



Année académique 2026-2027

MATH 2014 Compléments de mathématiques

Table des matières

1	Première partie	5
1.1	Rappels concernant les complexes	5
1.1.1	Définitions de l'ensemble des complexes et de deux opérations entre complexes . .	5
1.1.2	Propriétés	6
1.1.3	Introduction du complexe i et notations pratiques	7
1.1.4	Module et conjugué d'un complexe	8
1.1.5	Racines carrées d'un nombre complexe	9
1.1.6	Trinôme du second degré	9
1.1.7	Complexes et trigonométrie	9
1.2	(Rappels concernant les) fonctions de plusieurs variables	12
1.3	Rappels concernant le calcul intégral	13
1.4	Intégrales paramétriques	13
1.5	Transformation de Fourier des fonctions intégrables sur $\mathbb{R}$	14
1.5.1	Définition et interprétation	14
1.5.2	Exemples	16
1.5.3	Premières propriétés	17
1.5.4	Dérivation et transformation de Fourier	17
1.5.5	Intégration et transformation de Fourier	18
1.5.6	Théorème de transfert et de Fourier	18
1.6	Produit de composition	18
2	Seconde partie	21
2.1	Le développement limité de Taylor	21
2.2	Extrema : définitions et première propriété	22
2.3	Compléments de calcul matriciel	23
2.4	Retour aux extrema : cas des extrema dits « libres »	25
2.5	Application à la régression linéaire	27
2.6	Les extrema liés (extrema « sous contrainte »)	28
2.7	Les extrema sur des ensembles fermés bornés	29

Chapitre 1

Première partie

1.1 Rappels concernant les complexes

L'histoire veut qu'après les naturels on introduise les entiers pour résoudre des équations telles que $x + 4 = 1$. Ensuite on introduit les rationnels, qui permettent de résoudre des équations du type $3x = 1$. Les réels, quant à eux, donnent les solutions à des égalités du type ¹ $x^2 = 2$. Et les complexes permettent de résoudre par exemple $x^2 = -4$, comme on va le voir dans cette section.

1.1.1 Définitions de l'ensemble des complexes et de deux opérations entre complexes

Définition 1.1.1 L'ensemble des nombres complexes, noté $\mathbb{C}$, est l'ensemble des couples de réels

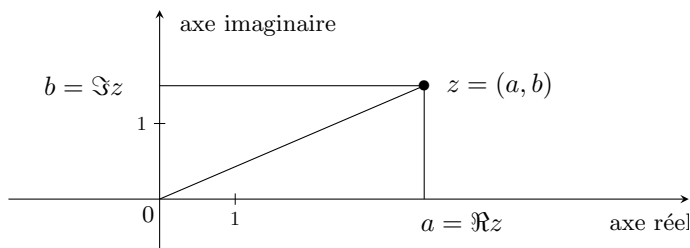
$$\mathbb{C} = \{(a, b) : a, b \in \mathbb{R}\}.$$

Par définition, deux complexes (a, b) et (a', b') sont égaux lorsque $a = a'$ et $b = b'$.

On a directement à notre disposition une représentation graphique de $\mathbb{C}$: si on considère le plan muni d'un repère orthonormé, tout point du plan définit un complexe et tout complexe définit un point du plan.

Si $z = (a, b)$ est un complexe, le réel a s'appelle *la partie réelle* du complexe et le réel b s'appelle *la partie imaginaire* du complexe. On utilise les notations

$$\Re z = a, \quad \Im z = b.$$



On dit que l'ensemble $\mathbb{R}$ des réels est inclus dans l'ensemble $\mathbb{C}$ des complexes en identifiant les réels aux couples $(a, 0)$, $a \in \mathbb{R}$. Les réels sont donc les complexes dont la partie imaginaire est nulle. Le complexe nul est le couple $(0, 0)$. Un nombre complexe dont la partie réelle est nulle et dont la partie imaginaire est non nulle est appelé *nombre complexe imaginaire pur*.

Dans l'ensemble des nombres complexes, on définit deux opérations fondamentales, l'*addition de deux complexes* et la *multiplication de deux complexes*. L'addition aura immédiatement une interprétation claire (elle se traduira par l'addition de deux vecteurs du plan). Quant à la multiplication, on verra son interprétation plus tard, à l'aide de rotations ; il faut bien se garder de l'interpréter à l'aide d'un produit quelconque de vecteurs !!

