrix}$$

Donc

$$\begin{aligned} \det J_{\mathbf{v}}(r,\theta,\phi) &= r\cos\theta\cos\phi\, r\sin\theta\cos\phi\cos\theta \\ &\quad + r\sin\theta\sin\phi\sin\theta\sin\phi\, r\sin\theta \\ &\quad + r\sin\theta\sin\phi\, r\cos\theta\sin\phi\cos\theta \\ &\quad + \sin\theta\cos\phi\, r\sin\theta\cos\phi\, r\sin\theta \\ &= r^2\sin\theta(\cos^2\phi\cos^2\theta + \sin^2\phi\sin^2\theta \\ &\quad + \cos^2\phi\sin^2\theta + \cos^2\phi\sin^2\theta) \\ &= r^2\sin\theta \end{aligned}$$

Pour $x>0, y>0, z>0$ l'application réciproque est donnée par

$$\begin{aligned} r &= w_1(x,y,z) = \sqrt{x^2+y^2+z^2} \\ \phi &= w_2(x,y,z) = \arcsin\frac{y}{\sqrt{x^2+y^2}} \\ \theta &= w_3(x,y,z) = \arccos\frac{z}{\sqrt{x^2+y^2+z^2}} \end{aligned}$$

4.2.3 Transformations du Gradient et du Laplacien

Transformation du Gradient. Soient $U, V \subset \mathbb{R}^2$ et $\mathbf{v} : U \longrightarrow V$ une application inversible de classe $C^1(U)$ et posons

$$\mathbf{v}(s,t) = (x,y).$$

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction réelle de classe C^1. On pose $g(s,t) = f(x,y) = f(\mathbf{v}(s,t))$ et réciproquement $f(x,y) = g(\mathbf{v}^{-1}(x,y)$. On peut écrire la matrice jacobienne de $\mathbf{v}(s,t)$ et de $\mathbf{v}^{-1}(x,y)$ sous la forme

$$J_{\mathbf{v}}(s,t) = \begin{pmatrix} D_s x & D_t x \\ D_s y & D_t y \end{pmatrix}, \quad \text{resp.} \quad J_{\mathbf{v}^{-1}}(x,y) = \begin{pmatrix} D_x s & D_y s \\ D_x t & D_y t \end{pmatrix}$$

En appliquant la règle de composition nous avons

$$\begin{aligned} \nabla_{x,y} g(s,t) &= \begin{pmatrix} \frac{\partial s}{\partial x}\frac{\partial g(s,t)}{\partial s} + \frac{\partial t}{\partial x}\frac{\partial g(s,t)}{\partial t} \\ \frac{\partial s}{\partial y}\frac{\partial g(s,t)}{\partial s} + \frac{\partial t}{\partial y}\frac{\partial g(s,t)}{\partial t} \end{pmatrix} \\ &= (J_{\mathbf{v}^{-1}}(x,y)^T \nabla_{s,t} g(s,t) \\ &= (J_{\mathbf{v}^{-1}}(\mathbf{v}(s,t))^T \nabla_{s,t} g(s,t). \end{aligned}$$

Par conséquent,

$$\begin{aligned} ||\nabla_{x,y} g(s,t)||_2^2 &= ||((D\mathbf{v}^{-1})(\mathbf{v}(s,t))^T \nabla_{s,t} g(s,t)||_2^2 \\ &= \nabla_{s,t} g(s,t)^T (D\mathbf{v}^{-1})(\mathbf{v}(s,t)\cdot((D\mathbf{v}^{-1})(\mathbf{v}(s,t))^T \nabla_{s,t} g(s,t), \end{aligned}$$

où

$$(D\mathbf{v}^{-1})(\mathbf{v}(s,t)\cdot((D\mathbf{v}^{-1})(\mathbf{v}(s,t))^T = \begin{pmatrix} ||\nabla_{x,y}s||_2^2 & \langle\nabla_{x,y}s,\nabla_{x,y}t\rangle \\ \langle\nabla_{x,y}s,\nabla_{x,y}t\rangle & ||\nabla_{x,y}t||_2^2. \end{pmatrix}$$

désigne le tenseur métrique des coordonnés (s,t).

Exemple - coordonnées polaires. Soient $V = \{(r,\phi) : r > 0, \phi \in]-\pi,\pi[\,\}$ et $W = \mathbb{R}^2 \setminus (]-\infty,0] \times \{0\})$. On considère l'application bijective $\mathbf{v} : V \longrightarrow W$ donné par

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{v}(r,\phi) = \begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix}.$$

Rapellons que sa matrice jacobienne est

$$J_{\mathbf{v}}(r,\phi) = \begin{pmatrix} \cos\phi & -r\sin\phi \\ \sin\phi & r\cos\phi \end{pmatrix}$$

et la matrice inverse de $J_{\mathbf{v}}(r,\phi)$,$(J_{\mathbf{v}}(r,\phi))^{-1}$ est donnée par

$$\begin{aligned}(J_{\mathbf{v}}(r,\phi))^{-1} &= \frac{1}{r}\begin{pmatrix} r\cos\phi & r\sin\phi \\ -\sin\phi & \cos\phi \end{pmatrix} \\ &= \frac{1}{x^2+y^2}\begin{pmatrix} x\sqrt{x^2+y^2} & y\sqrt{x^2+y^2} \\ -y & x \end{pmatrix} = J_{\mathbf{v}^{-1}}(x,y)\end{aligned}$$

Il en suit (également par un calcul direct) :

$$D_x g(r,\phi) = \cos\phi\, D_r g(r,\phi) - \frac{\sin\phi}{r}\, D_\phi g(r,\phi)$$

$$D_y g(r,\phi) = \sin\phi\, D_r g(r,\phi) + \frac{\cos\phi}{r}\, D_\phi g(r,\phi)$$

et

$$(D_x g(r,\phi))^2 + (D_y g(r,\phi))^2 = (D_r g(r,\phi))^2 + \frac{1}{r^2}\,(D_\phi g(r,\phi))^2.$$

Les formules de transformation pour les dérivés partielles

$$\frac{\partial\, g(r,\phi)}{\partial\, x} = \frac{\partial\, r}{\partial\, x}\,\frac{\partial\, g(r,\phi)}{\partial\, r} + \frac{\partial\,\phi}{\partial\, x}\,\frac{\partial\, g(r,\phi)}{\partial\,\phi}$$

et

$$\frac{\partial\, g(r,\phi)}{\partial\, y} = \frac{\partial\, r}{\partial\, y}\,\frac{\partial\, g(r,\phi)}{\partial\, r} + \frac{\partial\,\phi}{\partial\, y}\,\frac{\partial\, g(r,\phi)}{\partial\,\phi}.$$

sont donc calculées d'une façon très intuitive : la dérivée partielle par rapport à la variable x (resp. y) d'une fonction $g(r,\phi)$ est calculée en prenant la somme des dérivées partielles des coordonnés (r,ϕ) par rapport à x (resp. y) mulitpliée par la dérivée partielle de $g(r,\phi)$ par rapport à ces coordonnés. Noter également que

$$(D_x g(r,\phi))^2 + (D_y g(r,\phi))^2 = (D_r g(r,\phi))^2 + \frac{1}{r^2}\,(D_\phi g(r,\phi))^2.$$

Transformation du Laplacien. Soit $U \subset \mathbb{R}^n$ et $\mathbf{v} : U \longrightarrow \mathbb{R}^n$ une application inversible de classe $C^2(U)$. On pose

$$\mathbf{x} = \mathbf{v}(\mathbf{s}).$$

Pour une fonction $g : U \longrightarrow \mathbb{R}$ de classe $C^2(U)$ donnée comme fonction des coordonnées $\mathbf{s}$ on veut calculer

$$\Delta_{\mathbf{x}} g(\mathbf{s}) = \sum_{i=1}^{n} D_{x_i x_i} g(\mathbf{s})$$

Comme avant on calcul

$$\frac{\partial\, g(\mathbf{s})}{\partial\, x_i} = \sum_{j=1}^{n} \frac{\partial\, s_j}{\partial\, x_i} \frac{\partial\, g(\mathbf{s})}{\partial\, s_j}$$

et

$$\begin{aligned}\frac{\partial^2\, g(\mathbf{s})}{\partial\, x_i^2} = &\sum_{j=1}^{n} \frac{\partial^2\, s_j}{\partial\, x_i^2} \frac{\partial\, g(\mathbf{s})}{\partial\, s_j} \\ &+ \sum_{j=1}^{n} \frac{\partial\, s_j}{\partial\, x_i} \sum_{k=1}^{n} \frac{\partial\, s_k}{\partial\, x_i} \frac{\partial^2\, g(\mathbf{s})}{\partial\, s_j s_k}\end{aligned}$$

Donc

$$\begin{aligned}\Delta_{\mathbf{x}} g(\mathbf{s}) = &\sum_{j=1}^{n} \Delta_{\mathbf{x}} s_j \frac{\partial\, g(\mathbf{s})}{\partial\, s_j} \\ &+ \sum_{j=1}^{n} \sum_{k=1}^{n} \langle \nabla_{\mathbf{x}} s_j, \nabla_{\mathbf{x}} s_k \rangle \frac{\partial^2\, g(\mathbf{s})}{\partial\, s_j s_k}\end{aligned}$$

Exemple - coordonnées polaires. Soient $V = \{(r, \phi) : r > 0, \phi \in]-\pi, \pi[\,\}$ et $W = \mathbb{R}^2 \setminus (]-\infty, 0] \times \{0\})$. On considère l'application bijective $\mathbf{v} : V \longrightarrow W$ donné par

$$\mathbf{v}(r, \phi) = (x, y) = (r \cos\phi, r \sin\phi).$$

et $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ une fonction réelle de classe C^1. On pose $g(r, \phi) = f(x, y) = f(\mathbf{v}(r, \phi))$. On va calculer le laplacien de g directement :

$$\frac{\partial\, g(r, \phi)}{\partial\, x} = \frac{\partial\, r}{\partial\, x} \frac{\partial\, g(r, \phi)}{\partial\, r} + \frac{\partial\, \phi}{\partial\, x} \frac{\partial\, g(r, \phi)}{\partial\, \phi}$$

et

$$\begin{aligned}\frac{\partial^2\, g(r, \phi)}{\partial\, x^2} = &\frac{\partial^2\, r}{\partial\, x^2} \frac{\partial\, g(r, \phi)}{\partial\, r} + \left(\frac{\partial\, r}{\partial\, x}\right)^2 \frac{\partial^2\, g(r, \phi)}{\partial\, r^2} + \frac{\partial\, r}{\partial\, x} \frac{\partial\, \phi}{\partial\, x} \frac{\partial^2\, g(r, \phi)}{\partial\, \phi \partial\, r} \\ &+ \frac{\partial^2\, \phi}{\partial\, x^2} \frac{\partial\, g(r, \phi)}{\partial\, \phi} + \left(\frac{\partial\, \phi}{\partial\, x}\right)^2 \frac{\partial^2\, g(r, \phi)}{\partial\, \phi^2} + \frac{\partial\, r}{\partial\, x} \frac{\partial\, \phi}{\partial\, x} \frac{\partial^2\, g(r, \phi)}{\partial\, \phi \partial\, r}.\end{aligned}$$

Donc

$$\begin{aligned}\Delta_{x,y} g(r,\phi) =& \Delta_{x,y} r \frac{\partial g(r,\phi)}{\partial r} + ||\nabla_{x,y} r||_2^2 \frac{\partial^2 g(r,\phi)}{\partial r^2} \\ &+ \Delta_{x,y}\phi \frac{\partial g(r,\phi)}{\partial \phi} + ||\nabla_{x,y}\phi||_2^2 \frac{\partial^2 g(r,\phi)}{\partial \phi^2} \\ &+ 2\langle \nabla_{x,y} r, \nabla_{x,y}\phi\rangle \frac{\partial^2 g(r,\phi)}{\partial \phi \partial r}\end{aligned}$$

En utilisant la matrice jacobienne on voit que :

$$||\nabla_{x,y} r||_2^2 = 1, \quad ||\nabla_{x,y}\phi||_2^2 = \frac{1}{r^2}, \quad \langle \nabla_{x,y} r, \nabla_{x,y}\phi\rangle = 0$$

De plus,

$$\Delta_{x,y} r = \frac{d^2 r}{d r^2} + \frac{1}{r}\frac{d r}{d r} = \frac{1}{r}$$

et

$$\Delta_{x,y}\phi = -\frac{\partial}{\partial x}\frac{y}{r^2} + \frac{\partial}{\partial y}\frac{x}{r^2} = 0.$$

Par conséquent,

$$\Delta_{x,y} g(r,\phi) = \frac{\partial^2 g(r,\phi)}{\partial r^2} + \frac{1}{r}\frac{\partial g(r,\phi)}{\partial r} + \frac{1}{r^2}\frac{\partial^2 g(r,\phi)}{\partial \phi^2}.$$

4.3 Fonctions implicites

Introduction. Soit $U \subset \mathbb{R}^2$ ouvert et $f : U \longrightarrow \mathbb{R}$ une fonction de classe $C^1(U)$. Pour $c \in \mathbb{R}$ on considère l'ensemble de niveau c de f :

$$N_f(c) = \{(x,y) \in \mathbb{R}^2 : f(x,y) = c\}$$

Pour saisir la forme de $N_f(c)$ on peut essayer de résoudre l'équation $f(x,y) = c$ pour y en fonction de x (ou pour x en fonction de y). Dans le premier cas, $N_f(c)$ peut être représenter comme graphe d'une fonction réelle $g(x)$ et $y = g(x)$ est son équation.

Exemple. Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ donnée par

$$f(x,y) = x^2 + e^y - 1.$$

Pour déterminer la forme de $N_f(0)$ on essaye résoudre l'équation $x^2 + e^y - 1 = 0$ pour y en fonction de x. Une solution existe seulement pour $-1 \leq x \leq 1$ et on trouve

$$y = \ln(1 - x^2)$$

Donc $N_f(0)$ est donné par le graphe de la fonction

$$g(x) = \ln(1 - x^2).$$

La fonction g est dérivable en tout $x \in]-1, +1[$ et vérifie la rélation

$$x^2 + e^{g(x)} - 1 = 0.$$

On peut utilser cette rélation pour calculer la dérivée de $g(x)$. On appelle cette technique la différentiation implicite. En calculant la dérivée par rapport à x on obtient

$$2x + g'(x)e^{g(x)} = 0$$

i.e.

$$g'(x) = -2xe^{-g(x)} = -\frac{2x}{1-x^2}$$

Noter que $D_1 f(x,y) = 2x$ et $D_2 f(x,y) = e^y$. On a donc

$$g'(x) = -\frac{D_1 f(x, g(x))}{D_2 f(x, g(x))}.$$

Cette rélation est vérifiée dans le cas général :

Théorème - théorème des fonctions implicites. Soit $U \subset \mathbb{R}^2$ ouvert et $f : U \longrightarrow \mathbb{R}$ une fonction de classe $C^k(U)$, $k \geq 1$, telle que $f(a,b) = 0$ pour un point $(a,b) \in U$ et $D_2 f(a,b) \neq 0$. Alors, il existe un voisinage $B_\epsilon(a)$ et une unique fonction $g : B_\epsilon(a) \longrightarrow \mathbb{R}$ tels que

1. Le graphe de g est dans U : $G_g = \{(x, g(x)) : x \in B_\epsilon(a)\} \subset U$.
2. $g(a) = b$
3. $f(x, g(x)) = 0$ pour tout $x \in B_\epsilon(a)$.
4. La fonction g est de classe $C^k(B_\epsilon(a))$ et

$$g'(x) = -\frac{D_1 f(x, g(x))}{D_2 f(x, g(x))}$$

pour tout $x \in B_\epsilon(a)$.

Remarque. Le théorème affirme que près du point $(a,b) \in \mathbb{R}^2$ on peut résoudre l'équation $f(x,y) = 0$ (ou plus généralement $f(x,y) = c$ pour un $c \in \mathbb{R}$) en exprimant la variable y comme fonction de la variable x. Autrement dit, la fonction $y = g(x)$ est donnée implicitement par l'équation

$$f(x, g(x)) = 0.$$

Dans un voisinge du point $(a,b) \in \mathbb{R}^2$ l'ensemble de niveau $N_f(0)$ est donné par le graphe de g.

Remarque. La rélation

$$g'(x) = -\frac{D_1 f(x, g(x))}{D_2 f(x, g(x))}$$

est vérifie dans un voisinage de a car la continuité des dérivées implique qu'il existe un vosinage de (a,b) tel que $D_2 f(x,y) \neq 0$.

Démonstration. On définit une application $\mathbf{v} : U \longrightarrow \mathbb{R}^2$ de classe C^1 par

$$\mathbf{v}(x,y) = \begin{pmatrix} s \\ t \end{pmatrix} := \begin{pmatrix} x \\ f(x,y) \end{pmatrix}$$

Sa matrice jacobienne est donnée par

$$J_{\mathbf{v}}(x,y) \begin{pmatrix} \frac{\partial s}{\partial x} & \frac{\partial s}{\partial y} \\ \frac{\partial t}{\partial x} & \frac{\partial t}{\partial y} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \frac{\partial f(x,y)}{\partial x} & \frac{\partial f(x,y)}{\partial y} \end{pmatrix}.$$

Le jacobien est $\det J_{\mathbf{v}}(x,y) = \dfrac{\partial f(x,y)}{\partial y}$. En (a,b) :

$$\mathbf{v}(a,b) = \begin{pmatrix} a \\ 0 \end{pmatrix}, \quad \det J_{\mathbf{v}}(a,b) = D_2 f(a,b) \neq 0.$$

Par conséquent, $\mathbf{v}$ est inversible dans un voisinage de (a,b) et l'application réciproque est donnée par :

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x(s,t) \\ y(s,t) \end{pmatrix} = \begin{pmatrix} s \\ y(s,t) \end{pmatrix}.$$

Poser $g(s) := y(s,0)$. Par construction, g est de classe C^1 et son graphe est dans U. On a $g(a) = b$ et

$$\begin{pmatrix} s \\ 0 \end{pmatrix} = \mathbf{v}(x(s,0), y(s,0)) \begin{pmatrix} s \\ f(s, g(s)) \end{pmatrix}$$

d'où $0 = f(s, g(s))$. La propriété 4.suit de la règle de composition.

Exemple. L'équation $x \sin x - e^y \sin y = 0$ admet une solution unique $y = g(x)$ de classe C^1 (même C^∞, voir plus loin) dans le voisinage de $(x,y) = (0,0)$. En fait, $f(x,y) = x \sin x - e^y \sin y$ vérifie les hypothèses du théorème des fonctions implicites. En particulier, $f(0,0) = 0$ et $D_2 f(0,0) = -1$.

Généralisation du théorème des fonctions implicites. Soit $U \subset \mathbb{R}^n$ ouvert et $f : U \longrightarrow \mathbb{R}$ une fonction de classe $C^k(U)$, $k \geq 1$, telle que $f(a_1, \ldots, a_{n-1}, a_n) = 0$ pour un point $(a_1, \ldots, a_{n-1}, a_n) \in U$ et $D_n f(a_1, \ldots, a_{n-1}, a_n) \neq 0$. Alors, il existe un voisinage $B_\epsilon(a) \subset \mathbb{R}^{n-1}$ et une unique fonction $g : B_\epsilon(a_1, \ldots, a_{n-1}) \longrightarrow \mathbb{R}$ tel que

1. Le graph de g est dans U : $G_g = \{(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) : x \in B_\epsilon(a_1, \ldots, a_{n-1})\} \subset U$.
2. $g(a_1, \ldots, a_{n-1}) = a_n$
3. $f((x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}))) = 0$ pour tout $(x_1, \ldots, x_{n-1}) \in B_\epsilon(a_1, \ldots, a_{n-1})$.
4. La fonction g est de classe $C^k(B_\epsilon(a_1, \ldots, a_{n-1}))$ et

$$D_i g(x_1, \ldots, x_{n-1}) = -\frac{D_i f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}))}{D_n f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}))}$$

pour tout $x \in B_\epsilon(a)$.

Exemple. Montrer que près du point $(1,2,3)$,

$$6x^4 + xyz + y^4 - y^2z^2 + 8 = 0$$

est l'équation d'une surface dans $\mathbb{R}^3$. Trouver les équations du plan tangent de cette surface au point $(1,2,3)$. On pose

$$f(x,y,z) = 6x^4 + xyz + y^4 - y^2z^2 + 8.$$

La fonction f est de classe $C^1(\mathbb{R}^3$ (même de classe $C^k(\mathbb{R}^3$ pour tout $k \in \mathbb{N}$) et a pour dérivées partielles

$$D_1f(x,y,z) = 24x^3+yz, \quad D_2f(x,y,z) = xz+4y^3-2yz^2, \quad D_3f(x,y,z) = xy-2y^2z$$

on a $f(1,2,3) = 0$ et $D_3f(1,2,3) = -22 \neq 0$. Alors, il existe une unique fonction $g(x,y)$ définie dans un voisinage de $(1,2)$ telle que

$$g(1,2) = 3, \quad f(x,y,g(x,y)) = 0.$$

L'équation du plan tangent est donnée par

$$z = g(1,2) + \langle \nabla g(1,2), \begin{pmatrix} x-1 \\ y-2 \end{pmatrix} \rangle$$

Avec

$$D_1g(1,2) = -\frac{D_1f(1,2,g(1,2))}{D_3f(1,2,g(1,2))} = -\frac{D_1f(1,2,3)}{D_3f(1,2,3)}$$

et

$$D_2g(1,2) = -\frac{D_2f(1,2,g(1,2))}{D_3f(1,2,g(1,2))} = -\frac{D_2f(1,2,3)}{D_3f(1,2,3)}$$

on obtient

$$(z-3)D_3f(1,2,3) = -D_1f(1,2,3)(x-1) - D_2f(1,2,3)(y-2)$$

i.e.

$$0 = D_1f(1,2,3)(x-1)+D_2f(1,2,3)(y-2)+(z-3)D_3f(1,2,3) = \langle \nabla f(1,2,3), \begin{pmatrix} x-1 \\ y-2 \\ z-3 \end{pmatrix} \rangle$$

Cette équation confirme l'intiution géométrique que le gradient de f est orthogonal au plan tangent de la surface $N_f(0)$. Donc

$$z = \frac{19}{11} + \frac{15x}{11} - \frac{y}{22}$$

est l'équation du plan tangent.