1. A titre d'exercice, montrer que $\sqrt{2}$ n'est pas un nombre rationnel.

Définition 1.1.2 *Addition de deux complexes, opération notée $+$: $(a, b) + (c, d) = (a + c, b + d)$. Multiplication de deux complexes, opération notée $\times$ (ou encore par un blanc, comme dans le cadre réel) : $(a, b) \times (c, d) = (ac - bd, ad + bc)$.*

Remarquons que l'on a le cas particulier suivant (a, b, r sont des réels) :

$$(r, 0) \times (a, b) = (ra - 0b, rb + 0a) = (ra, rb) = (a, b) \times (r, 0),$$

ce qui s'écrit, si on identifie les réels aux couples de complexes de partie imaginaire nulle,

$$r(a, b) = (ra, rb).$$

En particulier, pour $r = 1$, c'est-à-dire le complexe $(1, 0)$, on a

$$(1, 0) \times (a, b) = (a, b) = (a, b) \times (1, 0),$$

ce qui signifie que 1 est neutre pour la multiplication. Et pour $a = 0$ et $b = 1$, on obtient (cela sera utilisé dans la suite)

$$(r, 0) \times (0, 1) = (r0 - 01, r1 + 0) = (0, r) = (0, 1) \times (r, 0)$$

ou encore

$$r(0, 1) = (0, r) = (0, 1)r.$$

Il est clair (et important !) de noter que ces opérations d'addition et de multiplication que l'on vient de définir, restreintes à $\mathbb{R}$, rendent les opérations usuelles de $\mathbb{R}$!

1.1.2 Propriétés

La première propriété est une généralisation de ce qui se passe dans $\mathbb{R}$.

Propriété(s) 1.1.3 *Soient (a, b) et (c, d) deux complexes. On a*

$$(a, b) \times (c, d) = (0, 0) \quad \text{si et seulement si} \quad (a, b) = (0, 0) \quad \text{ou} \quad (c, d) = (0, 0).$$

Autrement dit, le produit de deux complexes est nul si et seulement si l'un d'entre eux est nul.

Preuve. Voir par exemple le syllabus de MATH2007 $\square$

Passons aux autres propriétés essentielles des opérations introduites.

Propriété(s) 1.1.4 *Pour l'addition et la multiplication entre deux complexes, on a les propriétés suivantes :*

- l'ensemble $\mathbb{C}$ muni de l'addition est un groupe commutatif de neutre $0 = (0, 0)$ ²
- l'ensemble $\mathbb{C} \setminus \{(0, 0)\}$ muni de la multiplication est un groupe commutatif de neutre $(1, 0)$ ³
- la multiplication distribue l'addition⁴.

On dit que $\mathbb{C}$ muni de l'addition $+$ et de la multiplication $\times$ est un corps commutatif.

2. Cela signifie que l'addition possède les propriétés suivantes :

- associativité : pour tous complexes z_1, z_2, z_3 , on a $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$
- existence d'un neutre : le complexe $e = (0, 0)$ est tel que $e + z = z + e = z$ pour tout complexe z
- tout complexe possède un symétrique (ici, on parle aussi d'opposé) : pour tout z , il existe z' tel que $z + z' = e = z' + z$
- commutativité : pour tous complexes z, z' , on a $z + z' = z' + z$.

3. Cela signifie que la multiplication possède les propriétés suivantes :

- associativité : pour tous complexes z_1, z_2, z_3 , on a $(z_1 \times z_2) \times z_3 = z_1 \times (z_2 \times z_3)$
- existence d'un neutre : le complexe $e = (1, 0)$ est tel que $e \times z = z \times e = z$ pour tout complexe z
- tout complexe non nul possède un symétrique (ici, on parle plutôt d'inverse) : pour tout $z \neq 0$, il existe z' tel que $z \times z' = e = z' \times z$
- commutativité : pour tous complexes z, z' , on a $z \times z' = z' \times z$.

4. Cela signifie que pour tous complexes c, z_1, z_2 , on a $c \times (z_1 + z_2) = c \times z_1 + c \times z_2$.

Preuve. Tout se vérifie en appliquant les définitions. $\square$

Attention, on montre que dans $\mathbb{C}$ (avec les opérations définies ci-dessus), il n'y a *pas de relation d'ordre* compatible avec la structure de corps. **On ne doit donc JAMAIS écrire des inégalités faisant intervenir des nombres complexes.**

Parmi les propriétés ci-dessus, revenons sur celle qui dit que *pour tout complexe non nul (a, b) , il existe un complexe unique (c, d) tel que*

$$(a, b) \times (c, d) = (1, 0).$$

On dit que tout complexe non nul possède un inverse pour la multiplication. L'inverse du complexe non nul $z = (a, b)$ est noté $\frac{1}{z}$ ou encore z^{-1} ; il est donné par

$$\left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

1.1.3 Introduction du complexe i et notations pratiques

Avec la convention d'écriture d'un blanc en lieu et place du signe $\times$ et l'identification d'un réel comme étant un complexe particulier, la relation

$$(r, 0) \times (a, b) = (a, b) \times (r, 0)$$

s'écrit

$$rz = zr, \quad 1z = z1 = z \text{ si } r = 1$$

avec $z = (a, b)$.

Définition 1.1.5 *On pose*

$$i = (0, 1).$$

Propriété(s) 1.1.6 *1) En tenant compte de l'identification de $\mathbb{R}$ comme sous-espace de $\mathbb{C}$, tout nombre complexe (a, b) s'écrit*

$$z = (a, b) = a + bi.$$

2) On a

$$i^2 = (0, 1) \times (0, 1) = -1.$$

Preuve. 1) On a en effet

$$z = (a, b) = (a, 0) + (0, b) = a(1, 0) + b(0, 1) = a + bi.$$

2) Il suffit d'appliquer la définition du produit entre complexes. $\square$

Grâce à l'introduction du complexe i et aux propriétés vérifiées par l'addition et la multiplication, le **calcul algébrique entre complexes** apparaît comme une **généralisation naturelle du calcul dans $\mathbb{R}$** en tenant compte de $i^2 = -1$. Ainsi par exemple

$$(a, b) \times (c, d) = (ac - bd, ad + bc)$$

par définition. Si l'on écrit

$$z = (a, b) = a + bi, \quad z' = (c, d) = c + di$$

et que l'on applique les propriétés des opérations (d'abord la distributivité), on a

$$(a + ib)(c + id) = ac + iad + ibc + i^2bd = ac - bd + i(bc + ad),$$

ce qui est bien le complexe de partie réelle $ac - bd$ et de partie imaginaire $bc + ad$ comme annoncé.

L'inverse du complexe non nul $z = a + ib$ s'écrit donc, en multipliant le numérateur et le dénominateur par $a - ib$,

$$\frac{1}{z} = \frac{1}{a + bi} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2}i.$$

Par exemple, on a

$$\begin{aligned}
 (3i + 1) i (1 - i) &= (3i + 1)(i - i^2) = (3i + 1)(i + 1) \\
 &= 3i^2 + 3i + i + 1 \\
 &= -3 + 3i + i + 1 \\
 &= -2 + 4i.
 \end{aligned}$$

De manière analogue, les parties réelle et imaginaire de $z = \frac{i}{2i-1}$ sont $2/5$ et $-1/5$, c'est-à-dire

$$z = \frac{i}{2i-1} = \frac{2}{5} - \frac{1}{5}i.$$

En effet

$$\frac{i}{2i-1} = i \frac{1}{2i-1} = i \frac{-1-2i}{2^2 + (-1)^2} = \frac{1}{5}(2-i).$$

1.1.4 Module et conjugué d'un complexe

Définition 1.1.7 Soit $z = (a, b) = a + bi$ un complexe ($a, b \in \mathbb{R}$).