Exemple - Discussion. Si $D_2f(a,b) = 0$ la fonction $g(x)$ peut exister mais elle n'est pas forcement unique : Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ donnée par

$$f(x,y) = x^2 + y^4 - 1.$$

Soit $(a,b) = (1,0)$. L'équation $x^2 + y^4 - 1 = 0$ pour y en fonction de x admet deux solutions pour $-1 \leq x \leq 1$ et on trouve

$$y = \begin{cases} \sqrt[4]{1-x^2} & \text{si } y \geq 0, \\ -\sqrt[4]{1-x^2} & \text{si } y < 0. \end{cases}$$

De plus, la fonction

$$g(x) = \sqrt[4]{1-x^2}$$

n'est pas dérivable en $x = 0$. Par contre, pour le point $(a,b) = (0,1)$ le théorème des fonctions implicites est applicable et donne la fonction unique

$$g(x) = \sqrt[4]{1-x^2}$$

dans un voisinage de $a = 0$ prolongable pour tout $x \in]-1,+1[$ et dérivable sur cet intervalle avec la dérivée donnée par

$$g'(x) = -\frac{D_1 f(x,g(x))}{D_2 f(x,g(x))} = -\frac{x}{2\sqrt[4]{(1-x^2)^3}}.$$

Dérivées d'ordre supérieure. On considère le théorème des fonctions implicites pour $U \subset \mathbb{R}^2$ ouvert et une fonction $f : U \longrightarrow \mathbb{R}$ de classe $C^k(U)$, $k \geq 2$. Si les hypothèses du théorème sont vérifiées, il existe une fonction $g(x)$ telle que $f(x,g(x)) = 0$. La dérivée de $g(x)$ vérifie la rélation

$$D_1 f(x,g(x)) + g'(x) D_2 f(x,g(x)) = 0$$

En prenant la dérivée de cette identité on obtient

$$D_{11} f(x,g(x)) + 2g'(x) D_{12} f(x,g(x)) + g''(x) D_2 f(x,g(x)) + g'(x)^2 D_{22} f(x,g(x)) = 0$$

i.e.

$$g''(x) = -\frac{D_{11} f(x,g(x)) + 2g'(x) D_{12} f(x,g(x)) + g'(x)^2 D_{22} f(x,g(x))}{D_2 f(x,g(x))}.$$

en particulier, si $g'(a) = 0$ (tangent horizontal en a), alors

$$g''(a) = -\frac{D_{11} f(a,g(a))}{D_2 f(a,g(a))} = -\frac{D_{11} f(a,b)}{D_2 f(a,b)}.$$

Nous pouvons alors calculer les dérivés successives de la fonction implicite $g(x)$ en a sans déterminer $g(x)$ et, par conséquent, donner son développement limité en $x = a$.

4.3.1 La technique de la différentiation implicite

La technique appliquée ci-dessus pour calculer s'appelle la différentiation implicite. Nous pouvons l'appliquer, par exemple, aux changements de variables pour calculer les dérivées partielles sans invertir explicitement le changement de variables.

Exemple - coordonnées polaires. Soient $V = \{(r,\phi) : r > 0, \phi \in]-\pi,\pi[\,\}$ et $W = \mathbb{R}^2 \setminus (]-\infty, 0] \times \{0\})$. On considère l'application bijective $\mathbf{v} : V \longrightarrow W$ donné par

$$\mathbf{v}(r,\phi) = (x,y) = (r\cos\phi, r\sin\phi).$$

On va calculer les dérivées partielles $\dfrac{\partial r}{\partial x}, \dfrac{\partial r}{\partial y}, \dfrac{\partial \phi}{\partial x}, \dfrac{\partial \phi}{\partial y}$ par différentiation implicite en prenant les dérivées partielles par rapport à x et y dans les équations du changement de variables $x = r\cos\phi$ et $y = r\sin\phi$. On obtient quatre équations linéaires en les quatre dérivées partielles cherchées :

$$\begin{aligned} 1 &= \cos\phi\frac{\partial r}{\partial x} - r\sin\phi\frac{\partial \phi}{\partial x} \\ 0 &= \cos\phi\frac{\partial r}{\partial y} - r\sin\phi\frac{\partial \phi}{\partial y} \\ 0 &= \sin\phi\frac{\partial r}{\partial x} + r\cos\phi\frac{\partial \phi}{\partial x} \\ 1 &= \sin\phi\frac{\partial r}{\partial y} + r\cos\phi\frac{\partial \phi}{\partial y}. \end{aligned}$$

On obtient :

$$\frac{\partial r}{\partial x} = \cos\phi,\ \frac{\partial r}{\partial y} = \sin\phi,\ \frac{\partial \phi}{\partial x} = \frac{-\sin\phi}{r},\ \frac{\partial \phi}{\partial y} = \frac{\cos\phi}{r}.$$

4.4 Addendum - Quelques formules

Soient $f, g := \mathbb{R}^n \longrightarrow \mathbb{R}$, $\mathbf{u}, \mathbf{v}, \mathbf{w} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ suffisamment regulière.

$$\begin{aligned} \nabla(fg) &= f\nabla g + g\nabla f, && (4.13)\\ \operatorname{div}(f\mathbf{u}) &= \langle \nabla f, \mathbf{u}\rangle + f\operatorname{div}\mathbf{u}, && (4.14)\\ \operatorname{div}(\mathbf{v}\times\mathbf{w}) &= \langle \mathbf{w}, \nabla\times\mathbf{v}\rangle - \langle \mathbf{v}, \nabla\times\mathbf{w}\rangle, (n=2,3) && (4.15)\\ \nabla\times(f\mathbf{u}) &= \nabla f\times\mathbf{u} + f\nabla\times\mathbf{u}, (n=2,3). && (4.16) \end{aligned}$$

$$\begin{aligned} \Delta(fg) &= f\Delta g + 2\langle \nabla f, \nabla g\rangle + g\Delta f, && (4.17)\\ \Delta\ln(f^2) &= 2\frac{f\Delta f - \langle \nabla f, \nabla f\rangle}{f^2}, f \neq 0. && (4.18) \end{aligned}$$

$$\begin{aligned} \langle \mathbf{v}, \nabla\rangle\mathbf{w} &= J_{\mathbf{w}}\cdot\mathbf{v}, && (4.19)\\ \nabla(\langle \mathbf{v}, \mathbf{w}\rangle) &= J_{\mathbf{v}}^T\cdot\mathbf{w} + J_{\mathbf{w}}^T\cdot\mathbf{v}. && (4.20) \end{aligned}$$

Chapitre 5

Extremums locaux

Comme dans le cas des fonctions à une variable, la détermination d'extremums locaux d'une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ est important en vue des nombreuses applications.

5.1 Extremums et points stationnaires

Définition. Une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ atteint un maximum locale en $\mathbf{a} \in \mathbb{R}^n$ si il existe une boule $B_\delta(\mathbf{a})$ telle que $f(\mathbf{x}) \leq f(\mathbf{a})$ pour tout $\mathbf{x} \in B_\delta(\mathbf{a})$. Le maximum est appelé maximum stricte si $f(\mathbf{x}) < f(\mathbf{a})$ pour $\mathbf{x} \neq \mathbf{a}$.
Une fonction $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ atteint un minimum locale en $\mathbf{a} \in \mathbb{R}^n$ si il existe une boule $B_\delta(\mathbf{a})$ telle que $f(\mathbf{x}) \geq f(\mathbf{a})$ pour tout $\mathbf{x} \in B_\delta(\mathbf{a})$. Le maximum est appelé minimum stricte si $f(\mathbf{x}) > f(\mathbf{a})$ pour $\mathbf{x} \neq \mathbf{a}$.
On dit que f admet un extremum locale en $\mathbf{a} \in \mathbb{R}^n$ si f admet un maximum ou minimum locale en $\mathbf{a} \in \mathbb{R}^n$.

Définition. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction différentiable en $\mathbf{a} \in \mathbb{R}^n$ avec $\nabla f(\mathbf{a}) = \mathbf{0}$. On dit que $\mathbf{a}$ est un point stationnaire (ou point critique) de f. Un point stationnaire $\mathbf{a} \in \mathbb{R}^n$ est appelé point selle de f si il existe deux vecteurs $\mathbf{v}_1, \mathbf{v}_2 \in \mathbb{R}^n$ tels que la fonction $t \mapsto f(\mathbf{a}+t\mathbf{v}_1)$ admet un maximum locale stricte en $t = 0$ et $t \mapsto f(\mathbf{a} + t\mathbf{v}_2)$ admet un minimum locale stricte en $t = 0$.

Théorème - condition neccessaire. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ différentiable en $\mathbf{a} \in \mathbb{R}^n$. Si f admet un extremum locale en $\mathbf{a}$, alors $\mathbf{a}$ est un point stationnaire de f, c'est-à-dire $\nabla f(\mathbf{a}) = \mathbf{0}$.

Démonstration. Sans perdre la généralité on peut supposer que f admet un maximum local en $\mathbf{a}$. Soit $B_\delta(\mathbf{a})$ tel que $f(\mathbf{x}) \leq f(\mathbf{a})$ pour tout $\mathbf{x} \in B_\delta(\mathbf{a})$. Alors, pour tout $\mathbf{a}+\mathbf{h} \in B_\delta(\mathbf{a})$ la fonction $g : [-1, 1] \longrightarrow \mathbb{R}$ définie par $g(t) = f(\mathbf{a}+t\mathbf{h})$ est dérivable en $t = 0$ et admet un maximum local en $t = 0$, d'où

$$g'(0) = \langle \nabla f(\mathbf{a}), \mathbf{h} \rangle = 0.$$

pour tous $\mathbf{h} \in B_\delta(\mathbf{0})$ d'où $\nabla f(\mathbf{a}) = \mathbf{0}$.

5.2 Formes quadratiques

Soit $A \in M_{n,n}(\mathbb{R})$ une matrice symétrique et $q : \mathbb{R}^n \longrightarrow \mathbb{R}$ la forme quadratique définie par

$$q(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle.$$

La fonction $q(\mathbf{x})$ est de classe C^2 avec

$$\nabla q(\mathbf{x}) = \sum_{k=1}^{n} \langle A\mathbf{x}, \mathbf{e_k}\rangle \, \mathbf{e_k} = A\mathbf{x}$$

et

$$\text{Hess}\,(q)(\mathbf{x}) = A.$$

Pour toutes points $\mathbf{x}, \mathbf{a} \in \mathbb{R}^n$ la forme quadratique $q(\mathbf{x})$ admet la représentation

$$q(\mathbf{x}) = q(\mathbf{a}) + \langle \nabla q(\mathbf{a}), \mathbf{x} - \mathbf{a}\rangle + \frac{1}{2}\langle \text{Hess}\,(q)(\mathbf{a})(\mathbf{x} - \mathbf{a}), \mathbf{x} - \mathbf{a}\rangle.$$

En particulier, si $\mathbf{a}$ est un point stationnaire de $q(\mathbf{x})$, i.e. $A\mathbf{a} = 0$, on a

$$q(\mathbf{x}) = q(\mathbf{a}) + \frac{1}{2}\langle \text{Hess}\,(q)(\mathbf{a})(\mathbf{x} - \mathbf{a}), \mathbf{x} - \mathbf{a}\rangle.$$

Le point $\mathbf{a} = \mathbf{0}$ est toujours un point stationnaire. Pour étudier la nature des points stationnaires on introduit la terminologie suivante :

Définition. Soit $A \in M_{n,n}(\mathbb{R})$ une matrice symétrique. On dit que la matrice A est positive (respectivement définie positive) si la forme quadratique associée

$$q(\mathbf{x}) = \frac{1}{2}\langle A\mathbf{x}, \mathbf{x}\rangle$$

satisfait $q(\mathbf{x}) \geq 0$ (respectivement $q(\mathbf{x}) > 0$) pour tout $\mathbf{x} \neq \mathbf{0}$. Elle est dite négative (respectivement définie négative) si $q(\mathbf{x}) \leq 0$ (respectivement $q(\mathbf{x}) < 0$) pour tout $\mathbf{x} \neq \mathbf{0}$. La matrice A est dite indéfinie si A n'est ni négative ni positive.

Lien avec les valeurs propres. Pour tout matrice $A \in M_{n,n}(\mathbb{R})$ symétrique il existe une matrice orthogonale P telle que $P^{-1}AP = D$ où D est la matrice diagonale

$$D = diag(\lambda_1, \ldots, \lambda_n) = \sum_{i=1}^{n} \lambda_i E_{ii} \quad \text{et} \quad \lambda_1 \leq \ldots \leq \lambda_n.$$

Les $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres de A. Donc nous avons les équivalences suivantes :

$$\begin{array}{rcc} A \geq 0 & \Leftrightarrow & 0 \leq \lambda_1 \leq \ldots \leq \lambda_n \\ A > 0 & \Leftrightarrow & 0 < \lambda_1 \leq \ldots \leq \lambda_n \\ A \leq 0 & \Leftrightarrow & \lambda_1 \leq \ldots \leq \lambda_n \leq 0 \\ A < 0 & \Leftrightarrow & \lambda_1 \leq \ldots \leq \lambda_n < 0 \\ A \text{ indéfinie} & \Leftrightarrow & \lambda_1 < 0 < \lambda_n \end{array}$$

De plus

$$\lambda_1\langle \mathbf{x}, \mathbf{x}\rangle \leq \langle A\mathbf{x}, \mathbf{x}\rangle \leq \lambda_n\langle \mathbf{x}, \mathbf{x}\rangle$$

pour tout $\mathbf{x} \in \mathbb{R}^n$.

Exemple- le cas $n = 2$. Soit $A \in M_{2,2}(\mathbb{R})$ une matrice symétrique. Posons

$$A = \begin{pmatrix} p & q \\ q & r \end{pmatrix}.$$

Nous avons $\operatorname{tr} A = p + r = \lambda_1 + \lambda_2$ et $\det A = pr - q^2 = \lambda_1\lambda_2$. Par conséquent, pour une matrice $A \in M_{2,2}(\mathbb{R})$,

1. $A \geq 0$ si et seulement si $\det A \geq 0$ et $\operatorname{tr} A \geq 0$.
2. $A > 0$ si et seulement si $\det A > 0$ et $\operatorname{tr} A \geq 0$.
3. $A \leq 0$ si et seulement si $\det A \geq 0$ et $\operatorname{tr} A \leq 0$.
4. $A < 0$ si et seulement si $\det A > 0$ et $\operatorname{tr} A \leq 0$.
5. A est indéfinie si et seulement si $\det A < 0$.

Extremums locaux - conditions suffisantes.

1. Si $A < 0$, alors $\mathbf{0}$ est un maximum local stricte de $q(\mathbf{x})$.
2. Si $A > 0$, alors $\mathbf{0}$ est un minimum local stricte de $q(\mathbf{x})$.
3. Si A est indéfinie, alors $\mathbf{0}$ est un point selle de $q(\mathbf{x})$.

Remarque. Au ch. 5.3 nous allons étendre ces trois conditions aux points stationnaires des fonctions réelles de classe C^2 où la matrice A correspond à la matrice hessienne en ces points stationnaires.

Remarque. La forme quadratique $q(\mathbf{x})$ admet un point stationaire $\mathbf{a} \neq \mathbf{0}$ si et seulement si A a une valeur propre 0. Par conséquent, si $A < 0$ (respectivement $A > 0$), $\mathbf{a} = \mathbf{0}$ est le point stationnaire unique de $q(\mathbf{x})$ et par conséquent, la forme quadratique admet un extremum global en 0.

Remarque. Pour les formes quadratiques les conditions affaiblies $A \leq 0$ (respectivement $A \geq 0$) impliquent que A admet un maximum locale (respectivement minimum locale). Par contre, pour des fonctions générales ces conditions ne sont plus suffisantes. (voir ci-dessous).

5.3 Développement limité

Théorème - développement limité d'ordre 2. Soit $U \subset \mathbb{R}^n$ ouvert et$f : U \longrightarrow \mathbb{R}$ de classe $C^2(U)$. Alors, pour tous $\mathbf{a}, \mathbf{a} + \mathbf{h} \in U$:

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{h} \rangle + \frac{1}{2}\langle (\operatorname{Hess} f)(\mathbf{a})\mathbf{h}, \mathbf{h} \rangle + o(||\mathbf{h}||_2^2).$$

Démonstration. Pour $\mathbf{h}$ fixe on considère la fonction $g : [0, 1] \longrightarrow \mathbb{R}$ de classe C^2 définie par

$$g(t) = f(\mathbf{a} + t\mathbf{h})$$

Par le théorème du développement limité des fonctions d'une variable, il existe $\theta \in [0, 1]$ tel que

$$g(1) = g(0) + g'(0) + \frac{g''(0)}{2} + R_2(1)$$

et

$$R_2(1) = \frac{g''(\theta) - g''(0)}{2}(1-0)^2 = \frac{g''(\theta) - g''(0)}{2}.$$

Par définition de g on a $g(0) = f(\mathbf{a})$ et $g(1) = f(\mathbf{a}+\mathbf{h})$. Ensuite on calcul les dérivées de la fonction $g(t)$:

$$g'(t) = \langle \nabla f(\mathbf{a}+t\mathbf{h}), \mathbf{h}\rangle \quad g'(0) = \langle \nabla f(\mathbf{a}), \mathbf{h}\rangle$$

$$g''(t) = \langle (\mathrm{Hess}\, f)(\mathbf{a}+t\mathbf{h})\mathbf{h}, \mathbf{h}\rangle$$

et

$$o(||\mathbf{h}||_2^2) = \frac{1}{2}\langle ((\mathrm{Hess}\, f)(\mathbf{a}+\theta\mathbf{h}) - (\mathrm{Hess}\, f)(\mathbf{a}))\mathbf{h}, \mathbf{h}\rangle.$$

grâce à la continuité de $(\mathrm{Hess}\, f)$.

Remarque. On peut obtenir le développement limité d'ordre k, $k > 2$ à partir de la fonction $g(t)$.

5.4 Extremums locaux - conditions suffisantes

Le le théorème du développement limité nous permet de formuler des conditions suffisantes pour qu'un point stationnaire soit un extremum local.

Théorème - conditions suffisantes. Soit $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction de classe $C^2(\mathbb{R}^n)$. Soit $\mathbf{a} \in \mathbb{R}^n$ un point stationnaire de f, i.e. $\nabla f(\mathbf{a}) = \mathbf{0}$ et notons $A = (\mathrm{Hess}\, f)(\mathbf{a})$. Alors,

1. Si $A < 0$, alors $\mathbf{a}$ est un maximum local stricte de $f(\mathbf{x})$.
2. Si $A > 0$, alors $\mathbf{a}$ est un minimum local stricte de $f(\mathbf{x})$.
3. Si A est indéfinie, alors $\mathbf{a}$ est un point selle de $f(\mathbf{x})$.

Ce résultat permet de trouver les extremums relatifs d'une fonction sur $\mathbb{R}^n$ ou dans l'intérieur d'une région. Lorsqu'il s'agit de trouver les extremums sur un ensemble fermé il faut étudier aussi le comportement sur le bord. On rappelle au résultat suivant pour des fonctions continues (chap.1, p.13) :

Théorème. Soit $U \subset \mathbb{R}^n$ un ensemble borné et fermé et $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction continue sur U. Alors f atteint son maximum et son minimum sur U.

Remarque. Pour une fonction de classe $C^k(U)$ les extremums se trouvent soit parmi les points stationnaire dans l'intérieur de U soit sur le bord ∂U.

Exemple - problème 1. Déterminer la nature des points stationnaires de la fonction

$$f(x,y) = x^3 - 6x^2 + \frac{1}{8}y^3 - 6y.$$

Les points stationnaires sont donnés par la condition $\nabla f = 0$, i.e.

$$3x^2 - 12x = 0 \quad \text{et} \quad \frac{3}{8}y^2 - 6 = 0.$$

Il y a quatre points stationnaires de $f(x,y)$:

$$P_1 = (0,4), P_2 = (0,-4), P_3 = (4,4), P_4 = (4,-4)$$

On a

$$\text{Hess}\,(f)(x,y) = \begin{pmatrix} 6x-12 & 0 \\ 0 & \frac{3}{4}y \end{pmatrix}.$$

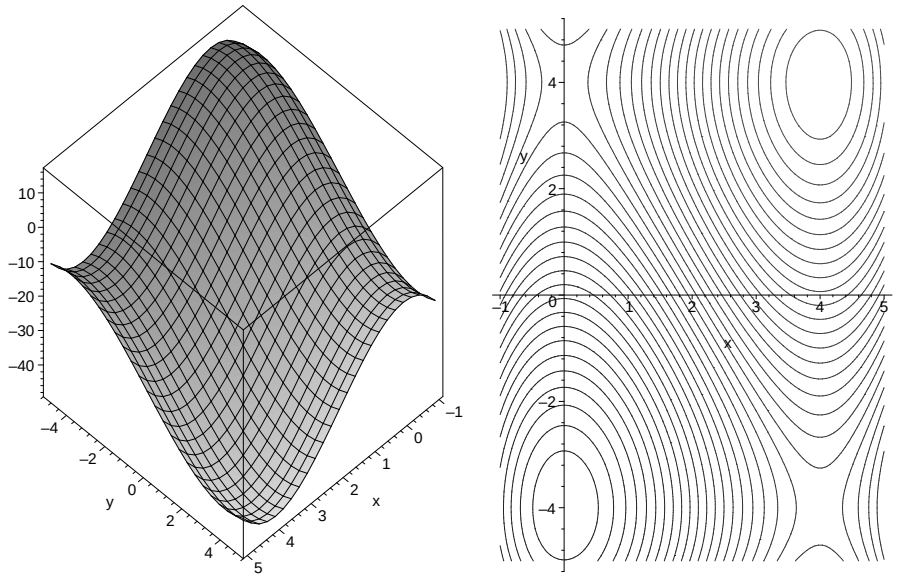
et par conséquent

$$\text{Hess}\,(f)(P_1) = \begin{pmatrix} -12 & 0 \\ 0 & 3 \end{pmatrix}, \quad f \text{ admet un point selle en } P_1, \quad f(P_1) = -16$$

$$\text{Hess}\,(f)(P_2) = \begin{pmatrix} -12 & 0 \\ 0 & -3 \end{pmatrix}, \quad f \text{ admet un maximum local stricte en } P_2, \quad f(P_2) = 16$$

$$\text{Hess}\,(f)(P_3) = \begin{pmatrix} 12 & 0 \\ 0 & 3 \end{pmatrix}, \quad f \text{ admet un minimum local stricte en } P_3, \quad f(P_3) = -48$$

$$\text{Hess}\,(f)(P_4) = \begin{pmatrix} 12 & 0 \\ 0 & -3 \end{pmatrix}, \quad f \text{ admet un point selle en } P_4, \quad f(P_4) = -16$$



$$f(x,y) = x^3 - 6x^2 + \tfrac{1}{8}y^3 - 6y$$

Exemple - problème 2. Donner le maximum de le minimum de $f(x,y)$ sur le rectangle $R = [-2,5] \times [0,6]$.
Le points stationnaires P_1 (point selle) et P_3 (minimum local stricte) sont dans R. On doit étudier la fonction f sur le bord de R qui consiste de quatre segments.

1. $S_1 = \{-2\} \times [0,6]$. La fonction $f(x,y)$ restreinte sur S_1 est donnée par

$$f|_{S_1}(x,y) = f(-2,y) = f_1(y) = \frac{1}{8}y^3 - 6y - 32, \quad 0 \le y \le 6$$

On a $f_1'(y) = \frac{3}{8}y^2 - 6$. $f_1(y)$ admet le point stationnaire $y_1 = 4$ dans $[0,6]$ (minimum locale stricte). On doit également calculer $f_1(y)$ sur le bord de $[0,6]$, i.e en $y = 0$ et $y = 6$. On trouve

$$f_1(0) = -32, \quad f_1(4) = -48, \quad f_1(6) = -41.$$

2. $S_2 = [-2, 5] \times \{0\}$. La fonction $f(x, y)$ restreinte sur S_2 est donnée par

$$f|_{S_2}(x, y) = f(x, 0) = f_2(x) = x^3 - 6x^2, \quad -2 \leq x \leq 5$$

On a $f_2'(x) = 3x^2 - 12x$. $f_2(x)$ admet les points stationnaires $x_1 = 0$ et $x_2 = 4$ dans $[-2, 5]$ (maximum local stricte respectivement minimum locale stricte). On doit également calculer $f_2(x)$ sur le bord de $[0 - 2, 5]$, i.e en $x = -2$ et $x = 5$. On trouve

$$f_2(-2) = -32, \quad f_2(0) = 0, \quad f_2(4) = -32, \quad f_2(5) = -25.$$

3. $S_3 = \{5\} \times [0, 6]$. La fonction $f(x, y)$ restreinte sur S_3 est donnée par

$$f|_{S_3}(x, y) = f(5, y) = f_3(y) = \frac{1}{8}y^3 - 6y - 25, \quad 0 \leq y \leq 6$$

Noter que $f_3(y) = f_1(y) + 7$. Donc $f_3(y)$ admet le même point stationnaire et

$$f_3(0) = -25, \quad f_3(4) = -41, \quad f_3(6) = -34.$$

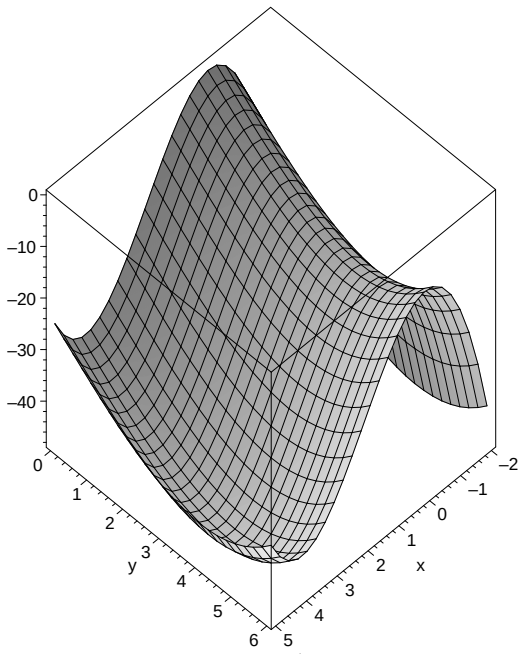
4. $S_4 = [-2, 5] \times \{6\}$. La fonction $f(x, y)$ restreinte sur S_4 est donnée par

$$f|_{S_4}(x, y) = f(x, 0) = f_4(x) = x^3 - 6x^2 - 9, \quad -2 \leq x \leq 5$$

Noter que $f_4(x) = f_2(x) - 9$. $f_4(x)$ admet les mêmes points stationnaires $x_1 = 0$ et $x_2 = 4$ dans $[-2, 5]$ que $f_2(x)$ et

$$f_2(-2) = -41, \quad f_2(0) = -9, \quad f_2(4) = -41, \quad f_2(5) = -34.$$

Par conséquent, $\max f|_R = 0$ atteint en $(0, 0)$ et $\min f|_R = -48$ atteint en P_1 et en $(-2, 4)$.



$f(x, y) = x^3 - 6x^2 + \frac{1}{8}y^3 - 6y$ sur le rectangle R.

Le cas de valeurs propres 0 **- Exemples.** Si une valeur popre de la matrice hessienne dans un point stationnaire est zéro le théorème sur les conditions suffisantes n'est plus applicable. Si $n \geq 3$ et si on a une valeur propre positive et une valeur propre négative, alors le point stationnaire est un point selle (par la définition du point selle). Par exemple, soit

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

Alors,

$$\nabla f(x, y, z) = \begin{pmatrix} 2x \\ 4y^3 \\ -2z \end{pmatrix}, \nabla f(0, 0, 0) = \mathbf{0}, \quad \text{Hess}\,(f)(0, 0, 0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

La matrice hessienne a les valeurs propre $2, 0, -2$. La fonction f admet un point selle en $\mathbf{0}$ puisque $f(x, 0, 0) > 0$ pour tout $x \neq 0$ et $f(0, 0, z) < 0$ pour tout $xz \neq 0$.
Si une valeur popre de la matrice hessienne dans un point stationnaire est zéro et les autres n'ont pas des signes opposées la situation est plus complexe. Par exemple la fonction $g(x, y) = x^2 + y^4$ admet un minimum local stricte (et même global) en $(0, 0)$. Les valeurs propres de la matrice hessienne en ce point stationnaire sont $0, 2$ et la fonction $h(x, y) = x^2 - y^4$ ayant le même point stationnaire et la même matrice hessienne en ce point admet un point selle en $(0, 0)$.

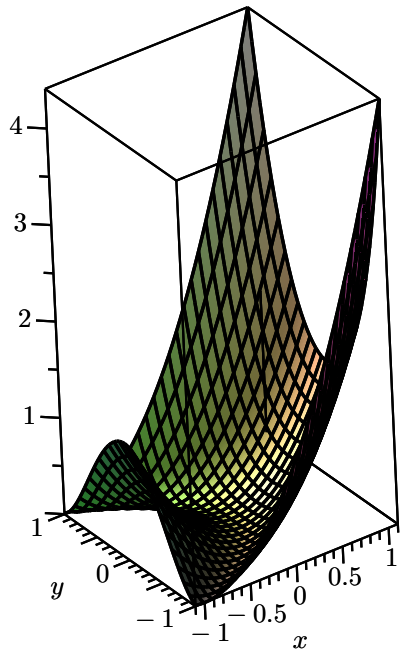
Finalement, soit $f(x, y) = x^2 + y^4 + 2xy^2$. On a

$$\nabla f(x, y) = \begin{pmatrix} 2(x + y^2) \\ 4(y^3 + xy) \end{pmatrix}, \quad \text{Hess}\,(f)(x, y) = \begin{pmatrix} 2 & 4y \\ 4y & 12y^2 + 4x \end{pmatrix}.$$

Les points stationnaires sont donnés par l'équation $x = -y^2$ et la matrice hessienne en ces points est

$$\text{Hess}\,(f)(-y^2, y) = \begin{pmatrix} 2 & 4y \\ 4y & 8y^2 \end{pmatrix}.$$

Elle admet des valuers propres 0 et $2 + 8y^2 > 0$. On ne peut pas déterminer la nature des points stationnaire par la matrice hessienne. Mais en notant que $f(x, y) = (x + y^2)^2 \geq 0$ on trouve que ces points sont de minimums locaux (mais pas strictes car la la courbe d'équation $x = -y^2$ est une courbe de niveau de f.



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

5.5 Extremums liés

Pour analyser la fonction sur le bord d'un rectangle (ou triangle) nous avons substitué l'équation du bord dans la fonction à étudier. Plus généralement, on cherche les extremums d'une fonction $h(\mathbf{x}) = h(x_1, \ldots, x_n)$ sur un ensemble donné par l'équation $f(\mathbf{x}) = 0$. Si on peut résoudre cette équation pour une variable réelle, par exemple $x_n = g(x_1, \ldots, x_{n-1})$, on peut substituer cette expression et étudier la fonction à $n-1$ variables réelles donnée par

$$\tilde{h}(x_1, \ldots, x_{n-1}) = h(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1}))$$

en appliquant la méthode présentée dans la section précédente. Souvent il est soit impossible de résoudre l'équation $f(\mathbf{x}) = 0$ pour une une variable réelle soit une telle substitution rend la fonction $\tilde{h}$ trop longue pour calculer facilement les dérivées. Pour ces situations il y a une autre méthode qui suit l'idée de la substitution en appliquant le théorème des fonctions implicites.

Théorème - Méthode des multiplicateurs de Lagrange I. Soit $U \subset \mathbb{R}^n$ ouvert. Soient $h, f : U \longrightarrow \mathbb{R}$ de classe $C^1(U)$ et

$$M = \{\mathbf{x} \in U : f(\mathbf{x}) = 0\}.$$

Soit $\mathbf{a} \in M$ tel que $\nabla f(\mathbf{a}) \neq \mathbf{0}$ et $h|_M$ admet un extremum locale en $\mathbf{a}$. Alors, il existe un $\lambda \in \mathbb{R}$ tel que

$$\nabla h(\mathbf{a}) + \lambda \nabla f(\mathbf{a}) = \mathbf{0}.$$

Démonstration. On peut supposer que $D_n f(\mathbf{a}) \neq 0$. Par le théorème des fonctions implicites il existe une unique fonction g définie dans un voisinage V de $(a_1, ..., a_{n-1}) \in \mathbb{R}^{n-1}$ telle que pour tout $(x_1, ..., x_{n-1}) \in V$:

$$f(x_1, ..., x_{n-1}, g(x_1, ..., x_{n-1})) = 0.$$

et

$$D_i f(\mathbf{a}) + D_i g(a_1, ..., a_{n-1}) D_n f(\mathbf{a}) = 0$$

pour tout $i = 1, ..., n-1$. On considère la fonction

$$H(x_1, ..., x_{n-1}) = h(x_1, ..., x_{n-1}, g(x_1, ..., x_{n-1})).$$

Par l'hypothèse, la fonction H adment un extremum local en $(a_1, ..., a_{n-1})$. Donc $D_i H(a_1, ..., a_{n-1}) = 0$ pour tout $i = 1, ..., n$. Les dérivées partielles de H en $(a_1, ..., a_{n-1})$ vérifont la rélation

$$D_i H(a_1, ..., a_{n-1}) = D_i h(\mathbf{a}) + D_i g(a_1, ..., a_{n-1}) D_n h(\mathbf{a}) = 0.$$

Si on pose

$$\lambda = -\frac{D_n h(\mathbf{a})}{D_n f(\mathbf{a})}$$

alors pour tout $i = 1, ..., n-1, n$:

$$D_i h(\mathbf{a}) + \lambda D_i f(\mathbf{a}) = 0.$$

Remarque. Le nombre λ est appelé le multiplicateur de Lagrange.

Remarque. Si M est borné les extremums de h existent et on les trouve comme suit :résoudre le système d'équations($n+1$ équations)

$$\nabla h(\mathbf{a}) + \lambda \nabla f(\mathbf{a}) = \mathbf{0}, \quad f(\mathbf{a}) = 0 \tag{5.1}$$

pour les $n+1$ inconnus $\mathbf{a}, \lambda$. Ensuite comparer les valeurs de h en ces points. Noter que ce théorème n'assure pas l'existence des extremums rélatifs. Les solutions de ce système ne sont pas forcement des extremums rélatifs.

Remarque. Si $\nabla f(\mathbf{a}) = \mathbf{0}$ pour un $\mathbf{a} \in M$ les extremums de $h|_M$ ne sont plus forcément parmi les solutions de (5.1). Autrement dit, un extremum local de $h|_M$ peut être en un tel point $\mathbf{a}$. Par exemple, considérer $f(x,y) = x^3 - y^7$ et $h(x,y) = x + y^2$. Alors, par substitution,

$$h|_M(x,y) = y|y|^{\frac{4}{3}} + y^2 > 0 \quad \text{si } y \neq 0$$

et $h(0,0) = 0$. Par conséquent, $h|_M$ admet un minumum local stricte en $(0,0) \in M$. Si on essaye d'appliquer la méthode des multiplicateurs de Lagrange par résolution du système (5.1) on doit résoudre les équations

$$1 + 3\lambda x^2 = 0, \quad 2y - 7\lambda y^6 = 0, \quad x^3 - y^7 = 0.$$

On constate que $(0,0)$ n'est pas une solution (ce système n'a pas de solution !). Donc la méthode des multiplicateurs de Lagrange ne fonctionne pas en les points stationnaires de f.

Théorème - Méthode des multiplicateurs de Lagrange II. Soit $U \subset \mathbb{R}^n$ ouvert. Soient $h, f_k : U \longrightarrow \mathbb{R}$, $k = 1, ..., p$ de classe $C^1(U)$ et

$$M_k = \{\mathbf{x} \in U : f_k(\mathbf{x}) = 0\}.$$

Soit $\mathbf{a} \in M_1 \cap, ..., \cap M_p$ tel que les vecteurs $\nabla f_k(\mathbf{a})$ sont linéairement indépendants et $h|_{M_1 \cap, ..., \cap M_p}$ admet un extremum locale en $\mathbf{a}$. Alors, il existe $\lambda_1, ..., \lambda_p \in \mathbb{R}$ tels que

$$\nabla h(\mathbf{a}) + \sum_{k=1}^{p} \lambda_k \nabla f_k(\mathbf{a}) = \mathbf{0}.$$

Exemple. Donner le maximum et le minimum de la fonction

$$h(x,y) = x^3 - 6x^2 + \frac{1}{8}y^3 - 6y.$$

sur $M = \{(x,y) \in \mathbb{R}^2 : x^2 + (y-2)(y-6) = 0\}$.
Noter d'abord que M est le cercle autour du point $(0,4)$ de rayon 2. M est borné et fermé. Par conséquent la fonction h atteint son minimum et son maximum sur M. De plus, avec $f(x,y) = x^2 + (y-2)(y-6)$ nous avons

$$\nabla f(x,y) = \begin{pmatrix} 2x \\ 2(y-4) \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \text{si} \quad \begin{pmatrix} x \\ y \end{pmatrix} \in M.$$

En effet, $\nabla f(x,y)$ s'annule si et seulement si $x=0$ et $y=4$, mais $(0,4)\notin M$. Nous pouvons appliquer le théorème, i.e. le minimum et le maximum de $h(x,y)$ se trouvent parmi les solutions du système d'équations

$$\nabla h(x,y)+\lambda\nabla f(x,y)=\mathbf{0},\quad f(x,y)=0.$$

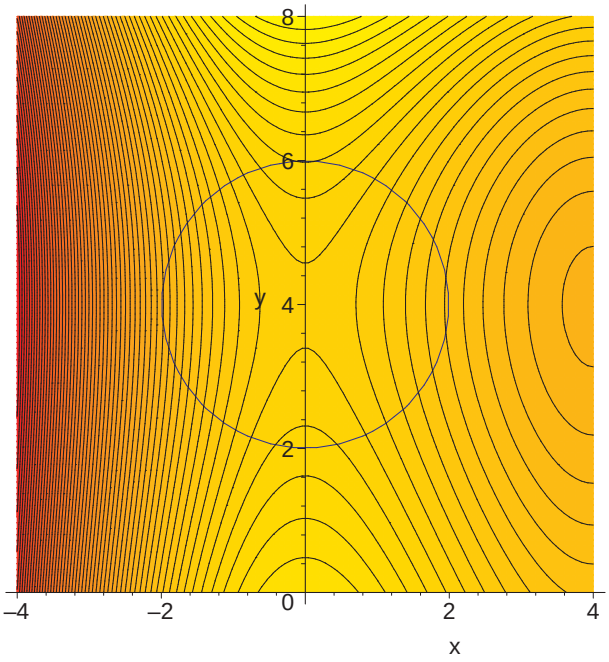
(et des telles solutions existent !). Le système d'équations s'écrit explicitement comme suit :

$$\begin{aligned}
&x(3x-12+2\lambda)=0\\
&(y-4)(\frac{3}{8}(y+4)+2\lambda)=0\\
&x^2+(y-2)(y-6)=0
\end{aligned}$$

Analysons ce système : Par la première équation nous avons soit $x=0$ ou $3x-12+2\lambda=0$. Si $x=0$, alors par la trosième équation $y=2$ (et donc $\lambda=-\frac{9}{8}$, mais il n'est pas néccessaire de calculer λ) ou $y=6$ (et donc $\lambda=-\frac{16}{8}$). Si $y=4$, alors par la trosième équation $x=2$ (et donc $\lambda=3$, mais de nouveau il n'est pas néccessaire de calculer λ) ou $x=-2$ (et donc $\lambda=9$). Il reste le cas $3x-12+2\lambda=\frac{3}{8}(y+4)+2\lambda=0$ donc $y=8x-36$ qui n'admet pas des solutions dans M. Nous avons alors les quatre solutions

$$\begin{aligned}
&(x,y)=(0,2)\quad\text{et}\quad h(0,2)=-11\\
&(x,y)=(0,6)\quad\text{et}\quad h(0,6)=-9\\
&(x,y)=(-2,4)\quad\text{et}\quad h(-2,4)=-48\\
&(x,y)=(2,4)\quad\text{et}\quad h(2,4)=-32
\end{aligned}$$

Donc $\min h|_M=-48$ et $\max h|_M=-9$.



$h(x,y)=x^3-6x^2+\frac{1}{8}y^3-6y$ sur M (cercle bleu). Dans les points d'extremums les lignes de niveau sont tangentielles au cercle.

Exemple. Donner le maximum et le minimum de la fonction

$$h(x, y, z) = x + y + z$$

sous les conditions $f_1(x, y, z) := x^2 + y^2 - 2 = 0$ et $f_2(x, y, z) := x + z - 1 = 0$. L'ensemble

$$M; +\{(x, y, z) \in \mathbb{R}^3 : f_1(x, y, z) = f_2(x, y, z) = 0\}$$

et borné et fermé. Par conséquent, la fonction $h(x, y, z)$ atteint son maximum et son minimum sur M. les vecteurs

$$\nabla f_1(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 0 \end{pmatrix}, \nabla f_2(x, y, z) = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

sont linéarement indépendants sur M. Le minimum et le maximum de $h(x, y, z)$ se trouvent parmi les solutions du système d'équations

$$\nabla h(x, y, z) + \lambda_1 \nabla f_1(x, y, z) + + \lambda_2 \nabla f_2(x, y, z) = \mathbf{0}, \quad f_1(x, y, z) = f_2(x, y, z) = 0.$$

(et des telles solutions existent !). Le système d'équations s'écrit explicitement comme suit :

$$\begin{aligned} &1 + 2\lambda_1 x + \lambda_2 = 0 \\ &1 + 2\lambda_1 y = 0 \\ &1 + \lambda_2 = 0 \\ &x^2 + y^2 - 2 = 0 \\ &x + z - 1 = 0 \end{aligned}$$

Analysons ce système : La première équation et la troisième èéquation nous donne $2\lambda_1 x = 0$ i.e. soit $x = 0$ ou $\lambda_1 = 0$. En vue de la deuxième équation ce dernier cas est impossible. Donc $x = 0$. La condition $f_1(x, y, z) = 0$ implique $y = \pm\sqrt{2}$ et la condition $f_1(x, y, z) = 0$ implique $z = 1$. De plus,

$$h(0, \sqrt{2}, 1) = 1 + \sqrt{2} \quad \text{et} \quad h(0, -\sqrt{2}, 1) = 1 - \sqrt{2}.$$

Par conséquent,$\min h|_M = 1 - \sqrt{2}$ et $\max h|_M = 1 + \sqrt{2}$.

Remarque. Pour le dernier exemple il y a une solution plus rapide : on peut directement insérer la relation $z = 1 - x$ dans h et chercher les extremums de $\tilde{h}(x, y) = y + 1$ sous la contrainte $x^2 + y^2 = 2$ pour obtenir la solution sans calcul.