Le complexe conjugué de z , noté $\bar{z}$, est le complexe

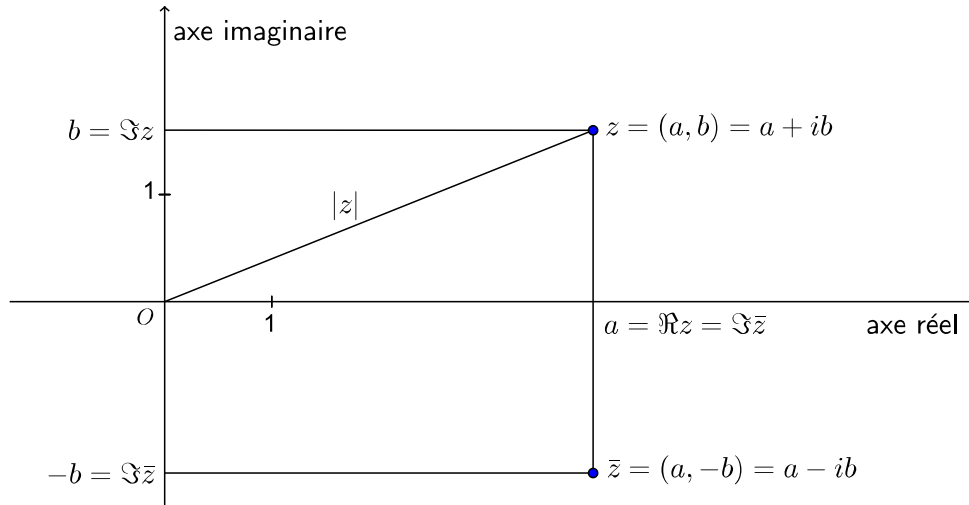
$$\bar{z} = (a, -b) = a - bi,$$

c'est-à-dire le complexe qui a la même partie réelle que z mais dont la partie imaginaire est l'opposé de celle de z . Graphiquement, il est le symétrique de z par rapport à l'axe réel.

Le module du complexe z , noté $|z|$, est le nombre réel positif

$$|z| = \sqrt{a^2 + b^2},$$

c'est-à-dire la longueur du vecteur d'origine O et dont l'extrémité est le point du plan de coordonnées (a, b) .



On vérifie directement les propriétés suivantes.

Propriété(s) 1.1.8 1. $|\bar{z}| = |z|$ pour tout complexe z

2. $|\operatorname{Re} z| \leq |z|$, $|\operatorname{Im} z| \leq |z|$ pour tout complexe z

3. $|z|^2 = z\bar{z}$ pour tout complexe z

4. $\frac{1}{z} = \frac{\bar{z}}{|z|^2}$ pour tout complexe non nul z

5. $|z_1 z_2| = |z_1| |z_2|$, $|z_1 + z_2| \leq |z_1| + |z_2|$ pour tous complexes z_1, z_2

6. $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$, $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$ pour tous complexes z_1, z_2 .

On voit aussi directement que

$$\frac{z + \bar{z}}{2} = \operatorname{Re} z, \quad \frac{z - \bar{z}}{2i} = \operatorname{Im} z$$

pour tout nombre complexe z .

1.1.5 Racines carrées d'un nombre complexe

Théorème 1.1.9 On a $z^2 = 0$ si et seulement si $z = 0$.

Si u est un complexe non nul, alors il possède deux racines carrées opposées. Cela signifie que l'équation en l'inconnue z

$$u = z^2$$

possède exactement deux solutions, qui sont des complexes opposés.

Preuve. Voir par exemple le syllabus de MATH2007 $\square$

Il ne faut pas oublier que la notation $\sqrt{}$ désigne « la fonction racine carrée », définie sur l'ensemble des réels positifs. Il est donc incorrect d'utiliser la notation $\sqrt{z}$ lorsque z est un complexe : cela n'a en effet aucun sens.

En guise d'exemples, cherchons les racines carrées des complexes

$$z_1 = -4, \quad z_2 = i, \quad z_3 = -5 + 12i.$$

On a $z_1 = 4i^2$; dès lors, les racines carrées de z_1 sont $2i$ et $-2i$.
Cherchons $x, y \in \mathbb{R}$ tels que $(x + iy)^2 = i$. On a

$$\begin{aligned} (x + iy)^2 = i &\Leftrightarrow \begin{cases} 0 = x^2 - y^2 \\ 1 = 2xy \end{cases} \\ &\Leftrightarrow \begin{cases} x = y \\ 1 = 2x^2 \end{cases} \quad \text{ou} \quad \begin{cases} x = -y \\ 1 = -2x^2 \end{cases} \\ &\Leftrightarrow \begin{cases} x = y \\ 1 = 2x^2 \end{cases}. \end{aligned}$$