Chapitre 6

Intégrales multiples

6.1 Intégrales qui dépendent d'un paramètre

Proposition 1. Soient $a < b$ et I un intervalle ouvert et $f : [a, b] \times I \longrightarrow \mathbb{R}$ une fonction continue. Alors, la fonction $g : I \longrightarrow \mathbb{R}$ définie par

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

est continue. Si f est continue et sa dérivée partielle $\frac{\partial f(x,y)}{\partial y}$ est continue, alors g est de classe $C^1(I)$ et, de plus, pour tout $y \in I$:

$$g'(y) := \int_a^b \frac{\partial f(x, y)}{\partial y}\, dx.$$

Démonstration - continuité. L'ingrédient principal de la démonstration est le fait que la fonction f est uniformément continue sur les domaines $[a, b] \times I_0$ pour tous $I_0 \subset I$ borné et fermé (voir le chapitre 1). Soit $y \in I$. Alors, il existe $I_0 \subset I$ borné et fermé tel que $y \in \overset{\circ}{I_0}$ (l'intérieur de I_0). Grâce à la continutité uniforme de f dans $[a, b] \times I_0$, pour tous $\epsilon > 0$, il existe $\delta > 0$ tel que pour tous $|h| < \delta$ pour tous $x \in [a, b]$:

$$(x, y), (x, y + h) \in [a, b] \times I_0 \quad \text{et} \quad |f(x, y + h) - f(x, y)| < \epsilon.$$

Par conséquent, pour tous $\epsilon > 0$, il existe $\delta > 0$ tel que pour tous $|h| < \delta$: $y + h \in I_0$ et

$$|g(y + h) - g(y)| \leq \int_a^b |f(x, y + h) - f(x, y)|\, dx \leq \epsilon(b - a)$$

prouvant la continutité de g puisque ϵ peut être choisi arbitrairement petit.

Démonstration - dérivabilité. On utilise la continuité uniforme de f et de $\partial f/\partial y$ sur les domaines $[a, b] \times I_0$. par le théorème des accroissements finis :

$$\frac{g(y + h) - g(y)}{h} = \int_a^b \frac{f(x, y + h) - f(x, y)}{h}\, dx = \int_a^b \frac{\partial f(x, y + h\theta_h)}{\partial y}\, dx$$

pour un $\theta_h \in]0,1[$. On écrit

$$\int_a^b \frac{\partial f(x,y+h\theta_h)}{\partial y}\,dx = \int_a^b \frac{\partial f(x,y)}{\partial y}\,dx + \int_a^b \frac{\partial f(x,y+h\theta_h)}{\partial y} - \frac{\partial f(x,y)}{\partial y}\,dx.$$

pour établir pour la dernière intégrale une estimation comme dans la première partie.

Applications. Cette proposition nous donne une nouvelle technique pour calculer des intégrales comme l'exemple suivant le montre :

Exemple 1. On veut calculer

$$\int_0^b x^2 \cos x\,dx$$

Au lieu d'effectuer des intégrations par partie (voir semestre 1 chapitre 6) on considère la fonction

$$g(y) := \int_0^b \cos xy\,dx$$

sur un intervalle ouvert contenant $y = 1$, par exemple $]1/2, 2[$. En appliquant la proposition 1d'abord à $g(y)$:

$$g'(y) := -\int_0^b x \sin xy\,dx$$

et ensuite à $g'(y)$:

$$g''(y) = -\int_0^b x^2 \cos xy\,dx,$$

donc

$$\int_0^b x^2 \cos x\,dx = -g''(1)$$

D'autre part on peut déterminer la fonction g explicitement :

$$g(y) = \frac{\sin xy}{y}\bigg|_0^b = \frac{\sin by}{y}.$$

Les dérivées de g sont données par

$$g'(y) = -\frac{\sin by}{y^2} + \frac{b\cos by}{y}$$

et

$$g''(y) = \frac{2\sin by}{y^3} - \frac{2b\cos by}{y^2} - \frac{b^2 \sin by}{y^2}.$$

Par consêquent,

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

Intégrales généralisées. La proposition 1 s'étend aux intégrales généralisées sur $[a,\infty[$ si l'on contrôle en plus le comportement de $f(x,y)$ et sa dérivée partielle $\frac{\partial f(x,y)}{\partial y}$ lorsque x tend vers l'infini :

Proposition 2. Soient I un intervalle ouvert et $f : [a,\infty[\times I \longrightarrow \mathbb{R}$ une fonction continue. Soit $\phi : [a,\infty[\longrightarrow \mathbb{R}_+$ une fonction dont l'intégrale généralisée

$$\int_a^\infty \phi(x)\,dx$$

converge telle que $|f(x,y)| \leq \phi(x)$ pour $(x,y) \in [a,\infty[\times I$.

Alors, la fonction $g : I \longrightarrow \mathbb{R}$ définie par

$$g(y) := \int_a^\infty f(x,y)\,dx$$

est continue. Si f est continue et sa dérivée partielle $\frac{\partial f(x,y)}{\partial y}$ est continue et satisfait $|\frac{\partial f(x,y)}{\partial y}| \leq \psi(x)$ pour $(x,y) \in [a,\infty[\times I$ pour une fonction $\psi : I \longrightarrow \mathbb{R}_+$ une fonction dont l'intégrale généralisée sur $[a,\infty[$ existe, alors g est de classe $C^1(I)$ et, de plus, pour tout $y \in I$:

$$g'(y) := \int_a^\infty \frac{\partial f(x,y)}{\partial y}\,dx.$$

Exemple 2. On veut calculer

$$\int_0^\infty x^3 e^{-x^2} : dx$$

On pose $I =]1/2, 2[$ et $f(x,y) = xe^{-yx^2}$. La fonction f satisfait les hypothèses de la proposition avec $\phi(x) = xe^{-x^2/2}$ et $\psi(x) = x^3 e^{-x^2/2}$. Donc

$$g(y) = \int_0^\infty xe^{-yx^2}\,dx.$$

est de class C^1 et

$$g'(y) = -\int_0^\infty x^3 e^{-yx^2}\,dx.$$

Notons que

$$g(y) = -\int_0^\infty \frac{d}{d\,x}\left(\frac{e^{-yx^2}}{2y}\right)dx = \frac{1}{2y}.$$

Donc $g'(y) = \frac{-1}{2y^2}$ et

$$\int_0^\infty x^3 e^{-x^2}\,dx = -g'(1) = \frac{1}{2}.$$

Intégrands et bornes d'intégration avec un paramètre. Si les bornes d'intégration dépendent également du paramètre y de classe C^1 on peut facilement généraliser le résultat du chapitre 6.4.2 de l'Analyse 1 aux intégrales qui dépendent du paramètre y.

Proposition 3. Soient I, J deux intervalles ouverts et $f : J \times I \longrightarrow \mathbb{R}$ une fonction de classe C^1 et $a, b : I \longrightarrow J$ deux fonctions de classe C^1. Alors, la fonction $g : I \longrightarrow \mathbb{R}$ définie par

$$g(y) := \int_{a(y)}^{b(y)} f(x, y)\, dx$$

est de classe $C^1(I)$ et, de plus, pour tout $y \in I$:

$$g'(y) := f(b(y), y)b'(y) - f(a(y), y)a'(y) + \int_{a(y)}^{b(y)} \frac{\partial f(x, y)}{\partial y}\, dx.$$

Démonstration :. Noter simplement que pour tout $y, y + h \in I$:

$$\begin{aligned}\frac{g(y+h) - g(y)}{h} =& \frac{1}{h}\int_{b(y)}^{b(y+h)} f(x, y+h)\, dx - \frac{1}{h}\int_{a(y)}^{a(y+h)} f(x, y+h)\, dx \\ &+ \frac{1}{h}\int_{a(y)}^{b(y)} f(x, y+h) - f(x, y)\, dx.\end{aligned}$$

L'affirmation suit de la continuité uniforme de f et du théorème de la valeur moyenne.

Exemple 3. Calculer

$$\lim_{y\to 0}\int_0^y \frac{(y-x)\cos x^2}{\sin y^2}\, dx.$$

On définit

$$g(y) = \int_0^y (y-x)\cos x^2\, dx.$$

Alors

$$g'(y) = \int_0^y \cos x^2\, dx, \quad g''(y) = \cos y^2.$$

Notons que $(\sin y^2)'' = (2y\cos y^2)' = 2\cos y^2 + 4y^2 \sin y^2$. Donc

$$\lim_{y\to 0} \frac{g''(y)}{(\sin y^2)''} = \frac{1}{2}$$

et par la règle de l'Hopital :

$$\lim_{y\to 0}\int_0^y \frac{(y-x)\cos x^2}{\sin y^2}\, dx \equiv \lim_{y\to 0}\frac{g(y)}{\sin y^2} = \lim_{y\to 0}\frac{g''(y)}{(\sin y^2)''} = \frac{1}{2}.$$

Extension à n paramètres. Dans les propositions 1 -3 on peut remplacer le paramètre y par n paramètres $\mathbf{y}$ en passant aux dérivées partielles de

$$g(\mathbf{y}) := \int_a^b f(x, \mathbf{y})\, dx \quad \text{ou} \quad g(\mathbf{y}) := \int_{a(\mathbf{y})}^{b(\mathbf{y})} f(x, \mathbf{y})\, dx.$$

Par exemple,

$$\frac{\partial g(\mathbf{y})}{\partial y_k} = \int_a^b \frac{\partial f(x, \mathbf{y})}{\partial y_k}\, dx.$$

6.2 Intégrales doubles

Soient $D \subset \mathbb{R}^2$ et $f : D \longrightarrow \mathbb{R}$ une fonction continue. On définit l'intégrale double

$$\iint_D f(x,y)\,dxdy.$$

pour des quelques classes de domaines D sans vouloir faire une théorie d'intégration.

6.2.1 Intégrale double sur un rectangle fermé

Soient $a < b$ et $c < d$ et $f : [a,b] \times [c,d] \longrightarrow \mathbb{R}$ une fonction continue. Alors, par la proposition 1, les fonctions $g : [c,d] \longrightarrow \mathbb{R}$ et $h : [a,b] \longrightarrow \mathbb{R}$ définies par

$$g(y) := \int_a^b f(x,y)\,dx \quad \text{et} \quad h(x) := \int_c^d f(x,y)\,dy$$

sont continues. Alors, les inteégrales

$$\int_c^d g(y)\,dy, \quad \int_a^b h(x)\,dx$$

existent. De plus, on a le

Théorème 1 - Théorème de Fubini pour des fonctions continues.

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

On peut donc définir l'intégrale double sur le rectangle fermé $D = [a,b] \times [c,d]$ par

$$\begin{aligned}\iint_D f(x,y)\,dxdy &= \int_c^d g(y)\,dy = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy \\ &= \int_a^b h(x)\,dx = \int_a^b \left(\int_c^d f(x,y)\,dy. \right) dx\end{aligned}$$

Démonstration. On définit la fonction $\psi : [c,d] \longrightarrow \mathbb{R}$ par

$$\psi(x,t) = \int_c^t f(x,y)\,dy.$$

et la fonction $\phi : [c,d] \longrightarrow \mathbb{R}$ par

$$\phi(t) = \int_a^b \psi(x,t)\,dx = \int_a^b \left(\int_c^t f(x,y)\,dy \right) dx.$$

On a $\psi(x,c) \equiv 0$ et donc $\phi(c) = 0$. Par la proposition 1 la fonction $\psi(x,t)$ est continue. Par le théorème fondamental du calcul intégrale sa dérivée partielle par rapport à t continue. Alors, par la proposition 1, la fonction $\phi(t)$ est dérivable et

$$\phi'(t) = \int_a^b \frac{\partial}{\partial t}\psi(x,t)\,dx = \int_a^b \frac{\partial}{\partial t}\left(\int_c^t f(x,y)\,dy \right) dx = \int_a^b f(x,t)\,dx = g(t)$$

Par conséquent,

$$\int_c^d g(t)\,dt = \int_c^d \phi'(t)\,dt = \phi(d) = \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \int_a^b h(x)\,dx.$$

L'intégrale double vérifie les proprietés de linéarité et monotonie.

Théorème 2. Soient $D = [a,b] \times [c,d]$ et $f, g : D \to \mathbb{R}$ deux fonctions continues. Alors, pour tout $\alpha, \beta \in \mathbb{R}$:

$$\iint_D (\alpha f(x,y) + \beta g(x,y))\,dxdy = \alpha \iint_D f(x,y)\,dxdy + \beta \iint_D g(x,y)\,dxdy \quad (6.1)$$

Si $f(x,y) \leq g(x,y)$ pour tout $(x,y) \in D$, alors

$$\iint_D f(x,y)\,dxdy \leq \iint_D g(x,y)\,dxdy. \quad (6.2)$$

$$\left| \iint_D f(x,y)\,dxdy \right| \leq \iint_D |f(x,y)|\,dxdy. \quad (6.3)$$

Théorème de la valeur moyenne. Soient $D = [a,b] \times [c,d]$ et $f, g : D \to \mathbb{R}$ deux fonctions continues et $g \geq 0$. Alors, il existe (au moins) un $(x_0, y_0) \in D$ tel que

$$\iint_D f(x,y)g(x,y)\,dxdy = f(x_0,y_0) \iint_D g(x,y)\,dxdy$$

En particulier, si $g \equiv 1$:

$$\iint_D f(x,y)\,dxdy = f(x_0,y_0)\text{Aire}(D) = f(x_0,y_0)|D|.$$

Par conséquent,

$$\iint_D dxdy = \text{Aire}(D) = |D|.$$

Exemples.

1. Soient $D = [a,b] \times [c,d]$ et $f : [a,b] \to \mathbb{R}$, $g : [c,d] \to \mathbb{R}$ deux fonctions continues. Alors,

   $$\iint_D f(x)g(y)\,dxdy = \int_a^b f(x)\,dx \cdot \int_c^d g(y)\,dy.$$

   Par exemple

   $$\iint_{[0,1]\times[0,2]} e^{x+2y}\,dxdy = \int_0^1 e^x\,dx \cdot \int_0^2 e^{2y}\,dy = (e-1)\cdot(\frac{e^4-1}{2}).$$

2. $$\iint_{[0,\pi/2]\times[0,\pi/2]} \sin(x+y)\,dxdy = \int_0^{\pi/2} (-\cos(x+\pi/2) + \cos x)\,dx$$
   $$= \int_0^{\pi/2} \cos x + \sin x\,dx = 2$$

Interprétation de l'intégrale double. Soit $D = [a, b] \times [c, d]$ et $f : D \to \mathbb{R}_+$ une fonction continue. On définit

$$E = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D : \text{ et } 0 \leq z \leq f(x, y)\}$$

Alors, le volume de E, noté $\text{Vol}(E)$, est donné par

$$\text{Vol}(E) = \iint_D f(x, y)\, dxdy.$$

Exemple - volume d'une pyramide. Soit $a > 0$ et $D = [-a/2, a/2] \times [-a/2, a/2]$ et $f : D \to \mathbb{R}_+$ définie par

$$f(x, y) = h(1 - \max(2|x|/a, 2|y|/a))$$

L'ensemble $E \in \mathbb{R}^3$ défini ci-dessus décrit une pyramide de hauteur h de base D. Pour calculer le volume de E on note que

$$\text{Vol}(E) = 4h \iint_{[0,a/2]\times[0,a/2]} f(x, y)\, dxdy.$$

Donc

$$\begin{aligned}
\text{Vol}(E) &= 4h \int_0^{a/2} \left(\int_0^y (1 - 2y/a)\, dx + \int_y^{a/2} (1 - 2x/a)\, dx \right) dy \\
&= 4h \int_0^{a/2} \left(y(1 - 2y/a) + (a/2 - y) - (\frac{(a/2)^2}{a} - \frac{y^2}{a}) \right) dy \\
&= 4h \int_0^{a/2} \frac{a}{4} - \frac{y^2}{a}\, dy \\
&= 4h \left(\frac{a^2}{8} - \frac{a^2}{24} \right) \\
&= \frac{ha^2}{3}
\end{aligned}$$

Rectangles fermés quasi disjoints. Un sous-ensemble D de $\mathbb{R}^2$ est appelé rectangle fermé s'il est donné par le produit cartésien de deux intervalles bornés et fermés : $D = [a, b] \times [c, d]$. On appelle $(D_k)_{k\in\mathbb{N}}$ une famille de rectangles fermés quasi disjoints si pour tout couple d'entiers naturels $n \neq m$:

$$\overset{\circ}{D}_n \cap \overset{\circ}{D}_m = \emptyset.$$

On a pour toute fonction continue $f : \mathbb{R}^{\not\vDash} \longrightarrow \mathbb{R}$:

$$\iint_{\bigcup_k D_k} f(x, y)\, dxdy = \sum_k \iint_{D_k} f(x, y)\, dxdy.$$

Si D est la rèunion des rectangles d'une famille des rectangles fermés quasi disjoints et $f, g : D \longrightarrow \mathbb{R}$ sont deux fonctions continues bornées telles que $f < g$ sur D. Alors, le volume de

$$E = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D : \text{ et } f(x, y) \leq z \leq g(x, y)\}$$

est donné par

$$\mathrm{Vol}(E) = \iint_D g(x,y) - f(x,y)\, dxdy.$$

Cette extension permets de traiter des intégrales généralisées comme, par exemple, f continue sur un rectangle ouvert :

$$\iint_{[0,1]\times[0,1]} \frac{1}{\sqrt{xy}}\, dxdy = \int_0^1 \frac{1}{\sqrt{x}}\, dx \int_0^1 \frac{1}{\sqrt{y}}\, dy = 4.$$

Remarquons encore que l'intégrale double sur ces domaines vérifie les propriétés de linéarité et de monotonie (6.1) - (6.3) et le théorème de la valeur moyenne.

Calcul des intégrale doubles. La technique introduit dans l'exemple precedent se généralise comme suit : Soient $\phi, \psi : [a,b] \longrightarrow \mathbb{R}$ deux fonctions continues telles que $\phi(x) < \psi(x)$ pour tout $x \in]a,b[$. On considère l'ensemble ouvert borné D défini par

$$D = \{(x,y) \in \mathbb{R}^2 : x \in]a,b[: \text{ et } \phi(x) < y < \psi(x)\}$$

Alors, pour toute fonction continue

$$f : \bar{D} = \{(x,y) \in \mathbb{R}^2 : x \in [a,b] : \text{ et } \phi(x) \leq y \leq \psi(x)\} \longrightarrow \mathbb{R},$$

on a

$$\iint_D f(x,y)\, dxdy = \int_a^b \Big(\int_{\phi(x)}^{\psi(x)} f(x,y)\, dy \Big) dx.$$

Pour toute $t \in]a,b[$ la fonction $A :]a,b[\longrightarrow \mathbb{R}$ définie par

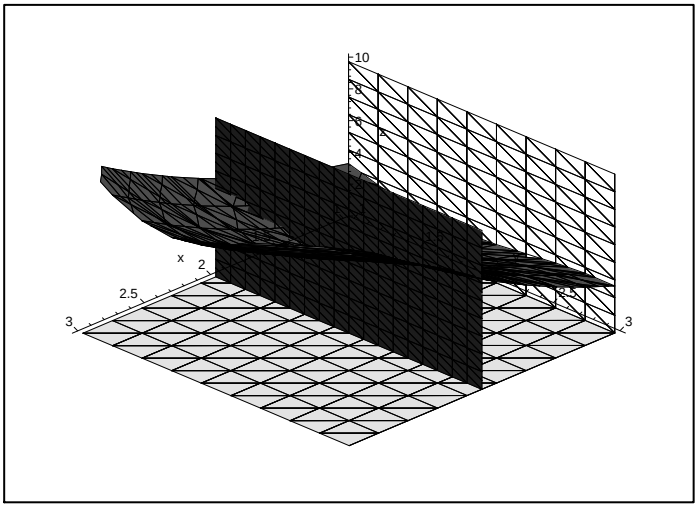
$$A(t) = \int_{\phi(t)}^{\psi(t)} f(t,y)\, dy$$

donne la surface plane obtenue en coupant le corps

$$E = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in D : \text{ et } 0 \leq z \leq f(x,y)\}$$

par le plan d'équation $x = t$. Le volume de E est

$$\mathrm{Vol}(E) = \int_a^b A(t)\, dt.$$



Plus généralement, si $g : \bar{D} \longrightarrow \mathbb{R}$ est une autre fonction continue telle que $f < g$ sur $D.$, alors pour toute $t \in]a, b[$ la fonction $A :]a, b[\longrightarrow \mathbb{R}$ définie par

$$A(t) = \int_{\phi(t)}^{\psi(t)} (g(t, y) - f(t, y))\, dy$$

donne la surface plane obteneu en coupant le corps

$$E = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D : \text{ et } f(x, y) \leq z \leq g(x, y)\}$$

par le plan d'équation $x = t$. Le volume de E est

$$\text{Vol}(E) = \int_a^b A(t)\, dt.$$

Dans l'exemple de la pyramide on peut couper par le plan d'équation $y = t$. On a $A(t) = (1 - \frac{t}{h})^2 a^2$, $t \in [0, h]$ d'où

$$\text{Vol}(E) = a^2 \int_0^h (1 - \frac{t}{h})^2\, dt = a^2 h \int_0^1 (1 - s)^2\, ds = \frac{A(0)h}{3}.$$

Exemple - domaine triangulaire. Soit $D \subset \mathbb{R}^2$ le triangle donné par

$$\begin{aligned} D &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 1,\ 0 \leq y \leq x\} \\ &= \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq 1,\ y \leq x \leq 1\} \end{aligned}$$

Alors, pour toute fonction continue sur D :

$$\begin{aligned} \iint_D f(x, y)\, dxdy &= \int_0^1 \left(\int_0^x f(x, y)\, dy \right) dx \\ &= \int_0^1 \left(\int_y^1 f(x, y)\, dx \right) dy \end{aligned}$$

Si $f(x, y) = \sinh x^2$ la seconde expression pour l'intégrale double sur D est difficilement applicable car il faut connaitre la fonction primitive de $\sinh x^2$. Par contre,

$$\iint_D f(x, y)\, dxdy = \int_0^1 \left(\int_0^x \sinh x^2\, dy \right) dx = \int_0^1 x \sinh x^2\, dx = \left. \frac{\cosh x^2}{2} \right|_0^1$$

Soit T est le triangle donné par les sommets $A = (0, 0)$, $B = (4, 0)$, $C = (0, 3)$. Alors,

$$\begin{aligned} T &= \{(x, y) \in \mathbb{R}^2 : 0 \leq x \leq 4,\ 0 \leq y \leq 3 - \frac{3x}{4}\} \\ &= \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq 3,\ 0 \leq x \leq 4 - \frac{4x}{3}\}. \end{aligned}$$

Alors,

$$\begin{aligned}\iint_T \frac{x}{3x+4y}\,dxdy &= \int_0^4 dx \int_0^{3-\frac{3x}{4}} dy \frac{x}{3x+4y} \\ &= \int_0^4 dx\,\frac{x}{4}\ln(3x+4y)\Big|_{y=0}^{y=3-\frac{3x}{4}} \\ &= \int_0^4 dx\,\frac{x}{4}(\ln 12 - \ln 3x) = \frac{1}{4}\int_0^4 dx\,(2\ln 2)x - x\ln x \\ &= \frac{(4\ln 2)x^2 - 2x^2\ln x + x^2}{16}\Big|_{x=0}^{x=4} = 1.\end{aligned}$$

Exemple - la paritédu domaine et de la fonction. Soit $D = B_1(\mathbf{0}) \subset \mathbb{R}^2$ la boule d'unité. Calculer

$$\iint_D x^2 \sin y\,dxdy.$$

On a

$$\iint_D x\,dxdy = \int_{-1}^1 x^2\Big(\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \sin y\,dy\Big)\,dx = 2\int_{-1}^1 0\,dx = 0.$$

Plus géneralement, soit D un domaine tel que $(x,y) \in D$ implique $(x,-y) \in D$. Si f est une fonction continue sur D et impaire en y, alors

$$\begin{aligned}\iint_D f(x,y)\,dxdy &= \iint_{D\cap\{y>0\}} f(x,y)\,dxdy + \iint_{D\cap\{y<0\}} f(x,y)\,dxdy \\ &= \iint_{D\cap\{y>0\}} f(x,y)\,dxdy + \iint_{D\cap\{y>0\}} f(x,-y)\,dxdy \\ &= \iint_{D\cap\{y>0\}} f(x,y)\,dxdy + \iint_{D\cap\{y>0\}} -f(x,y)\,dxdy = 0\end{aligned}$$

6.2.2 Intégrale double sur $\mathbb{R}^2$

Ici on définit l'intégrale double

$$\iint_{\mathbb{R}^2} f(x,y)\,dxdy$$

pour une fonction continue $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ par la limite des intégrales sur des rectangles bornés D_k tels que $D_k \to \mathbb{R}^2$ lorsque $k \to \infty$ (en analogie avec l'Analyse 1). Une propriété importante de l'intégrale double sur $\mathbb{R}^2$ est son invariance sous translations : pour tous $(x_0, y_0) \in \mathbb{R}^2$,

$$\iint_{\mathbb{R}^2} f(x-x_0, y-y_0)\,dxdy = \iint_{\mathbb{R}^2} f(x,y)\,dxdy.$$

Cette invariance sous translations reste valable si $x_0 = x_0(y)$ et $y_0 = y_0(x)$ sont des fonctions continues d'une variable.

Exemple. Pour calculer

$$I := \iint_{\mathbb{R}^2} \frac{e^{x-\frac{y^2}{2}}}{\left(1+e^{x+\frac{y^2}{2}}\right)^2}\, dxdy$$

on note d'abord que

$$\begin{aligned}\int_{\mathbb{R}} \frac{e^{x-\frac{y^2}{2}}}{\left(1+e^{x+\frac{y^2}{2}}\right)^2}\, dx &= e^{-y^2}\int_{\mathbb{R}} \frac{e^{x+\frac{y^2}{2}}}{\left(1+e^{x+\frac{y^2}{2}}\right)^2}\, dx \\ &= e^{-y^2}\int_{\mathbb{R}} \frac{e^{x}}{\left(1+e^{x}\right)^2}\, dx \\ &= e^{-y^2}\int_{\mathbb{R}} \frac{d}{dx}\big(\frac{-1}{1+e^{x}}\big)\, dx = e^{-y^2}.\end{aligned}$$

Par conséquent, par le changement de variable $y^2 = t$:

$$I = \int_{\mathbb{R}} e^{-y^2}\, dy = 2\int_0^{\infty} e^{-y^2}\, dy = \int_0^{\infty} t^{-\frac{1}{2}} e^{-t}\, dt = \Gamma(\frac{1}{2}) = \sqrt{\pi}.$$

6.2.3 Changement de variables dans $\mathbb{R}^2$

Une intérpretation du determinant. Soient $\mathbf{v}, \mathbf{w} \in \mathbb{R}^2$ linéairement indépendants. L'ensemble

$$D = \{s\mathbf{v} + t\mathbf{w}, 0 \le s \le 1, 0 \le t \le 1\}$$

d'écrit un parallélogramme dans $\mathbb{R}^2$. Son aire est donnée par

$$|D| := \text{Aire}(D) = |\det(\mathbf{v}, \mathbf{w})|$$

Le changement des aires par des applications linéaires. Soit $A \in M_{2,2}(\mathbb{R})$ inversible et $\lambda_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ l'application linéaire définie par $\lambda_A(\mathbf{x}) = A\mathbf{x}$. Soit $C = [0,1]\times[0,1]$. Alors, l'ensemble $\lambda_A(C)$ est un parallélogramme dans $\mathbb{R}^2$. Son aire est donnée par

$$|\lambda_A(C)| := \text{Aire}(\lambda_A(C)) = |\det(A)|.$$

.

Formule du changement des variables pour des applications linéaires. Soit $A \in M_{2,2}(\mathbb{R})$ inversible

$$\iint_C f(\lambda_A(s,t))|\det(A)|\, dsdt = \iint_{\lambda_A(C)} f(x,y)\, dxdy.$$

Exemple. Pour la matrice diagonale $A = \text{diag}(\lambda, \mu)$ on a

$$\int_0^1\int_0^1 f(\lambda t, \mu s)|\lambda\mu|\, dsdt = \int_0^{\lambda}\int_0^{\mu} f(x,y)\, dxdy.$$

Exemple. Soient $\mathbf{v} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \mathbf{w} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Calculer l'intégral suivant sur le parallélogramme $D = \{s\mathbf{v} + t\mathbf{w}, 0 \leq s \leq 1, 0 \leq t \leq 1\}$:

$$\iint_D x + y \, dxdy = \int_0^1 \int_0^1 ((s + 3t) + (2s + 4t)) \cdot 2 \, dsdt = 10.$$

Le changement des variables. Soit $\psi : U \longrightarrow V$ une application bijective de classe $C^1(U)$. Posons $\psi(s, t) = (x, y)$. Alors, pour toute fonction $f : V \longrightarrow \mathbb{R}$ intégrable

$$\iint_U f(\psi(s,t))|\det J_\psi(s,t)| \, dsdt = \iint_V f(x,y) \, dxdy.$$

Exemple - coordonnés polaires. Soit $\psi(r, \phi) = (r\cos\phi, r\sin\phi) = (x, y)$. Donc $\det J_\psi(r, \phi) = r$. Soit $D = \{(x, y) : x^2 + y^2 \leq R^2\}$ le disque de rayon R. Alors,

$$\iint_{[0,R]\times[0,2\pi]} f(r\cos\phi, r\sin\phi) \, r \, drd\phi = \iint_D f(x,y) \, dxdy.$$

En particulier,

$$\iint_{[0,\infty]\times[0,2\pi]} f(r\cos\phi, r\sin\phi) \, r \, drd\phi = \iint_{\mathbb{R}^2} f(x,y) \, dxdy.$$

Ce changement de variable nous permet de calculer facilement l'intégral de Gauss

$$I := \int_{\mathbb{R}} e^{-\frac{x^2}{2}} \, dx$$

car

$$\begin{aligned} I^2 &= \left(\int_{\mathbb{R}} e^{-\frac{x^2}{2}} \, dx \right)^2 \\ &= \int_{\mathbb{R}} e^{-\frac{x^2}{2}} \, dx \int_{\mathbb{R}} e^{-\frac{y^2}{2}} \, dy \\ &= \iint_{\mathbb{R}^2} e^{-\frac{x^2+y^2}{2}} \, dxdy \\ &= \iint_{[0,\infty]\times[0,2\pi]} e^{-\frac{r^2}{2}} \, r \, drd\phi \\ &= 2\pi \int_0^\infty e^{-\frac{r^2}{2}} \, r \, dr \\ &= -2\pi e^{-\frac{r^2}{2}}\Big|_0^\infty \\ &= 2\pi \end{aligned}$$

Si la fonction $f(x, y)$ est une fonction à symétrie sphérique, i.e. $f(x, y) = g(r)$, alors

$$\iiint_{B_R} f(x,y) \, dxdy = 2\pi \int_{[0,R]} g(r) r \, dr.$$

Le facteur 2π correspond à la longueur d'arc du cercle d'unité $\mathbf{S}_1 \subset \mathbf{R}^2$.

6.2.4 Calcul des intégrales multiples

Intégrales multiples. $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ continue. En généralisant la construction de l'intégrale double décrite ci-dessus on peut définir

$$\iiint_D f(x,y,z)\,dxdydz$$

où $D = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ ou D est un borné ouvert dans $\mathbb{R}^3$ ou encore $D = \mathbb{R}^3$. En particulier,

$$\mathrm{Vol}(D) = |D| = \iiint_D dxdydz.$$

Plus généralement encore pour $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ continue et on peut définir

$$\int ... \int_D f(x_1, ..., x_n)\,dx_1...dx_n$$

pour des domains D convenables.

Le changement des variables. Soit $\psi : U \longrightarrow V$ une application bijective de classe $C^1(U)$. Posons $\psi(t_1, ...t_n) = (x_1, ..., x_n)$. Alors, pour toute fonction $f : V \longrightarrow \mathbb{R}$ intégrable

$$\int ... \int_U f(\psi(t_1, ...t_n))|\det J_\psi(t_1, ...t_n)|dt_1, ...dt_n = \int ... \int_V f(x_1, ..., x_n)dx_1...dx_n.$$

Exemple - coordonnés sphériques. Soit

$$\psi(r, \theta, \phi) = (r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta) = (x, y, z)$$

où $r > 0, \theta \in [0, \pi], \phi \in [0, 2\pi]$. Donc $\det J_\psi(r, \theta, \phi) = r^2\sin\theta > 0$. Soit $D = B_R = \{(x, y, z) : x^2 + y^2 + z^2 \leq R^2\}$ la boule de rayon R. Alors,

$$\iiint_{[0,R]\times[0,\pi]\times[0,2\pi]} f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\,r^2\sin\theta\,drd\theta d\phi =$$
$$\iiint_D f(x,y,z)\,dxdydz.$$

En particulier,

$$\iiint_{[0,\infty[\times[0,\pi]\times[0,2\pi]} f(r\sin\theta\cos\phi, r\sin\theta\sin\phi, r\cos\theta)\,r^2\sin\theta\,drd\theta d\phi =$$
$$\iiint_{\mathbb{R}^3} f(x,y,z)\,dxdydz.$$

L'intégrale en coordonnés sphériques consiste alors d'une intégrale sur la sphère d'unité (les angles θ, ϕ) et les rayons r. Si la fonction $f(x, y, z)$ est une fonction à symétrie sphérique, i.e. $f(x, y, z) = g(r)$, alors

$$\iiint_{B_R} f(x,y,z)\,dxdydz = 4\pi\int_{[0,R]} g(r)r^2\,dr.$$

Le facteur 4π correspond à l'aire de la sphère d'unité $\mathbf{S}_2 \subset \mathbf{R}^3$ et $4\pi r^2$ représente l'aire de la sphère de rayon r. En particulier,

$$\mathrm{Vol}(B_R) = \iiint_{B_R} dxdydz = 4\pi\int_0^R r^2\,dr = \frac{4\pi R^3}{3}.$$

Exemple - distribution des masses. Soient $D \subset \mathbf{R^3}$ borné et $f : D \longrightarrow \mathbb{R}_+$ continue. On peut interpréter f comme densité de masse. Alors,

$$M = \iiint_D f(x, y, z)\, dxdydz$$

représante la masse du corps D et les intégrales

$$R_x = \frac{1}{M} \iiint_D xf(x, y, z)\, dxdydz,$$
$$R_y = \frac{1}{M} \iiint_D yf(x, y, z)\, dxdydz,$$
$$R_z = \frac{1}{M} \iiint_D zf(x, y, z)\, dxdydz$$

les coordonnés du centre de gravité de D. Si $D = B_R$ et $f = e^{-r}$ alors

$$M = \iiint_{B_R} e^{-r}\, dxdydz = 4\pi \int_{[0,R]} e^{-r} r^2\, dr = 8\pi(1 - (1 + R + R^2/2)e^{-R})$$

et $R_x = R_y = R_z = 0$.

Exemple - volume d'un corps de révolution. Soit $f : [a, b] \longrightarrow \mathbf{R}$ une fonction continue. Le sous-ensemble E de $\mathbf{R}^3$ obtenu par la rotation autour de l=axe Ox de la surface délimitée par le graphe de f est donné par

$$E = \{(x, y, z) \in \mathbf{R}^3 : x \in [a, b], y^2 + z^2 < f^2(x)\}$$

Alors, en prenant des coordonnés polaires pour y, z, on obtient

$$\mathrm{Vol}(E) = \iiint_E dxdydz = 2\pi \int_a^b \left(\int_0^{|f(x)|} r\, dr \right) dx = \pi \int_a^b f^2(x)\, dx.$$

Exemple - Fonction radiale en $\mathbf{R}^n$ et volume de la boule d'unité. Soit $f : \mathbf{R}^n \longrightarrow \mathbf{R}$ une fonction radiale (à symétrie sphérique),c'est-à-dire $f(\mathbf{x}) = g(||\mathbf{x}||_2$ pour tous $\mathbf{x} \in \mathbf{R}^n$. Alors, (voir également les exercices) :

$$\int_{\mathbf{R^n}} f(\mathbf{x})\, d\mathbf{x} = |S_{n-1}| \int_0^\infty g(r) r^{n-1}\, dr$$

où $|S_{n-1}| = \mathrm{Vol}_{n-1}(\{\mathbf{x} \in \mathbf{R}^n : ||\mathbf{x}||_2 = 1\})$ désigne l'aire de la sphère unitaire dans $\mathbf{R}^n$. En prenant $g(r) = 1$ si $0 \leq r \leq 1$ et $g(r) = 0$ sinon on trouve pour le volume de la boule d'unité $B_1(\mathbf{0}) = \{\mathbf{x} \in \mathbf{R}^n : ||\mathbf{x}||_2 \leq 1\}$:

$$B_n := \mathrm{Vol}_n(B_1(\mathbf{0})) = |S_{n-1}| \int_0^1 r^{n-1}\, dr = \frac{|S_{n-1}|}{n}.$$

Par conséquent, pour tout $R > 0$

$$\mathrm{Vol}_n(B_R(\mathbf{0})) = B_n R^n.$$

Pour calcluer B_n on utilise les fonctions gaussiennes $f(\mathbf{x}) = e^{-||\mathbf{x}||_2^2} = \prod_{k=1}^{n} e^{-x_k^2}$.
D'une part

$$\int_{\mathbf{R^n}} f(\mathbf{x})\, d\mathbf{x} = |S_{n-1}| \int_0^\infty e^{-r^2} r^{n-1}\, dr = \frac{|S_{n-1}|}{2} \int_0^\infty e^{-s} s^{\frac{n-1}{2}}\, ds = \frac{|S_{n-1}|}{2}\Gamma(\frac{n}{2})$$

par le changement de variable $s = r^2$. D'autre part

$$\int_{\mathbf{R^n}} f(\mathbf{x})\, d\mathbf{x} \prod_{k=1}^{n} \int_{\mathbf{R}} e^{-x_k^2}\, dx_k = \pi^{\frac{n}{2}}$$

puisque

$$\int_{\mathbf{R}} e^{-x_k^2}\, dx_k = 2\int_0^\infty e^{-x_k^2}\, dx_k = \int_0^\infty e^{-s} s^{\frac{-1}{2}}\, ds = \Gamma(\frac{1}{2}) = \sqrt{\pi}.$$

Il en suit

$$B_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2}+1)}. \tag{6.4}$$

6.3 Quelques applications en Physique

6.3.1 Le théorème de Newton

Soit $\mathbf{R^3} : D \longrightarrow \mathbb{R}_+$. Comme avant on peut interpréter f comme densité de masse (ou de charge). Le potentiel gravitationnel en $\mathbf{x} \in \mathbf{R^3}$ y associé est donné par

$$\Phi(\mathbf{x}) = \iiint_{\mathbf{R^3}} \frac{f(\mathbf{y})}{||\mathbf{x}-\mathbf{y}||_2}\, dy_1 dy_2 dy_3$$

Le théorème de Newton affirme que si f est à symétrie sphérique, i.e. $f(\mathbf{x}) = g(||\mathbf{x}||_2)$, alors

$$\begin{aligned}\Phi(\mathbf{x}) =& \frac{1}{||\mathbf{x}||_2} \iiint_{||\mathbf{y}||_2<||\mathbf{x}||_2} g(||\mathbf{y}||_2)\, dy_1 dy_2 dy_3 \\ &+ \iiint_{||\mathbf{y}||_2\geq||\mathbf{x}||_2} \frac{1}{||\mathbf{y}||_2}\, g(||\mathbf{y}||_2)\, dy_1 dy_2 dy_3\end{aligned}$$

En particulier, si $g(||\mathbf{x}||_2) = 0$ pour $||\mathbf{x}||_2 > R$, alors

$$\Phi(\mathbf{x}) = \frac{1}{||\mathbf{x}||_2} \iiint_{||\mathbf{y}||_2<||\mathbf{x}||_2} g(||\mathbf{y}||_2)\, dy_1 dy_2 dy_3 \quad \text{si } ||\mathbf{x}||_2 > R.$$

Démonstration. Il est commode de simplifier le terme $||\mathbf{x}-\mathbf{y}||_2$ en utilisant l'invariance de la norme euclidienne sous transformation orthogonale avant passer aux coordonnées sphériques pour $\mathbf{y}$. Soit $\mathbf{x}$ fixe. Il existe une matrice orthogonale H telle que $H\mathbf{x} = ||\mathbf{x}||_2\mathbf{e}_3$ (voir l'exercice du chapitre 1 sur les matrices de Householder). En utilisant $H^{-1}H = 1$ et H^{-1} orthogonal on a pour tous $\mathbf{y}$:

$$||\mathbf{x}-\mathbf{y}||_2 = ||H^{-1}H(\mathbf{x}-\mathbf{y})||_2 = ||H\mathbf{x} - H\mathbf{y}||_2.$$

Alors,

$$\Phi(\mathbf{x}) = \iiint_{\mathbf{R^3}} \frac{g(||\mathbf{y}||_2)}{||\mathbf{x}-\mathbf{y}||_2}\, dy_1 dy_2 dy_3 = \iiint_{\mathbf{R^3}} \frac{g(||H\mathbf{y}||_2)}{||H\mathbf{x}-H\mathbf{y}||_2}\, dy_1 dy_2 dy_3$$

Par le changement de variable $\mathbf{y} = H\mathbf{y}'$ (noter que $|\det H| = 1$) :

$$\Phi(\mathbf{x}) = \iiint_{\mathbf{R^3}} \frac{g(||\mathbf{y}'||_2)}{||r\mathbf{e}_3 - \mathbf{y}'||_2} |\det H|\, dy_1' dy_2' dy_3', \quad r = ||\mathbf{x}||_2.$$

En passant aux coordonnées sphériques pour $\mathbf{y}'$ (en notant s la variable radiale) on obtient par une première simplification :

$$\begin{aligned}\Phi(\mathbf{x}) &= \iiint_{[0,\infty[\times[0,\pi]\times[0,2\pi]} \frac{g(s)}{\sqrt{r^2+s^2-2rs\cos\theta}}\, s^2 \sin\theta\, ds d\theta d\phi \\ &= 2\pi \iint_{[0,\infty[\times[0,\pi]} \frac{g(s)s^2}{\sqrt{r^2+s^2-2rs\cos\theta}}\, \sin\theta\, ds d\theta \\ &= 2\pi \iint_{[0,\infty[\times[0,\pi]} \frac{g(s)s^2}{rs} \frac{d}{d\theta}\sqrt{r^2+s^2-2rs\cos\theta}\, ds d\theta.\end{aligned}$$

En évaluant $\sqrt{r^2+s^2-2rs\cos\theta}$ en $\theta = \pi$ et $\theta = 0$ on trouve

$$\sqrt{r^2+s^2-2rs\cos\theta}\Big|_{\theta=0}^{\theta=\pi} = \sqrt{(r+s)^2} - \sqrt{(r-s)^2} = \begin{cases} 2s, & \text{si } s \le r\,; \\ 2r, & \text{si } s > r. \end{cases}$$

Par conséquent, en divisant l'intervalle de l'intégration pour la variable radiale s en deux partie : $[0,\infty[= [0,r] \cup [r,,\infty[$:

$$\Phi(\mathbf{x}) = \frac{4\pi}{r} \int_{[0,r]} g(s)\, s^2\, ds + 4\pi \int_{[r,\infty[} \frac{g(s)s^2}{s}\, ds$$

ce qui donne l'affirmation.