Les deux racines carrées de i sont donc $\frac{\sqrt{2}}{2}(1 + i)$ et $-\frac{\sqrt{2}}{2}(1 + i)$.
Cherchons $x, y \in \mathbb{R}$ tels que $(x + iy)^2 = -5 + 12i$. On a

$$(x + iy)^2 = -5 + 12i \Leftrightarrow \begin{cases} -5 = x^2 - y^2 \\ 6 = xy. \end{cases}$$

En remplaçant y par $6/x$ dans la première équation, on trouve

$$x^4 + 5x^2 - 36 = 0.$$

Comme $\Delta = 25 + 4 \cdot 36 = 169 = 13^2$, cette équation a comme solutions $x = 2$ et $x = -2$. Dès lors, les deux racines carrées de $-5 + 12i$ sont $2 + 3i$ et $-(2 + 3i)$.

1.1.6 Trinôme du second degré

Propriété(s) 1.1.10 Le polynôme $z \mapsto P(z) = az^2 + bz + c$ où a, b, c sont des complexes et $a \neq 0$ admet toujours deux zéros (deux zéros distincts ou un zéro double).

Plus précisément, si on pose $\Delta = b^2 - 4ac$ et si z_0 est un complexe tel que $z_0^2 = \Delta$ alors les zéros de ce polynôme sont

$$\frac{-b + z_0}{2a}, \quad \frac{-b - z_0}{2a}.$$

Si a, b, c sont réels, et si $\Delta < 0$ alors les zéros sont des complexes conjugués.

Preuve. Voir par exemple le syllabus de MATH2007 $\square$

1.1.7 Complexes et trigonométrie

On démontre que quel que soit le complexe z , la suite

$$\sum_{m=0}^M \frac{z^m}{m!}, \quad M \in \mathbb{N}_0$$

converge. Sa limite est appelée l'exponentielle de z et est notée

$$\lim_{M \rightarrow +\infty} \sum_{m=0}^M \frac{z^m}{m!} = \sum_{m=0}^{+\infty} \frac{z^m}{m!} = \exp(z) = e^z = 1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots$$

Cela étant, on démontre que, quels que soient les complexes z, z' , on a

$$\exp(z + z') = \exp(z) \exp(z'),$$

ce qui explique d'ailleurs pourquoi on utilise la notation e^z qui rappelle les puissances (lesquelles ont justement la propriété $a^m a^n = a^{n+m}$ si a est un réel et n, m des naturels).

On a également

$$\overline{\exp(z)} = \exp(\bar{z}), \quad \exp(z) \overline{\exp(z)} = |\exp(z)|^2 = \exp(z + \bar{z}).$$

On définit alors rigoureusement les fonctions sinus et cosinus comme suit.

Définition 1.1.11 *Pour tout réel x , on définit*

$$\cos(x) = \Re(e^{ix}) \quad \text{et} \quad \sin(x) = \Im(e^{ix}).$$

On en déduit que

$$\cos(x) + i \sin(x) = \exp(ix) = e^{ix}, \quad \cos(x) - i \sin(x) = \overline{e^{ix}} = e^{-ix}$$

pour tout réel x . De plus, comme $\Re z = \frac{z + \bar{z}}{2}$ et $\Im z = \frac{z - \bar{z}}{2i}$ pour tout complexe z , on déduit aussi de la définition que

$$\cos(x) = \frac{e^{ix} + e^{-ix}}{2} \quad \text{et} \quad \sin(x) = \frac{e^{ix} - e^{-ix}}{2i}.$$

Des définitions précédentes, on déduit également ce qui suit.

Propriété(s) 1.1.12 *1) Pour tout $x \in \mathbb{R}$, on a*

$$|e^{ix}| = 1.$$

2) On a $\cos^2(x) + \sin^2(x) = 1$ pour tout réel x .

3) On a $(\cos(x) + i \sin(x))^m = \cos(mx) + i \sin(mx)$ pour tout naturel m et tout réel x .

4) Pour tout $z \in \mathbb{C}$ tel que $|z| = 1$, il existe $x \in [0, 2\pi[$ unique tel que

$$z = e^{ix}.$$

Preuve. 1) On a $z\bar{z} = |z|^2$ pour tout complexe z . Il s'ensuit que

$$|e^{ix}|^2 = e^{ix} \overline{e^{ix}} = e^{ix} e^{-ix} = e^{ix - ix} = 1,$$

d'où la conclusion car le module d'un complexe est un réel positif ou nul.