6.3.2 Les équations d'Euler-Lagrange

Soit $L : [a,b]\times\mathbb{R}\times\mathbb{R} \longrightarrow \mathbb{R}$ une fonction de classe C^2. En physique la fonction L représente le Lagrangien du système (ici d'un mouvement unidimensionnel) et on note $L = L(t,q,\dot{q})$ en vu de l'interprétation des variables comme le temps, la position et la vitesse. Soit E l'espace vectoriel de chemins (trajectoires) de classe C^2 entre x_a et x_b :

$$E := \{\phi = \phi(t) \in C^2([a,b]) :\ q(a) = x_a,\ q(b) = x_b\}$$

Pour tout élément on $\phi = \phi(t) \in E$ on définit une action $S : E \longrightarrow \mathbb{R}$ par

$$S(\phi) := \int_a^b L(t, \phi(t), \dot{\phi}(t))\, dt$$

Les équations d'Euler-Lagrange donne une conditions nécéssaire pour des extremums (locaux) de l'action S :

Théroème. Soit $q = q(t) \in E$ tel que $S(q) = \min\{S(\phi),\ \phi \in E\}$. Alors

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(t, q(t), \dot{q}(t)) - \frac{\partial L}{\partial q}(t, q(t), \dot{q}(t)) = 0 \tag{6.5}$$

Démonstration. Pour tout $\psi \in C^2([a,b])$ tel que $\psi(a) = \psi(b) = 0$ le segment $k(\epsilon) = q + \epsilon\psi$ est dans E et $S(k(\epsilon))$ atteint son minimum en $\epsilon = 0$ d'où

$$\frac{d}{d\epsilon}\bigg|_{\epsilon=0} S(k(\epsilon)) = 0.$$

Par la règle de composition (en appliquant les résultats sur les intégrales qui dépendent d'une paramètre) on trouve :

$$\frac{d}{d\epsilon}S(k(\epsilon)) = \int_a^b \frac{d}{d\epsilon}\, L(t, q(t) + \epsilon\psi(t), \dot{q}(t) + \epsilon\dot{\psi}(t))\, dt$$
$$= \int_a^b \psi(t)\frac{\partial L}{\partial q}(t, q(t) + \epsilon\psi(t), \dot{q}(t) + \epsilon\dot{\psi}(t)) + \dot{\psi}(t)\frac{\partial L}{\partial \dot{q}}(t, q(t) + \epsilon\psi(t), \dot{q}(t) + \epsilon\dot{\psi}(t))\, dt$$

On intègre le second terme par partie (en utilisant $\psi(a) = \psi(b) = 0$) :

$$\int_a^b \dot{\psi}(t)\frac{\partial L}{\partial \dot{q}}(t, q(t)+\epsilon\psi(t), \dot{q}(t)+\epsilon\dot{\psi}(t))dt = -\int_a^b \psi(t)\frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(t, q(t)+\epsilon\psi(t), \dot{q}(t)+\epsilon\dot{\psi}(t))dt.$$

Il en suit qu'en $\epsilon = 0$:

$$0 = \frac{d}{d\epsilon}\bigg|_{\epsilon=0} S(k(\epsilon)) = \int_a^b \psi(t)\left(\frac{\partial L}{\partial q}(t, q(t), \dot{q}(t)) - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}}(t, q(t), \dot{q}(t))\right) dt$$

pour tous $\psi \in C^2([a,b])$ tel que $\psi(a) = \psi(b) = 0$ et donc l'affirmation.

Exemple. Soit $L = L(q, \dot{q}) = \frac{m\dot{q}^2}{2} - V(q)$, $m > 0$, $V \in C^2$. Alors,

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

Donc l'équation d'Euler-Lagrange est équivalente à l'équation différentielle ordinaire (voir chapitre 7)

$$m\frac{d}{dt}\dot{q}(t) = -V'(q(t)) \quad \Leftrightarrow m\ddot{q}(t) = -V'(q(t)).$$

C'est l'équation de mouvement (en une dimension) d'une particule dans le potentiel V.

Les équations d'Euler-Lagrange en $\mathbb{R}^n$. Soit $L : [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}$ une fonction de classe C^2. Soit E l'espace vectoriel de chemins (trajectoires) de classe C^2 entre $\mathbf{x}_a$ et $\mathbf{x}_b$:

$$E := \{\phi = \phi(t) \in C^2([a,b]) :\ \mathbf{q}(a) = \mathbf{x}_a,\ \mathbf{q}(b) = \mathbf{x}_b\}$$

Pour tout élément on $\phi = \phi(t) \in E$ on définit une action $S : E \longrightarrow \mathbb{R}$ par

$$S(\phi) := \int_a^b L(t, \phi(t), \dot{\phi}(t))\, dt$$

Les équations d'Euler-Lagrange donne une conditions nécéssaire pour des extremums (locaux) de l'action S :

Théroème. Soit $\mathbf{q} = \mathbf{q}(t) \in E$ tel que $S(\mathbf{q}) = \min\{S(\phi),\, \phi \in E\}$. Alors, pour tous $k = 1, \ldots, n$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_k}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) - \frac{\partial L}{\partial q_k}(t, \mathbf{q}(t), \dot{\mathbf{q}}(t)) = 0 \tag{6.6}$$

Exemple. Soit $L = L(\dot{\mathbf{q}}) = \frac{m\langle \dot{\mathbf{q}}, \dot{\mathbf{q}}\rangle}{2} - V(q)$, $m > 0$,. Alors,

$$\frac{\partial L}{\partial \dot{q}_k} = m\dot{q}_k, \quad \frac{\partial L}{\partial q_k} = 0.$$

Donc l'équation d'Euler-Lagrange est équivalente aux système dl'équations différentielles ordinaires (voir chapitre 7)

$$m\frac{d}{dt}\dot{\mathbf{q}}(t) = \mathbf{0} \quad \Leftrightarrow m\ddot{\mathbf{q}}(t) = \mathbf{0}.$$

C'est l'équation de mouvement (en n dimensions) d'une particule libre ou (des particules libres). La solution telle que $\mathbf{q}(a) = \mathbf{x}_a$, $\mathbf{q}(b) = \mathbf{x}_b$ est donnée par

$$\mathbf{q}(t) = \frac{t-a}{b-a}\,\mathbf{x}_b + \frac{b-t}{b-a}\,\mathbf{x}_a.$$

6.3.3 Ensemble microcanonique et l'équation d'un gaz parfait

On considère N particules identiques libres de masse $m > 0$ enfermées dans une boite $D \subset \mathbb{R}^3$ de volume $V < \infty$. L'énergie (cinétique) de la i-ième particule est donnée par $H_i = \dfrac{\langle \mathbf{p}_i, \mathbf{p}_i\rangle}{2m}$, , $\mathbf{p}_i \in \mathbb{R}^3$. L'énergie totale est $H = \displaystyle\sum_{i=1}^{N} H_i$. On définit l'entropie du système $S = S(E, V)$ par

$$S(E, V) = k \ln \Phi(E, V) + \text{const}$$

où k désigne la constante de Boltzmann (rappel : $R = N_A \cdot k$, $N_A \simeq 6 \cdot 10^2 3$ est la constante d'Avrogardo) et $\Phi(E, V)$ est donné par

$$\Phi(E, V) = \frac{1}{h^{3N}} \int\cdots\int_{\{\mathbf{x}_i \in D, H < E\} \subset \mathbb{R}^{6N}} d\mathbf{p}_1 \cdots d\mathbf{p}_N d\mathbf{x}_1 \cdots d\mathbf{x}_N$$

où h dénote la constante de Planck. On peut calculer $\Phi(E, V)$ explicitement : Les intégrales par rapport aux $\mathbf{x}_i$ donnent V^N. Ensuite en notant que $H = \dfrac{\langle \mathbf{p}, \mathbf{p}\rangle}{2m}$ avec $\mathbf{p} \in \mathbb{R}^{3N}$ on obtient

$$\begin{aligned}\Phi(E, V) &= \frac{V^N}{h^{3N}} \int\cdots\int_{\{\langle \mathbf{p}, \mathbf{p}\rangle < 2mE\} \subset \mathbb{R}^{3N}} d\mathbf{p}_1 \cdots d\mathbf{p}_N \\ &= \frac{V^N}{h^{3N}} (\sqrt{2mE})^{3N} B_{3N} = \frac{V^N (2\pi m E)^{\frac{3N}{2}}}{h^{3N}\Gamma(\frac{3N}{2} + 1)}\end{aligned}$$

en utilisant (6.4). Par conséquent, en mettant la constante additive égale à zéro :

$$S(E,V) = k \ln \frac{V^N (2\pi m E)^{\frac{3N}{2}}}{h^{3N} \Gamma(\frac{3N}{2}+1)}$$

On résoud cette identité pour E :

$$E = E(S,V) = \frac{h^2 \Gamma(\frac{3N}{2}+1)^{\frac{2}{3N}}}{2\pi m} \frac{e^{\frac{2S}{3Nk}}}{V^{\frac{2}{3}}}.$$

On définit la temperature T et la pression P par les dérivées partielles de $E(V,S)$:

$$T := \frac{\partial E(S,V)}{\partial S}, \quad P := -\frac{\partial E(S,V)}{\partial V}$$

c'est-à-dire

$$T = \frac{2E}{3kN}, \quad P = \frac{2E}{V}.$$

On élimine E de ces équations pour trouver

$$PV = NkT.$$

En particulier, si $N = N_A$ (un mol) :

$$PV = RT.$$

Chapitre 7

Equations différentielles

7.1 Classification des équations différentielles

Soit $I \subset \mathbb{R}$ un intervalle ouvert et $f : I \times \mathbb{R} \longrightarrow \mathbb{R}$ une fonction continue. On appelle équation différentielle du premier ordre une équation de la forme

$$y' = f(t, y) \quad \text{ou} \quad y' = f(x, y) \quad \text{ou encore} \quad \dot{y} = f(t, y). \tag{7.1}$$

$y = y(t)$ est dite une solution de $y' = f(t, y)$ sur l'intervalle I si $y(t) \in C^1(I)$ telle que

$$\frac{dy(t)}{dt} = f(t, y(t))$$

Dans l'équation différentielle (1) on appelle y la variable dépendante et t la variable indépendante.

Interprétation géométrique. On définit un champs vectoriel continu $\mathbf{v} : I \times \mathbb{R} \longrightarrow \mathbb{R}^2$ par

$$\mathbf{v}(t, y) = \begin{pmatrix} 1 \\ f(t, y) \end{pmatrix}$$

Le graphe d'une solution $y(t)$ définit une courbe $\mathbf{k} : I \longrightarrow \mathbb{R}^2$ de classe $C^1(I)$ par

$$\mathbf{k}(t) = \begin{pmatrix} t \\ y(t) \end{pmatrix}$$

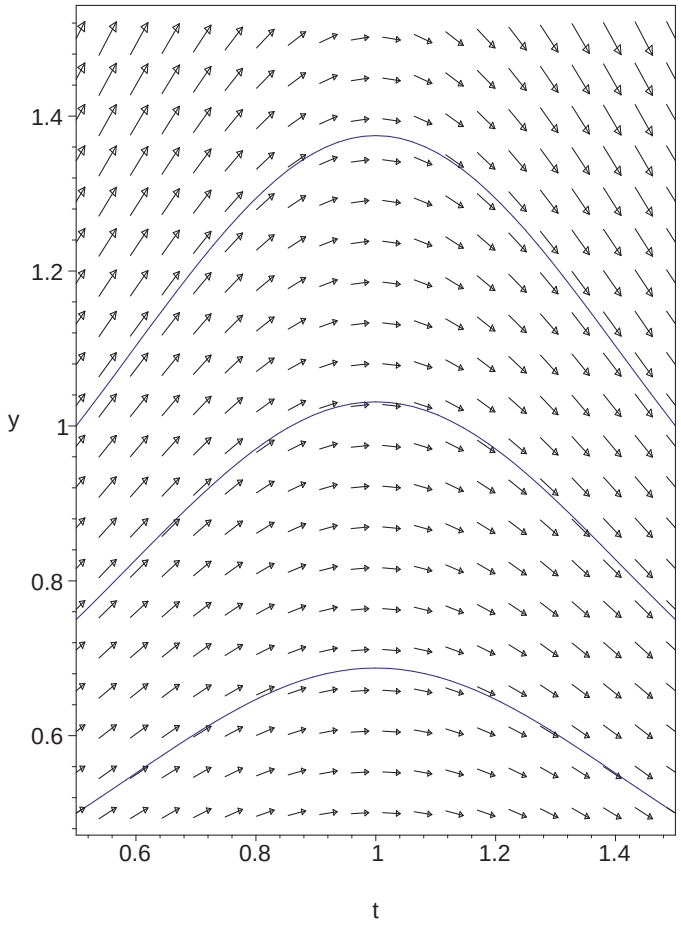
Cette courbe est tangentielle au champs vectoriel $\mathbf{v}$ car

$$\frac{d\,\mathbf{k}(t)}{dt} = \mathbf{v}(t, y(t)).$$

On appelle une telle courbe aussi une courbe intégrale du champs $\mathbf{v}$.

Problème de Cauchy. En cherchant une courbe intégrale passante par le point (t_0, y_0) on cherche une solution du problème de Cauchy

$$y' = f(t, y) \quad \text{et} \quad y(t_0) = y_0. \tag{7.2}$$



$\mathbf{v}(t, y)$ et ses courbes intégrales.

L'équation

$$y'' = f(t, y, y') \tag{7.3}$$

est appelée l'équation différentielle du second ordre où $f : I \times \mathbb{R}^2 \longrightarrow \mathbb{R}$. En mécanique l'équation de Newton $F = ma$ décrivante le mouvement unidimensionnelle d'une particule de masse m sous l'influence de la force F est une équation différentielle du second ordre : t represente le temps, $y(t)$ la position de la masse au temps t, $v(t) = y'(t)$ sa vitesse et $a(t) = y''(t)$ l'acceleration. La force $F = f(t, y, y')$ peut dépendre de trois variables t, y, y'. En mécanique on dénote souvent les derivées par rapport au temps t par $\dot{y}, \ddot{y}$ etc. Il existe des nombreux exemples dans des autres domaines de physique, par exemple en éltricité l'équation

$$L\ddot{y} + R\dot{y} + C^{-1}y = 0$$

décrit un circuit de resistance R, de capacité C et de l'inductivité L ou $y(t)$ repésente la charge. C'est l'équation d'un oscillateur harmonique (avec frottement et sans force extérieure).

$y = y(t)$ est dite une solution de $y'' = f(t, y, y')$ sur l'intervalle I si $y(t) \in C^2(I)$ telle que

$$\frac{d^2y(t)}{dt^2} = f\big(t, y(t), \frac{dy(t)}{dt}\big).$$

Le problème de Cauchy pour l'équation $y'' = f(t, y, y')$ est donné par

$$y'' = f(t, y, y') \quad \text{et} \quad y(t_0) = y_0,\ y'(t_0) = v_0. \tag{7.4}$$

En mécanique y_0 et v_0 représentent la position respectivement la vitesse au temps t_0. Pour l'équation $y'' = f(t, y, y')$ on peut également fixer des condistions "aux bords" (d'où la notion du *problème aux bords* dans la littérature), par exemple :

$$y'' = f(t, y, y') \quad \text{et} \quad y(t_0) = y_0,\ y(t_1) = y_1. \tag{7.5}$$

Si $f = f(t, y)$ cette équation représente l'équation d'Euler- Lagrange (6.5) pour le Lagrangien

$$L = \frac{1}{2}y'^2 - F(t, y), \quad \frac{\partial F(t, y)}{\partial y} = f(t, y).$$

Plus généralement, on appelle l'équation

$$y^{(n)} = f(t, y, y', ..., y^{(n-1)})$$

léquation différentielle d'ordre n. Si $\mathbf{y} = \mathbf{y}(t) \in \mathbb{R}^n$ et $\mathbf{f} : I \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ l'équation

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$$

est appelé système des n équations différentielles (du premier ordre). Toute équation d'ordre n peut s'écrire comme système des n équations différentielles. Par exemple, si $n = 2$ on définit

$$\mathbf{y} = \begin{pmatrix} y \\ y' \end{pmatrix}$$

et

$$\mathbf{f} = \mathbf{f}(t, \mathbf{y}) = \begin{pmatrix} y' \\ f(t, y, y') \end{pmatrix}.$$

Si y est une solution de $y'' = f(t, y, y')$ alors $\mathbf{y}$ est une solution de

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$$

et vice versa. Par exemple, l'oscillateur harmonique s'écrit comme suit :

$$\frac{d}{dt}\begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{R}{L} & -\frac{1}{LC} \end{pmatrix}\begin{pmatrix} y \\ y' \end{pmatrix}$$

Dans la pratique un objectif est de trouver une solution explicite d'une équation différentielle et du problème de Cauchy associé. Souvent c'est impossible et on doit faire une étude dite qualitative du problème et résoudre le problème numériquement. Pour ces études qualitatives il faut d'abord savoir que des solutions existent. Un résultat mathématique fondamental est l'existence et l'unicité de solution du problème de Cauchy

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \quad \text{et} \quad \mathbf{y}(t_0) = \mathbf{y}_0$$

sous des hypothèses assez générales.

7.2 Existence et unicité locale des solutions

D'abord nous transformons le problème de Cauchy pour le système d'équations différentielles dans un système d'équations intégrales.

Théorème - Équation intégrale. Soient I un intervalle, $t_0 \in \overset{\circ}{I}$, $D \subset \mathbb{R}^n$ et $\mathbf{f} : I \times D \longrightarrow \mathbb{R}^n$ continu. Alors, $\mathbf{y} : I \longrightarrow \mathbb{R}^n$ est une solution de classe C^1 de

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \quad \text{et} \quad \mathbf{y}(t_0) = \mathbf{y}_0 \tag{7.6}$$

si et seulement si $\mathbf{y} : I \longrightarrow \mathbb{R}^n$ est une solution continue de l'équation intégrale

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_{t_0}^{t} \mathbf{f}(s, \mathbf{y}(s))\, ds. \tag{7.7}$$

Démonstration. Si $\mathbf{y} : I \longrightarrow \mathbb{R}^n$ est une solution de classe C^1 de (7.2), alors en intégrant le système d'équations différentielles entre t_0 et $t \in I$:

$$\int_{t_0}^{t} \mathbf{y}'(s)\, ds = \int_{t_0}^{t} \mathbf{f}(s, \mathbf{y}(s))\, ds$$

c'est-à-dire

$$\mathbf{y}(t) - \mathbf{y}_0 = \int_{t_0}^{t} \mathbf{f}(s, \mathbf{y}(s))\, ds.$$

Si $\mathbf{y} : I \longrightarrow \mathbb{R}^n$ est une solution continue de l'équation intégrale (7.7), alors $\mathbf{y}$ est de classe C^1 puisque le membre de droite de (7.7) est l'intégrale indéfinie d'une fonction continue (de la variable s) donc C^1. En prenant la dérivée de (7.7) par rapport à t on obtient le système d'équations différentielles de (7.2). De plus, par l'équation intégrale (7.7) on a $\mathbf{y}(t_0) = \mathbf{y}_0$.

Définitions. Soit $t_0 \in \mathbb{R}$, $\mathbf{y}_0 \in \mathbb{R}^n$. On définit le "cylindre" fermé $R_{a,b} \subset \mathbb{R}^{n+1}$ par

$$R_{a,b} := \{ \begin{pmatrix} t \\ \mathbf{y} \end{pmatrix} \in \mathbb{R}^{n+1} : |t - t_0| \leq a,\ ||\mathbf{y} - \mathbf{y}_0||_2 \leq b \}$$

Une fonction continue $\mathbf{f} : R_{a,b} \longrightarrow \mathbb{R}^n$ est appelée lipschitzienne sur $R_{a,b}$ par rapport à $\mathbf{y}$ s'il existe une constante $L > 0$ telle que pour tous $\begin{pmatrix} t \\ \mathbf{y}_1 \end{pmatrix}, \begin{pmatrix} t \\ \mathbf{y}_2 \end{pmatrix} \in R_{a,b}$:

$$||\mathbf{f}(t, \mathbf{y}_2) - \mathbf{f}(t, \mathbf{y}_1)||_2 \leq L||\mathbf{y}_2 - \mathbf{y}_1||_2. \tag{7.8}$$

Théorème - Existence et unicité locale. Soit $\mathbf{f} : R_{a,b} \longrightarrow \mathbb{R}^n$ lipschitzienne sur $R_{a,b}$ par rapport à $\mathbf{y}$. Soit $M = M_{a,b} = \max\{||\mathbf{f}(t, \mathbf{y})||_2 : (t, \mathbf{y}) \in R_{a,b}\}$. Alors, le problème de Cauchy (7.2)

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \quad \text{et} \quad \mathbf{y}(t_0) = \mathbf{y}_0 \tag{7.9}$$

admet une solution unique $\mathbf{y} : [t_0 - \alpha, t_0 + \alpha] \longrightarrow \mathbb{R}^n$ avec $\alpha = \min(a, \dfrac{b}{M})$.

Remarque - comprendre α. La constante α donne une condition suffisante pour que la solution $\mathbf{y}(t)$ ne quitte pas le cylindre $R_{a,b}$. En fait, par l'inégalité triangulaire (voir chap. 2, une inégalité pour la longueur) et l'équation différentielle

$$||\mathbf{y}(t) - \mathbf{y}_0||_2 \leq \int_{[t_0,t]} ||\mathbf{y}'(s)||_2 \, ds = \int_{[t_0,t]} ||\mathbf{f}(s,\mathbf{y}(s))||_2 \, ds \leq M|t-t_0|.$$

Si $M|t-t_0| \leq b$ et $|t-t_0| \leq a$, alors $\mathbf{y}(t) \in R_{a,b}$. Si $Ma \leq b$ on peut prendre $\alpha = a$ sinon il faut choisir $\alpha = \frac{b}{M}$.

Démonstration. Sur $I := [t_0 - \alpha, t_0 + \alpha]$ on considère l'espace $E = C^0(I)$ des fonctions (courbes) continues $\mathbf{y}(t) : I \longrightarrow \mathbb{R}^n$ muni de la norme

$$||\mathbf{y}||_E := \sup\{||\mathbf{y}(t)||_2 e^{-2L|t-t_0|} : \, t \in I\} = \max\{||\mathbf{y}(t)||_2 e^{-2L|t-t_0|} : \, t \in I\}.$$

Alors, $(E, ||\cdot||_E)$ est une espace de Banach (la norme $||\cdot||_E$ est équivalente à la norme habituelle de la convergence uniforme, c'est-à-dire sans le facteur exponentiel). Soit F l'ensemble des courbes $\mathbf{y} \in E$ tel que leur graphe est dans $R_{a,b}$. Alors, $F \subset E$ est fermé car $R_{a,b}$ est fermé dans $\mathbb{R}^{n+1}$. On considère l'application $T : F \longrightarrow E$ définie par

$$(T\mathbf{y})(t) := \mathbf{y}_0 + \int_{t_0}^{t} \mathbf{f}(s,\mathbf{y}(s)) \, ds.$$

Grâce à l'estimation $||(T\mathbf{y})(t) - \mathbf{y}_0||_2 \leq M\alpha \leq b$ on a $T : F \longrightarrow F$. De plus, en utilisant

$$\begin{aligned}
e^{-2L|t-t_0|}||(T\mathbf{y})(t) - (T\mathbf{x})(t)||_2 &\leq e^{-2L|t-t_0|} \int_{[t_0,t]} ||\mathbf{f}(s,\mathbf{y}(s)) - \mathbf{f}(s,\mathbf{x}(s))||_2 \, ds \\
&\leq e^{-2L|t-t_0|} \int_{[t_0,t]} L||\mathbf{y}(s) - \mathbf{x}(s)||_2 \, ds \\
&= Le^{-2L|t-t_0|} \int_{[t_0,t]} e^{2L|s-t_0|} e^{-2L|s-t_0|} ||\mathbf{y}(s) - \mathbf{x}(s)||_2 \, ds \\
&\leq Le^{-2L|t-t_0|} ||\mathbf{y} - \mathbf{x}||_E \int_{[t_0,t]} e^{2L|s-t_0|} \, ds \\
&\leq Le^{-2L|t-t_0|} ||\mathbf{y} - \mathbf{x}||_E \frac{e^{2L|t-t_0|} - 1}{2L} \\
&\leq \frac{1}{2} ||\mathbf{y} - \mathbf{x}||_E
\end{aligned}$$

on voit en prenant le maximum à gauche que $T : F \longrightarrow F$ est une contraction :

$$||T(\mathbf{y}) - T(\mathbf{x})||_E \leq \frac{1}{2} ||\mathbf{y} - \mathbf{x}||_E.$$

Alors, T admet un unique point fixe $\mathbf{y} = \mathbf{y}(t)$ dans F :

$$\mathbf{y}(t) = (T\mathbf{y})(t) := \mathbf{y}_0 + \int_{t_0}^{t} \mathbf{f}(s,\mathbf{y}(s)) \, ds.$$

Fonction localement lipschitzienne. Soient I un intervalle, $D \subset \mathbb{R}^n$. Une fonction $\mathbf{f} : I \times D \longrightarrow \mathbb{R}^n$ continue est appelée localement lipschitzienne par rapport à $\mathbf{y}$ sur $I \times D$ si pour tous $t_0 \in \overset{\circ}{I}$, $\mathbf{y}_0 \in \overset{\circ}{D}$, il existe un cylindre $R_{a,b} \subset I \times D$ et une constante $L := L(t_0, \mathbf{y}_0) > 0$ telle que

$$||\mathbf{f}(t, \mathbf{y}_2) - \mathbf{f}(t, \mathbf{y}_1)||_2 \leq L||\mathbf{y}_2 - \mathbf{y}_1||_2$$

dans $R_{a,b}$. Une fonction $\mathbf{f} : I \times D \longrightarrow \mathbb{R}^n$ continue telle que les dérivées partielles $\dfrac{\partial f_i(t, \mathbf{y})}{\partial y_j}$ sont continues pour tous ($1 \leq i, j \leq n$ est toujours localement lipschitzienne par rapport à $\mathbf{y}$. En fait, par la règle de composition

$$\begin{aligned}\mathbf{f}(t, \mathbf{y}_2) - \mathbf{f}(t, \mathbf{y}_1) &= \int_0^1 \frac{d}{d\sigma}\mathbf{f}(t, \sigma\mathbf{y}_2 + (1-\sigma)\mathbf{y}_1)\, d\sigma \\ &= \int_0^1 J_{\mathbf{f}}^{\mathbf{y}}(t, \sigma\mathbf{y}_2 + (1-\sigma)\mathbf{y}_1)(\mathbf{y}_2 - \mathbf{y}_1)\, d\sigma\end{aligned}$$

où $J_{\mathbf{f}}^{\mathbf{y}}$ signifie la matrice $n \times n$ contenante les dérivées partielles $\dfrac{\partial f_i(t, \mathbf{y})}{\partial y_j}$. Par la continuité de ces dérivées partielles

$$||\mathbf{f}(t, \mathbf{y}_2) - \mathbf{f}(t, \mathbf{y}_1)||_2 \leq \int_0^1 ||J_{\mathbf{f}}^{\mathbf{y}}(t, \sigma\mathbf{y}_2 + (1-\sigma)\mathbf{y}_1)||_2 d\sigma ||\mathbf{y}_2 - \mathbf{y}_1||_2 \leq L||\mathbf{y}_2 - \mathbf{y}_1||_2$$

avec $L = \max\{||J_{\mathbf{f}}^{\mathbf{y}}(t, \mathbf{y})||_2 : (t, \mathbf{y}) \in R_{a,b}$. Par exemple, la fonction $f := [0, 1] \times]0, \infty[\longrightarrow \mathbb{R}$ donnée par $f(t, y) = \dfrac{t}{y}$ est localement lipschitzienne par rapport à y. Elle n'est pas lipschitzienne par rapport à y pour $y \in]0, \infty[$.

Corollaire. Soit $\mathbf{f} : I \times D \longrightarrow \mathbb{R}^n$ localement lipschitzienne par rapport à $\mathbf{y}$ sur $I \times D$. Pour tous $t_0 \in \overset{\circ}{I}$, $\mathbf{y}_0 \in \overset{\circ}{D}$, il existe un intervalle J tel que le problème de Cauchy admet une solution unique dans J.

Autres propriéts. La solution $\mathbf{y}(t)$ du problème de Cauchy dépend de ces conditions initiales $t_0, \mathbf{y}_0$. On peut donc écrire $\mathbf{y} = \mathbf{y}(t, t_0, \mathbf{y}_0)$. Par un argument du type "point fixe" on peut démontrer que y est continue en t_0 et $\mathbf{y}_0$. De même si la fonction $\mathbf{f}$ est une fonction continue d'un paramètre $\lambda : \mathbf{f} = \mathbf{f}(t, \mathbf{y}, \lambda$. Alors, la solution $\mathbf{y} = \mathbf{y}(t, \lambda)$ est continue en λ.

7.3 Résolution des certains équations différentielles du premier ordre

7.3.1 $\mathbf{y' = f(t)}$

La solution générale de

$$y' = f(t)$$

pour une fonction $f : I \longrightarrow \mathbb{R}$ continue est donnée par la famille des primitives de $f(t)$:

$$y(t) = \int^t f(s)\, ds + C, \quad C \in \mathbb{R}.$$

Le problème de Cauchy

$$y' = f(t) \quad \text{et} \quad y(t_0) = y_0$$

admet une solution unique donnée par

$$\int_{t_0}^t y'(s)\, ds = \int_{t_0}^t f(s)\, ds$$

i.e.

$$y(t) = \int_{t_0}^t f(s)\, ds + y(t_0).$$

Exemples.

1. La solution générale de
$$y' = 2 \sin t \cos t$$
est donnée par
$$y(t) = \sin^2 t + C, \quad C \in \mathbb{R}.$$

2. On considère
$$y' = \frac{2}{1 - t^2}$$
sur $I =]-1, 1[$. La solution du problème de Cauchy avec $y(-0.5) = 3$ est donnée par
$$\int_{-0.5}^t y'(s)\, ds = \int_{-0.5}^t \frac{2}{1 - s^2}\, ds = \int_{-0.5}^t \frac{1}{1 + s} + \frac{1}{1 - s}\, ds$$
i.e.
$$y(t) = \ln(1 + t) - \ln(1 - t) - (\ln 1/2 - \ln 3/2) + 3 = \ln \frac{1 + t}{1 - t} + \ln 3 + 3.$$
Noter aussi que
$$\ln \frac{1 + t}{1 - t} = 2 \operatorname{arctanh} t.$$

7.3.2 $\mathbf{y'(x) + a(x)y(x) = b(x)}$

Trouver les solutions de l'équation linéaire $\mathbf{y'(x) + a(x)y(x) = b(x)}$. Soient $a, b : I \longrightarrow \mathbb{R}$ continues. Pour résoudre l'équation linéaire homogène (i.e. $b(x) \equiv 0$)

$$y'(x) + a(x)y(x) = 0$$

on définit

$$u(x) = e^{A(x)}y(x)$$

où $A(x)$ est un primitive de $a(x)$, i.e. $A'(x) = a(x)$. La fonction $u(x)$ satisfait

$$u'(x) = e^{A(x)}(y'(x) + a(x)y(x)) = 0$$

Donc $u(x) = c$ pour une constante $c \in \mathbb{R}$ et $y(x) = c\,e^{-A(x)}$ est la solution générale de l'équation linéaire homogène. Pour trouver une solution particulière de l'équation linéaire inhomogène notons que $u(x)$ satisfait

$$u'(x) = b(x)e^{A(x)}.$$

Donc

$$u(x) = \int^x b(s)e^{A(s)}\,ds = \int_{x_0}^x b(s)e^{A(s)}\,ds + c.$$

pour un $x_0 \in \mathbb{R}$, i.e.

$$y(x) = c\,e^{-A(x)} + e^{-A(x)}\int_{x_0}^x b(s)e^{A(s)}\,ds.$$

Formules - l'équation linéaire. Solution générale :

$$y(x) = c\,e^{-A(x)} + e^{-A(x)}\int^x b(s)e^{A(s)}\,ds, \quad A'(x) = a(x).$$

Solution du problème de Cauchy $y(x_0) = y_0$:

$$y(x) = y_0e^{A(x_0)-A(x)} + e^{-A(x)}\int_{x_0}^x b(s)e^{A(s)}\,ds, \quad A'(x) = a(x).$$

Notons qu'il existe toujours une solution unique du problème de Cauchy.

Propriétés générales.

1. Soient $y_1(x), y_2(x)$ deux solutions de l'équation homogène $y'(x)+a(x)y(x) = 0$. Alors toute combinaison linéaire $\alpha y_1(x)+\beta y_2(x)$ de $y_1(x), y_2(x)$ est aussi une solution de l'équation homogène.
2. Soient $z_1(x), z_2(x)$ deux solutions de l'équation inhomogène $y'(x)+a(x)y(x) = b(x)$. Alors $z_1(x) - z_2(x)$ est une solution de l'équation homogène. Ou, si $y(x)$ est une solution de l'équation homogène et $z(x)$ est une solution de l'équation inhomogène, alors $y(x) + z(x)$ est une solution de l'équation inhomogène. C'est pourquoi la solution générale de l'équation inhomogène s'écrit comme la somme de la solution générale de l'équation homogène et d'une solution particulièère de l'équation inhomogène.

Méthode de variation des constantes . Une méthode pratique pour trouver une solution particulière de l'équation inhomogène est de remplacer la constante c dans la solution génrale $y(x) = c\,e^{-A(x)}$ par une fonction $c(x)$ est insérer cette ansatz dans l'équation inhomogène. On trouve alors

$$\big(c(x)\,e^{-A(x)}\big)' + a(x)c(x)\,e^{-A(x)} = b(x)$$

et finalement

$$c'(x) = b(x)e^{A(x)}$$

Exemple. Donner la solution générale de

$$y' + \frac{t}{1+t^2}y = \frac{t}{(1+t^2)^2}$$

Noter que $A(t) = \frac{1}{2}\ln(1+t^2)$. Par conséquent, la solution générale de l'équation homogène associé est de la forme

$$c\,e^{-A(t)} = \frac{c}{\sqrt{1+t^2}}, \quad c \in \mathbb{R}.$$

Une solution particulière de l'équation inhomogène est donnée par

$$e^{-A(t)}\int_0^t b(s)e^{A(s)}\,ds = \frac{1}{\sqrt{1+t^2}}\int_0^t \frac{s}{(1+t^2)^{3/2}} = \frac{-1}{1+t^2}.$$

Donc la solution générale est de la forme

$$y(t) = \frac{c}{\sqrt{1+t^2}} - \frac{1}{1+t^2}, \quad c \in \mathbb{R}.$$

7.3.3 $\mathbf{y' = f(y)}$

Cette équation est appelée équation autonôme car la fonction f ne dépend pas de la variable indépendante t. L'équation différentielle autonôme a la propriéte d'être "invariante sous translations en t", i.e. si $y(t)$ est une solution de $y' = f(y)$ alors $y(t+s)$ est une solution de $y' = f(y)$ pour tout s réel fixe. Si y_0 est tel que $f(y_0) = 0$, alors la fonction constante $y(t) = y_0$ est une solution de l'équation différentielle $y' = f(y)$. Une telle solution est appelée une solution stationnaire ou un point stationnaire.

Solutions de l'équation autonôme. On considère le problème de Cauchy

$$y' = f(y), \quad y(t_0) = y_0$$

pour une fonction f continue telle que $f(y_0) \neq 0$ (i.e. y_0 n'est pas une solution stationnaire). Soit F une primitive de la fonction $1/f$, i.e.

$$\frac{d\,F(y)}{dy} = \frac{1}{f(y)}$$

alors, la solution unique du problème de Cauchy $y(t)$ satisfait la rélation

$$F(y(t)) - F(y_0) = t - t_0.$$

Ensuite il faut resoudre cette équation pour $y(x)$. Noter que F est localment inversible autour de y_0 car $\frac{d\,F(y)}{dy}|_{y=y_0} = \frac{1}{f(y_0)} \neq 0$. Par conséquent, dans un voisinage de t_0 nous avons

$$y(x) = F^{-1}\big(F(y_0) + t - t_0\big).$$

Remarque. La solution $y(t)$ est unique pour tout x tel que $f(y(x)) \neq 0$. En effet, si $z(t)$ est une autre solution avec $f(z(x)) \neq 0$, alors $F(z(t)) = F(y(t))$ car

$$\frac{d\,F(z(t))}{dt} = \frac{z'(t)}{f(z(t))} = 1 = \frac{y'(t)}{f(y(t))} = \frac{d\,F(y(t))}{dt}$$

et $F(z(t_0)) = F(y(t_0))$. Par le théorème des accroissements finis il existe $c = c(t)$ enter $y(t)$ et $z(t)$ tel que $f(c(t)) \neq 0$ et

$$0 = F(z(t)) - F(y(t)) = \frac{1}{f(c(t))}\big(z(t) - y(t)\big)$$

i.e. $z(t) = y(t)$.

Remarque. Pour l'unicité de la solution du problème de Cauchy la condition $f(y_0) \neq 0$ est nécessaire et suffisante. Par exemple,

$$y' = 2\sqrt{|y|}, \quad y(0) = 0$$

admet deux solutions distinctes $y_1(t) = 0$ et $y_2(t) = t^2$. Pour garantir l'unicité de solution stationnaire f doit vérifier des hypothèses plus fortes : Une condition suffisante est que f soit de classe C^1.

Remarque. Formellement on peut écrire l'équation différentielle $\frac{d\,y}{dt} = f(y)$ comme

$$dy = f(y)\,dt \quad \text{ou} \quad \frac{d\,y}{f(y)} = dt$$

et l'on intègre pour obtenir

$$\int_{y_0}^{y} \frac{d\,\eta}{f(\eta)}\,d\eta = \int_{t_0}^{t} d\xi$$

i.e.

$$F(y) - F(y_0) = t - t_0.$$

Remarque. Donner la solution du problème de Cauchy

$$y' = 1 - y^2, \quad y(1) = 0.$$

On a

$$\int_{0}^{y} \frac{d\,\eta}{1-\eta^2}\,d\eta = \int_{1}^{t} d\xi$$

ou

$$\operatorname{arctanh} y(t) = t - 1 \quad \text{i.e.} \quad y(t) = \tanh(t-1)$$

7.3.4 La technique du changement des variables

On peut essayer de transformer l'équation différentielle $y' = f(t, y)$ par l'application $(t, y) \mapsto (s, u)$ où $s = s(t, y)$ et $u = h(t, y)$. Dans la suite ,nous présentons quelques exemples :

Exemple 1. Le changement des variables

$$s = s(t), \quad y(t) = a(t)u(s), \quad a'(t), a(t) > 0,$$

permet de transformer l'équation différentielle

$$y' = a'(t)f(y(t)/a(t))$$

par $y'(t) = a'(t)u(s)+a(t)s'(t)\dot{u}(s)$ avec $\dot{u}(s) = \dfrac{d\,u(s)}{ds}$ en l'équation différentielle

$$a(t)s'(t)\dot{u}(s) = a'(t)f(u(s)) - a'(t)u(s)$$

que devient une équation autonôme si on pose $s(t) = \ln a(t)$.

L'équation de Riccati. Il y a des transformations qui donnent une équation différentielle du second ordre. Pour trouver la solution générale de

$$y' = -y^2 + q(x)$$

on pose $y = \dfrac{u'}{u}$, $u > 0$. On obtient

$$u'' = q(x)u$$

Systèmes autonômes - une équation pour les orbits dans l'espace de phase. Considérons le système autonôme

$$\dot{x} = -g(x, y), \quad \dot{y} = f(x, y) \tag{7.10}$$

Si on ne peut pas résoudre ce système on peut formuler une équation différentielle du prémier ordre en considérant y comme fonction de x. En posant $y(t) = Y(x(t))$ on obtient

$$f(x, Y(x)) = \dot{y} = \frac{d\,Y}{dx}\dot{x} = -\frac{d\,Y}{dx}\, g(x, Y(x)).$$

Ceci donne (on maintien par commodité la notation y au lieu de Y) une équation différentielle du premier ordre :

$$f(x, y) + g(x, y)\frac{d\,y}{dx} = 0, \quad \text{ou} \quad \frac{d\,y}{dx} = -\frac{f(x, y)}{g(x, y)}.$$

7.3.5 Équations différentielles exactes

La forme la plus générale d'une équation différentielle est $f(x, y, y') = 0$ appelée équation différentielle implicite. Un cas particulier est donné par des équations différentielles de la forme $f(x, y) + g(x, y)y' = 0$ pour lesquelles on peut sous certaines conditions trouver des solutions donnée par une équation de la forme $H(x, y(x)) = 0$. En fait, si$D \subset \mathbb{R}^2$ et $H : D \longrightarrow \mathbb{R}$ est de classe C^1, alors pour $(t_0, x_0) \in \overset{\circ}{D}$ une solution du probleème de Cauchy

$$\frac{\partial H(x,y)}{\partial x} + \frac{\partial H(x,y)}{\partial y}\, y' = 0, \quad y(x_0) = y_0$$

est donné par l'équation $H(x, y(x)) = H(x_0, y_0)$.

Équation différentielle exacte. Soient $D =: [a, b] \times [c, d]$ et $f, g : D \longrightarrow \mathbb{R}$ de classe C^1. L'équation différentielle

$$f(x, y) + g(x, y)y' = 0 \tag{7.11}$$

est dite exacte si

$$\frac{\partial f(x,y)}{\partial y} = \frac{\partial g(x,y)}{\partial x} \tag{7.12}$$

Dans ce cas il existe $H : D \longrightarrow \mathbb{R}$ de classe C^2 donné par (voir l'exercice 27 du chapitre 6 - construction du potentiel)

$$H(x, y) = \int_0^1 f(x_t, y_t)(x - x_0) + g(x_t, y_t)(y - y_0)\, dt + \text{const.} \tag{7.13}$$

où $(x_t, y_t) = (tx + (1 - t)x_0, ty + (1 - t)y_0) \in \overset{\circ}{D}$. Alternativement, on peut trouver un $H(x, y)$ comme suit : Soit

$$F(x, y) = \int_{x_0}^{x} f(z, y)\, dz$$

donc $\dfrac{\partial F(x,y)}{\partial x} = f(x, y)$. On pose $H(x, y) = F(x, y) + K(y)$. Evidemment, $\dfrac{\partial H(x,y)}{\partial x} = f(x, y)$ et $\dfrac{\partial H(x,y)}{\partial y} = g(x, y)$ implique $K'(y) = g(x, y) - \dfrac{\partial F(x,y)}{\partial y}$. Le membre de droite est effectivement indéépendant de x (justifiant notre pose de H) puisque

$$\frac{\partial}{\partial x}\left(g(x, y) - \frac{\partial F(x,y)}{\partial y}\right) = \frac{\partial g(x,y)}{\partial x} - \frac{\partial f(x,y)}{\partial x} = 0.$$

On doit résoudre une équation différentielle autonôme pour K.

Application aux systèmes autonômes. Si les fonctions f et g vérifiont la condition (7.12) le système autonôme (7.10) s'écrit comme le système "Hamiltonien"

$$\dot{x} = -\frac{\partial H(x,y)}{\partial y}, \quad \dot{y} = \frac{\partial H(x,y)}{\partial x}. \tag{7.14}$$

Si $(x(t), y(t))$ est une solution de (7.14), alors $H(x(t), y(t)) = const$ interprété comme la conservation de "l'énergie $H(x, y)$". Si on identifie le système (7.14) avec un système mécanique (mouvement d'une particule en une dimension) , alors x correspond à la quantité de mouvement et y correspond à la position (voir aussi section 7.4.5).

Facteur intégrant. Si l'équation différentielle (7.11) n'est pas exacte on peut dans certain cas multiplier (7.11) par une fonction $\lambda(x, y)$ (appelé facteur intégrand) qui rend (7.11) exacte, c'est-à-dire que

$$\frac{\partial(\lambda(x,y)f(x,y))}{\partial y} = \frac{\partial(\lambda(x,y)g(x,y))}{\partial x}.$$

Par exemple, l'équation différentielle linéaire inhomogène $y' + a(x)y - b(x) = 0$ n'est pas excacte ($g(x, y) = 1, f(x, y) = a(x)y - b(x)$). Un facteur intégrand est $\lambda(x) = e^{A(x)}$, $A' = a$. On trouve $H(x, y) = e^{A(x)}y - k(x)$ où $k'(x) = b(x)e^{A(x)}$.

7.4 Equations différentielles du second ordre

7.4.1 Equations linéaires homogènes - propriétés générales

Soient $p, q : I \longrightarrow \mathbb{R}$ continues.On appelle équation différentielle linéaire homogène du second ordre une équation de la forme

$$y''(t) + p(t)y'(t) + q(t)y(t) = 0. \tag{7.15}$$

A l'execption des quelques cas particuliers, on ne peut pas donner les solutions explicites de cette équation.