2) On a

$$1 = |e^{ix}|^2 = (\Re(e^{ix}))^2 + (\Im(e^{ix}))^2 = \cos^2(x) + \sin^2(x).$$

3) On a

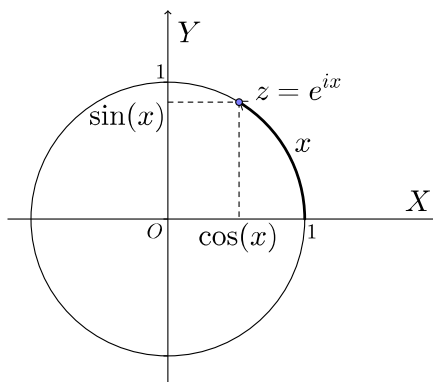
$$(\cos(x) + i \sin(x))^m = (e^{ix})^m = e^{imx} = \cos(mx) + i \sin(mx).$$

4) Résultat admis. $\square$

Ajoutons quelques autres propriétés.

1) Etant donné un complexe z de module 1, c'est-à-dire un point du plan situé sur le cercle centré à l'origine et de rayon 1, on sait qu'il existe un réel unique $x \in [0, 2\pi[$ tel que $z = e^{ix}$. On montre aussi que la longueur de l'arc de cercle joignant le complexe 1 au complexe z vaut x .

On obtient donc la représentation suivante.



2) La forme trigonométrique d'un nombre complexe consiste simplement à écrire celui-ci en se servant des coordonnées polaires du point du plan qu'il détermine.

Etant donné $z \in \mathbb{C}$, $z \neq 0$, on sait qu'il existe un réel $x \in [0, 2\pi[$, unique, tel que

$$\frac{z}{|z|} = e^{ix}.$$

En posant

$$r = |z|,$$

on a

$$z = re^{ix};$$

c'est ce que l'on appelle la forme trigonométrique du complexe z . Les réels r et x constituent également les coordonnées polaires du point P d'abscisse $\Re z$ et d'ordonnée $\Im z$.

3) Interprétons à présent ⁵ la *multiplication de deux complexes*. Soient z, z' deux complexes non nuls. On peut écrire

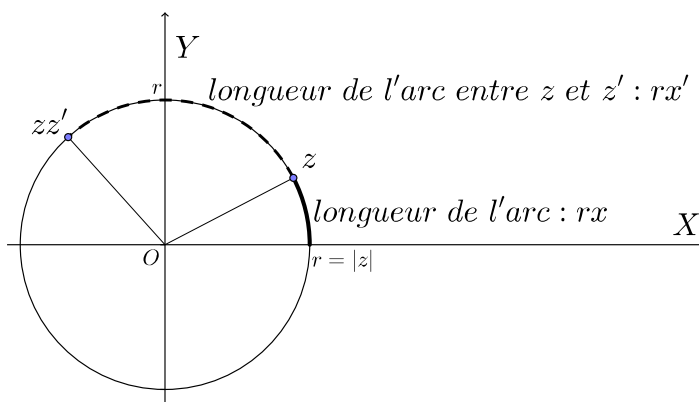
$$z = re^{ix}, \quad z' = r'e^{ix'}$$

donc

$$zz' = rr'e^{i(x+x')}.$$

La multiplication de z par z' consiste donc en une multiplication par le réel r' (qui s'interprète comme la multiplication d'un vecteur par un réel) et en une rotation d'un angle x' .

Illustration lorsque $|z'| = r' = 1$.



⁵. Cf aussi MATH2007 où l'interprétation est faite grâce à la trigonométrie, sans faire appel aux propriétés de la fonction exponentielle

1.2 (Rappels concernant les) fonctions de plusieurs variables

Nous renvoyons au syllabus du cours MATH1009 et dans une certaine mesure, à celui du cours MATH0009 (la matière de ce syllabus concernant les fonctions de plusieurs variables n'est pas toujours complètement couverte lors du cours enseigné, par manque de temps).