Proposition. Soient $y_1(t), y_2(t)$ deux solutions de (7.3). Alors pour tout $\alpha, \beta \in \mathbb{R}$ la combinaison linéaire $\alpha y_1(t) + \beta y_2(t)$ est aussi une solution de (7.3)

Proposition. Le problème de Cauchy pour l'équation différentielle linéaire homogène du second ordre, i.e.

$$y''(t) + p(t)y'(t) + q(t)y(t) = 0, \quad \text{et} \quad y(t_0) = y_0, y_0'(t_0) = v_0$$

admet une solution unique pour tout couple $(y_0, v_0) \in \mathbb{R}^2$.

Définition. Les deux fonctions $y_1, y_2 : I \longrightarrow \mathbb{R}$ sont dites linéairement indépendantes si

$$(\alpha y_1(t) + \beta y_2(t) = 0 \quad \text{pour tout}\, t \in I) \Rightarrow \alpha = \beta = 0$$

Définition. Soient $y_1, y_2 : I \longrightarrow \mathbb{R}$ deux fonctions de classe $C^1(I)$. On appelle le wronskien de y_1, y_2 la fonction $w : I \longrightarrow \mathbb{R}$ définie par

$$w(t) = y_1(t)y_2'(t) - y_1'(t)y_2(t)$$

Proposition. Soient $y_1(t), y_2(t)$ deux solutions de (7.3). Alors, le wronskien de y_1, y_2 est donné par

$$w(t) = w(t_0) \exp\big(- \int_{t_0}^{t} p(s)\, ds\big)$$

Démonstration. $w'(t) = y_1(t)y_2''(t) - y_1''(t)y_2(t) = -p(t)w(t)$.

Remarque. Par conséquent on a l'alternative suivante : ou bien le wronskien est toujours nul su I ou béien il ne s'annule jamais sur I.

Proposition. Deux solutions $y_1(t), y_2(t)$ de (7.3) sont linéairement indépendantes si et seulement si leur wronskien est non-zéro.

Proposition. L'équation (7.3) admet deux solution linéairement indépendantes $y_1(t), y_2(t)$ est toute solution $y(t)$ de (7.3) est de la forme

$$y(t) = c_1 y_1(t) + c_2 y_2(t)$$

où c_1 et c_2 sont des constantes.

Remarque. Pour trouver la solution unique du problème de Cauchy à partir de la solution générale on résoud le système d'équations linéaires

$$y_0 = c_1 y_1(t_0) + c_2 y_2(t_0) \quad v_0 = c_1 y_1'(t_0) + c_2 y_2'(t_0)$$

pour c_1 et c_2.

Remarque. Si on a une solution $y_1(t)$ de (7.3) on trouve une autre solution $y_2(t)$ linéairement indépendante de $y_1(t)$ par le wronskien : pour un $t_0 \in I$ on pose $w(t_0) = 1$ et on résoud la relation

$$w(t) = \exp\big(- \int_{t_0}^{t} p(s)\, ds\big)$$

pour $y_2(t)$ en utilisant que

$$w(t) = y_1(t)y_2'(t) - y_1'(t)y_2(t) = y_1^2(t)\left(\frac{y_2(t)}{y_1(t)}\right)'.$$

Par conséquent, une autre solution $y_2(t)$ est donnée par

$$y_2(t) = y_1(t) \int_{t_0}^{t} \frac{w(s)}{y_1^2(s)}\, ds = y_1(t) \int_{t_0}^{t} \frac{e^{-\int_{t_0}^{s} p(\tau)\, d\tau}}{y_1^2(s)}\, ds$$

Interprétation comme système des équations du premier ordre. Les notions "linéairement indépendant" et le wronskien sont plus inituitives pour le système des système des équations du premier ordre associé à (7.3) :

$$\frac{d}{dt}\begin{pmatrix} y \\ y' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -q(t) & -p(t) \end{pmatrix}\begin{pmatrix} y \\ y' \end{pmatrix} \tag{7.16}$$

Deux solutions $y_1(t), y_2(t)$ sont dites linéarement indépendantes si les vecteurs $\mathbf{y}_1(t) = \begin{pmatrix} y_1(t) \\ y_1'(t) \end{pmatrix}, \mathbf{y}_2(t) = \begin{pmatrix} y_2(t) \\ y_2'(t) \end{pmatrix}$ sont deux vecteurs linéarement indépendants pour tout $t \in I$. Le wronskien est le determinant de la matrice formée par les vecteurs $\mathbf{y}_1(t), \mathbf{y}_2(t)$:

$$w(t) = \det\begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix}$$

Le wronskien donne l'aire (plus precisement l'aire orienté car le wronskien peut être négatif) du parallélogramme

$$D(t) = \{\alpha\mathbf{y_1}(\mathbf{t}) + \beta\mathbf{y_2}(\mathbf{t}), 0 \leq \alpha \leq 1, 0 \leq \beta \leq 1\}.$$

Par les résultats ci-dessus les vecteurs $\mathbf{y}_1(t), \mathbf{y}_2(t)$ sont linéarement indépendants pour tout $t \in I$ si et seulement si $\mathbf{y}_1(t_0), \mathbf{y}_2(t_0)$ sont linéarement indépendants pour un $t_0 \in I$. Par conséquent, les solutions du système (7.4) forment un espace vectoriel de dimension deux.

7.4.2 Equations linéaires homogènes à coefficients constants

On considère

$$y''(t) + p\,y'(t) + q\,y(t) = 0 \quad \text{où} \quad p, q \in \mathbb{R}. \tag{7.17}$$

On essaie si $y(t) = exp(\lambda t)$ est une solution pour un λ donné. On trouve que λ doit être une racine du polynôme charactéristique de la matrice assocciée

$$A = \begin{pmatrix} 0 & 1 \\ -q & -p \end{pmatrix}$$

, i.e.

$$\lambda^2 + p\,\lambda + q = 0.$$

Noter que les racines donnent les valeurs propres λ_1, λ_2 de la matrice A. On distinct les cas suivants :

Deux valeurs propres réelles distinctes. $\lambda_1, \lambda_2 \in \mathbb{R}$ et $\lambda_1 \neq \lambda_2$. La solution générale est donnée par

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}$$

Deux valeurs propres distinctes complexes conjugées. $\lambda_1 = \mu + i\omega, \lambda_2 = \mu - i\omega$, $\mu, \omega \in \mathbb{R}, \omega \neq 0$. Les fonctions exponentielles sont des solutions complexes. pour obtenir des solutions réelles on note que la partie imaginaire et la partie réelle sont également des solutions. La solution générale est donnée par

$$y(t) = c_1 e^{\mu t}\cos(\omega t) + c_2 e^{\mu t}\sin(\omega t)$$

Une valeur propre réelle double. $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$. On trouve une deuxième solution par la formule du wronskien donnée par $te^{\lambda\, t}$. La solution générale est donnée par

$$y(t) = c_1 e^{\lambda t} + c_2 t\, e^{\lambda t} = (c_1 + c_2 t)e^{\lambda t}$$

7.4.3 Equations linéaires inhomogènes

Soient $p, q, f : I \longrightarrow \mathbb{R}$ continues.On appelle équation différentielle linéaire inhomogène du second ordre une équation de la forme

$$y''(t) + p(t)y'(t) + q(t)y(t) = f(t). \tag{7.18}$$

Proposition. Toute solution $y(t)$ de (7.6) est de la forme

$$y(t) = c_1 y_1(t) + c_2 y_2(t) + y_p(t)$$

où $y_1(t), y_2(t)$ sont deux solutions linéairement indépendantes de l'équation homogène associée et $y_p(t)$ est une solution particulière de (7.6)

Construction d'une solution particulière dans le cas général. Soient $y_1(t), y_2(t)$ deux solutions linéarement indépendantes de l'équation homogéène (7.3) et $w(t) = y_1(t)y_2'(t) - y_1'(t)y_2(t)$ son wronskien. On définit la fonction $k : I \times I \longrightarrow \mathbb{R}$ par

$$k(s,t) = \frac{y_1(s)y_2(t) - y_1(t)y_2(s)}{y_1(s)y_2'(s) - y_1'(s)y_2(s)} = \frac{y_1(s)y_2(t) - y_1(t)y_2(s)}{w(s)}.$$

Proposition. Une solution particulière de (7.6) $y_p(t)$ est donnée par

$$y_p(t) = \int_{t_0}^{t} k(s,t)f(s)\, ds$$

Démonstration. Notant que $k(t,t) = 0$ nous avons

$$y_p'(t) = k(t,t)f(t) + \int_{t_0}^{t} D_t k(s,t)f(s)\, ds = \int_{t_0}^{t} D_t k(s,t)f(s)\, ds$$

et

$$y_p''(t) = D_t k(t,t)f(t) + \int_{t_0}^{t} D_{tt} k(s,t)f(s)\, ds.$$

Donc

$$y_p''(t)+p(t)y_p'(t)+q(t)y_p(t) = D_t k(t,t)f(t)+\int_{t_0}^{t}\Big(D_{tt}k(s,t)+p(t)D_t k(s,t)+q(t)k(s,t)\Big)f(s)ds.$$

En utilisant

$$D_t k(s,t) = \frac{y_1(s)y_2'(t) - y_1'(t)y_2(s)}{w(s)}$$

et

$$D_{tt} k(s,t) = \frac{y_1(s)y_2''(t) - y_1''(t)y_2(s)}{w(s)}$$

on voit que

$$D_t k(t,t) = \frac{y_1(t)y_2'(t) - y_1'(t)y_2(t)}{w(t)} = 1$$

et

$$D_{tt}k(s,t) + p(t)D_t k(s,t) + q(t)k(s,t) = 0,$$

donc $y_p(t)$ est une solution particulière de (7.6).

Remarque. Bien que cette méthode est générale il y a souvent des techniques plus efficaces pour trouver une solution particulière.

7.4.4 Equations linéaires inhomogènes à coefficients constants

Soit $f : I \longrightarrow \mathbb{R}$ continue. On considère

$$y''(t) + p\,y'(t) + q\,y(t) = f(t). \tag{7.19}$$

avec $p, q \in \mathbb{R}$. Calculons $k(s,t)$:

Deux valeurs propres réelles distinctes. $\lambda_1, \lambda_2 \in \mathbb{R}$ et $\lambda_1 \neq \lambda_2$.

$$k(s,t) = \frac{e^{\lambda_2(t-s)} - e^{\lambda_1(t-s)}}{\lambda_2 - \lambda_1}$$

Deux valeurs propres distinctes complexes conjugées. $\lambda_1 = \mu + i\omega, \lambda_2 = \mu - i\omega$, $\mu, \omega \in \mathbb{R}, \omega \neq 0$. Il est plus facile de calculer $k(s,t)$ à partir des solutions complexes :

$$k(s,t) = e^{\mu(t-s)}\frac{e^{i\omega(t-s)} - e^{-i\omega(t-s)}}{2i\omega} = e^{\mu(t-s)}\frac{\sin\omega(t-s)}{\omega}$$

Une valeur propre réelle double. $\lambda_1 = \lambda_2 = \lambda \in \mathbb{R}$. Alors

$$k(s,t) = (t-s)e^{\lambda(t-s)}$$

Exemple. Résoudre le problème de Cauchy

$$y''(t) + p\,y'(t) = -g, \quad \text{et} \quad y(0) = h, y_0'(0) = 0$$

où$g > 0$. Cet problème décrit la chute libre avec la résitance d'air ($p > 0$) en partant de l'hauteur h.

Application de la méthode générale. Les valeurs propres sont 0 et $-p$, donc $y_1(t) = e^{-pt}$ et $y_2(t) = 1$ deux solutions linéairement indépendantes de l'équation homogèène. On a

$$k(s,t) = \frac{1 - e^{-p(t-s)}}{p}$$

et

$$y_p(t) = -g\int_0^t k(s,t)\,ds = g\frac{1 - p\,t - e^{-p\,t}}{p^2}$$

On en déduit que $y_p(0) = 0$ et $y_p'(0) = 0$. Donc

$$y(t) = h + y_p(t)$$

est une solution du problème de Cauchy.

Méthode alternative. L'équation est du premier ordre pour y'. Par 7.2.2 une solution particulière est

$$y'(t) = -g\frac{1-e^{-p\,t}}{p}$$

d'où après intégration $y(t) = h + y_p(t)$.

Equations avec $\mathbf{f(t) = E\sin(\omega t)}$ ou $\mathbf{f(t) = E\cos(\omega t)}$. Pour trouver une solution particulière de (7.7) où (i)$f(t) = E\sin(\omega t)$ ou (ii) $f(t) = E\cos(\omega t)$ on passe l'équation dans les nombres complexes et on cherche une solution particulière $z_p : \mathbb{R} \longrightarrow \mathbb{C}$ de

$$z'' + pz' + qz = Ee^{i\omega t}$$

Ensuite on prend $y_p = \mathrm{Im} z_p$ dans le cas (i) et $y_p = \mathrm{Re} z_p$ dans le cas (ii). On essaie l'Ansatz $z = Ae^{i\omega t}$, $A \in \mathbb{C}$. Son insertion dans l'équation différentielle donne la condition

$$(-\omega^2 + ip\omega + q)Ae^{i\omega t} = Ee^{i\omega t}.$$

Si $-\omega^2 + ip\omega + q \neq 0$ la solution est donnée par $A = \dfrac{E}{-\omega^2 + ip\omega + q}$ et par conséquent

$$z_p(t) = \frac{E}{-\omega^2 + ip\omega + q}\, e^{i\omega t}.$$

Exemple. Donner une solution particulière de

$$y''(t) + 2y'(t) + y(t) = E\sin(\omega\, t)$$

où$E > 0$, $\omega \neq 0$. On trouve la condition

$$A = \frac{E}{1-\omega^2+2i\omega} = \frac{E(1-\omega^2-2i\omega)}{(1+\omega^2)^2}$$

Donc

$$y_p(t) = \mathrm{Im}(A\, e^{i\,\omega\, t}) = E\frac{(1-\omega^2)\sin(\omega\, t) - 2\omega\cos(\omega\, t))}{(1+\omega^2)^2}$$

Exemple. Donner une solution particulière de

$$y''(t) + \omega_0^2 y(t) = E\sin(\omega\, t).$$

Si $\omega^2 \neq \omega_0^2$, alors

$$y_p(t) = \mathrm{Im}(\frac{E}{-\omega^2+\omega_0^2}\, e^{i\,\omega\, t}) = \frac{E\sin\omega t}{\omega_0^2 - \omega^2}.$$

Si $\omega^2 = \omega_0^2$ on fait un Ansatz de variation des constantes $z(t) = A(t)e^{i\,\omega\, t}$ pour trouver $A(t) = \dfrac{-iEt}{2\omega}$ donc

$$y_p(t) = \mathrm{Im}(\frac{-iEt}{2\omega}\, e^{i\,\omega\, t}) = \frac{-iEt\cos(\omega t)}{2\omega} = \frac{-iEt\sin(\omega t + \frac{\pi}{2})}{2\omega}.$$

La "phase" $\dfrac{\pi}{2}$ signifie le retard du système.

7.4.5 Equations autonômes et systèmes conservatifs

Soit $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$ continue. On appelle équation différentielle autonôme du second ordre une équation de la forme

$$\ddot{y}(t) = -f(y(t), \dot{y}(t)).$$

Cetteéquation différentielle a de nouveau la propriéte d'être "invariante sous translations en t", i.e. si $y(t)$ est une solution alors $y(t+s)$ est une solution pour tout s réel fixe. Si la fonction f ne dépend pas de y', i.e. $f = f(y)$ on appelle cette équation conservative. En introduisant un paramètre $m > 0$ (la masse de la particule) le problème de Cauchy associé est

$$m\ddot{y}(t) = -f(y) \quad \text{et} \quad y_(t_0) = y_0, \dot{y}(t)(t_0) = v_0 \tag{7.20}$$

Définition. Soit F une primitive de f et $E : \mathbb{R}^2 \longrightarrow \mathbb{R}$ définie par

$$E(y, p) = \frac{1}{2m}p^2 + F(y)$$

Si $y(t)$ est une solution de (7.20) on appelle $E(y(t), m\dot{y}(t))$ l'énergie de $y(t)$.

Proposition. Soit $y(t)$ une solution de (7.20), alors

$$\frac{d}{dt}\, E(y(t), m\dot{y}(t)) = 0$$

Par conséquent,

$$E(y, m\dot{y}) = \frac{m\dot{y}^2}{2} + F(y) = \text{const} = E(y_0, v_0) := E_0$$

et on peut reduire le problème à une équation autonôme du premier ordre :

$$\dot{y} = \pm\sqrt{\frac{2E_0 - 2F(y)}{m}}\,.$$

Remarquons encore que l'équation différentielle de (7.20) est équivalente à un système Hamiltonien de la forme (7.14) (voir section 7.3.5). En fait si on pose $p = m\dot{y}$, alors

$$\dot{p} = -\frac{\partial E(y,p)}{\partial y} = -f(y), \quad \dot{y} = \frac{\partial E(y,p)}{\partial p} = \frac{p}{m}. \tag{7.21}$$

7.5 Systèmes d' équations différentielles linéares à coefficients constantes

7.5.1 Systèmes homogènes

Soit $A \in M_{n,n}(\mathbb{R})$ (ou même $A \in M_{n,n}(\mathbb{R})$) . On considère le système homogène d'équations différentielles

$$\mathbf{y}'(t) = A\mathbf{y}(t) \tag{7.22}$$

et le problème de Cauchy y asscoié :

$$\mathbf{y}'(t) = A\mathbf{y}(t), \quad \mathbf{y}(0) = \mathbf{y}_0. \tag{7.23}$$

Le théorème sur l'existence et l'unicité locale des solutions nous assure l'existence d'une unique solution $\mathbf{y}(t)$ du problème (7.23). Cette solution peut être prolongée sur tout $\mathbb{R}$. En fait, $f(\mathbf{y}) = A\mathbf{y}$ est une fonction lipschitzienne avec $L = ||A||_2$ sur tout $\mathbb{R}^n$. On peut donc choisir $R_{a,b}$ tel que $\alpha = ||A||_2$ pour tout condition initiale $\mathbf{y}_0$. Par itération on prolonge la solution sur $\mathbb{R}$. De même pour les solutions de (7.22).

Proposition - Superposition des solutions. Soient $\mathbf{y}_1(t), \mathbf{y}_2(t)$ de (7.22). Alors, pour tous $\alpha, \beta \in \mathbb{R}$ la combinaison linéaire $\alpha\mathbf{y}_1(t)+\beta\mathbf{y}_2(t)$ est une solution de (7.22).

En analogie avec la discussion dans le ch. 7.3 n solutions $\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t)$ sont linéairement indépendantes si et seulement si les vecteurs $\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t)$ sont linéairement indépendants pour tout t. Rappelons que $\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t)$ sont linéairement indépendants si et seulement si la matrice

$$Y(t) = (\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t))$$

est inversible ou bien si et seulement si "le wronskien" $w(t) = \det Y(t)$ est non zéro. Pour trouver n solutions linéairement indépendantes nous cherchons en générale $\mathbf{y}_1(t), \ldots, \mathbf{y}_n(t)$ telles que $\mathbf{y}_1(0) = \mathbf{e}_1, \ldots, \mathbf{y}_n(0) = \mathbf{e}_n$, c'est-à-dire $Y(0) = Id_n$ (la matrice d'identité) Il en suit :

Proposition : Toute solution $\mathbf{y}(t)$ de (7.23) est de la forme

$$\mathbf{y}(t) = \sum_{i=1}^{n} c_i \mathbf{y}_i(t)$$

Équation différentielle matricielle. Par conséquent, nous cherchons à résoudre une équation différentielle pour une matrice $n \times n$ $Y(t)$:

$$Y'(t) = AY(t), \quad Y(0) = Id_n = 1. \tag{7.24}$$

Son analogie avec l'équation pour une seule fonction est parfaite : dans ce cas la solution était $y(t) = \exp(At)$ (voir ch. 7.2) Dans notre cas nous définissons pour une matrice $A \in M_{n,n}(\mathbb{C})$ l'exponentiel par la série

$$\exp(A) = \sum_{k=0}^{\infty} \frac{A^k}{k!} = 1 + A + \frac{A^2}{2} + \ldots.$$

Cette série est absolument convergente par rapport à la norme $||A||_2$. Grâce aux propriétés de séries absolument convergente nous avons le résultat suivant :

Theorème - solution de l'équation différentielle matricielle. La solution de (7.24) est donnée par

$$Y(t) = \exp(At) \tag{7.25}$$

et la solution du problème de Cauchy

$$\mathbf{y}'(t) = A\mathbf{y}(t), \quad \mathbf{y}(0) = \mathbf{y}_0 \tag{7.26}$$

est donnée par

$$\mathbf{y}(t) = \exp(At)\mathbf{y}_0 \tag{7.27}$$

Dans la suite nous présentons quelques résultats utiles pour le calcul de $\exp(A)$.

Proposition - Propriétés de la fonction exponentielle. Soient $A, B \in M_{n,n}(\mathbb{C})$. Alors,

1. $$\exp(A)\exp(B) = \exp(A+B) \quad \text{si } [A,B] := AB - BA = 0,$$
2. $$\exp(B^{-1}AB) = B^{-1}\exp(A)B \quad \text{si } \det B \neq 0.$$
3. $$\exp(A) = \lim_{k\to\infty}\left(1 + \frac{A}{k}\right)^k.$$
4. $$\det(\exp(A)) = e^{\mathrm{Tr}(A)}.$$
5. Pour une matrice diagonale :
$$\exp(\sum_{k=1}^{n}\lambda_k E_{kk}) = \sum_{k=1}^{n} e^{\lambda_k} E_{kk}.$$

Exemples.

1. $$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \exp(At) = E_2 + At = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}$$
car $A^2 = A^3 = \ldots = 0$.
2. $$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \exp(At) = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}$$
car $A^{2n} = (-1)^n E_2$, $A^{2n+1} = (-1)^n A$.
3. $$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \quad \exp(At) = \begin{pmatrix} e^{at} & 0 \\ 0 & e^{bt} \end{pmatrix}$$

L'idée principale consists à simplifier la matrice par des transfomations (par exemple la diagonaliser, si possible) pour calculer son exponentiel. Cette procedure est liée à la détemination des valeurs propres et de vecteurs propres.