Dans le plan, on définit les coniques (ellipse, hyperbole, parabole), qui sont des courbes très particulières. De même dans l'espace, on introduit la notion de quadriques (ellipsoïde, hyperboloïde à une nappe et à deux nappes, paraboloides elliptique, paraboloides hyperbolique, cylindre elliptique, cylindre hyperbolique, cylindre parabolique), qui sont des surfaces très particulières.

Dans le cadre des fonctions de plusieurs variables, on définit la dérivabilité de la même manière qu'à une variable, en « bloquant » successivement toutes les variables sauf une. On introduit ainsi la notion de dérivée partielle (voir les syllabi MATH1009 et MATH0009). Par ailleurs, la notion de « fonction de fonction » devient la notion de « fonction composée », en partant du même principe qu'à une variable. Rappelons ici le résultat concernant la dérivation des fonctions composées.

Proposition 1.2.1 a) Soient

- f une fonction de deux variables réelles continûment dérivable sur un ouvert U de $\mathbb{R}^2$,
- f_1, f_2 deux fonctions d'une variable réelle dérivables sur un ouvert J de $\mathbb{R}$.

Alors la fonction définie par

$$F(t) = f(f_1(t), f_2(t))$$

est dérivable sur $I = \{t \in J : (f_1(t), f_2(t)) \in U\}$ et on a

$$DF(t) = D_1f(X, Y) Df_1(t) + D_2f(X, Y) Df_2(t)$$

avec $t \in I$ et $X = f_1(t)$, $Y = f_2(t)$.

b) Soient

- f une fonction de deux variables réelles continûment dérivable sur un ouvert U de $\mathbb{R}^2$,
- f_1, f_2 deux fonctions de deux variables réelles dérivables sur un ouvert V de $\mathbb{R}^2$.

Alors la fonction définie par

$$F(x, y) = f(f_1(x, y), f_2(x, y))$$

est dérivable sur $\Omega = \{(x, y) \in V : (f_1(x, y), f_2(x, y)) \in U\}$ et on a

$$D_1F(x, y) = D_1f(X, Y) D_1f_1(x, y) + D_2f(X, Y) D_1f_2(x, y)$$

$$D_2F(x, y) = D_1f(X, Y) D_2f_1(x, y) + D_2f(X, Y) D_2f_2(x, y)$$

avec $(x, y) \in \Omega$ et $X = f_1(x, y)$, $Y = f_2(x, y)$.

c) Soient

- f une fonction d'une variable réelle dérivable sur un ouvert I de $\mathbb{R}$,
- g une fonction de deux variables réelles dérivable sur l'ouvert V de $\mathbb{R}^2$.

Alors la fonction définie par

$$F(x, y) = f(g(x, y))$$

est dérivable sur $\Omega = \{(x, y) \in V : g(x, y) \in I\}$ et on a

$$D_1F(x, y) = Df(X) D_1g(x, y)$$

$$D_2F(x, y) = Df(X) D_2g(x, y)$$

avec $(x, y) \in \Omega$ et $X = g(x, y)$.

d) On a bien sûr des résultats analogues pour des fonctions du type $f(f_1, \dots, f_n)$ avec f fonction de n variables et $f_1, \dots, f_n$ fonctions de p variables, à valeurs réelles.

Remarque. Pour avoir la dérivabilité de la fonction composée ET l'expression explicite des dérivées présentées dans le résultat précédent, il importe de bien vérifier les hypothèses.

1.3 Rappels concernant le calcul intégral

Voir les notes du cours de MATH2007, MATH0009 et MATH1009.

Rappelons que l'intégrale d'une fonction d'une variable réelle à valeurs réelles positives permet de définir la notion d'« aire sous une courbe » et que l'intégrale d'une fonction de deux variables réelles à valeurs réelles positives permet de définir la notion de « volume sous une surface ».

Rappelons tout de même quelques résultats concernant l'intégration à plusieurs variables ; nous ne rappelons que ce qui nous sera utile par la suite.

Définition 1.3.1 Soit f une fonction continue sur le rectangle $R =]a, b[\times]c, d[$. On dit que f est intégrable sur R lorsque

- $\forall x \in]a, b[$, la fonction $y \mapsto |f(x, y)|$ est intégrable sur $]c, d[$
et
- la fonction $x \mapsto \int_c^d |f(x, y)| dy$ est intégrable sur $]a, b[$
OU lorsque
- $\forall y \in]c, d[$, la fonction $x \mapsto |f(x, y)|$ est intégrable sur $]a, b[$
et
- la fonction $y \mapsto \int_a^b |f(x, y)| dx$ est intégrable sur $]c, d[$.

Remarquons que puisque $\|f\| = |f|$, la fonction f est intégrable sur R si et seulement si la fonction $|f|$ l'est.

Cela étant, si f est intégrable sur $R =]a, b[\times]c, d[$, on démontre que

- $\forall x \in]a, b[$, la fonction $y \mapsto f(x, y)$ est intégrable sur $]c, d[$
et
- la fonction $x \mapsto \int_c^d f(x, y) dy$ est intégrable sur $]a, b[$
ET
- $\forall y \in]c, d[$, la fonction $x \mapsto f(x, y)$ est intégrable sur $]a, b[$
et
- la fonction $y \mapsto \int_a^b f(x, y) dx$ est intégrable sur $]c, d[$
ET (permutation de l'ordre d'intégration sans changer la valeur de l'intégrale)

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

où la notation

$$\int \int_R f(x, y) dx dy$$

désigne l'intégrale de f (fonction de deux variables) sur l'ensemble R .

1.4 Intégrales paramétriques

Les intégrales paramétriques (et leur dérivation) représentent un outil très important de l'analyse. Elles apparaissent notamment dans le cadre de l'analyse de Fourier et du produit de convolution de fonctions.

Dans le présent cours, nous ne considérerons ces intégrales que dans le cas de l'intégration à une variable et dans le cas d'un seul paramètre réel.

Comme leur nom l'indique déjà, les « intégrales paramétriques » sont des fonctions qui sont définies par des intégrales, comme suit

$$\lambda \mapsto \int_A f(x, \lambda) dx.$$

Dès le départ, il convient de fixer un cadre qui donne sens à cette définition :

- A est un sous-ensemble de $\mathbb{R}$ (le plus souvent un intervalle) ; le paramètre λ varie dans un sous-ensemble de $\mathbb{R}$ également, noté par exemple Λ , et qui sera considéré ouvert lorsqu'il sera question de dérivation

- f est une fonction définie sur le produit cartésien $A \times \Lambda$ et, pour tout $\lambda \in \Lambda$, la fonction $x \mapsto f(x, \lambda)$ est intégrable sur A .

Enonçons à présent le « Théorème de dérivation des intégrales paramétriques » dans le cas d'une seule dérivation.

Théorème 1.4.1 (Dérivation des intégrales paramétriques) *On reprend les notations et les hypothèses naturelles ci-dessus. En outre, on suppose que*

- *quel que soit $x \in A$, la fonction $\lambda \mapsto f(x, \lambda)$ appartient à $C_1(\Lambda)$*
- *quel que soit $\lambda \in \Lambda$, la fonction $x \mapsto f(x, \lambda)$ est intégrable sur A*
- *(*) quel que soit l'ensemble borné fermé K inclus dans l'ouvert Λ , il existe une fonction g_K intégrable sur A telle que*

$$|D_\lambda f(x, \lambda)| \leq g_K(x), \quad \forall x \in A, \forall \lambda \in K$$

Sous ces conditions, la fonction $\lambda \mapsto \int_A f(x, \lambda) dx$ appartient à $C_1(\Lambda)$ et « on peut dériver sous le signe intégral », ce qui signifie que

$$D_\lambda \int_A f(x, \lambda) dx = \int_A D_\lambda f(x, \lambda) dx, \quad \forall \lambda \in \Lambda.$$

Remarque 1.4.2 1) *Lorsque $f \in C_1(A' \times \Lambda)$ avec A' ouvert contenant A et lorsque A est borné et fermé, l'hypothèse (*) est automatiquement satisfaite.*

2) *Pour n dérivations (n naturel strictement plus grand que 1), on demande que $\lambda \mapsto f(x, \lambda)$ appartienne à $C_n(\Lambda)$, que les dérivées (par rapport à λ) jusqu'à l'ordre $n - 1$ soient intégrables sur A et que la majoration qui apparaît dans (*) fasse intervenir $D_\lambda^n f(x, \lambda)$. On obtient alors que la fonction $\lambda \mapsto \int_A f(x, \lambda) dx$ appartient à $C_n(\Lambda)$ et que la dérivation s'effectue encore de la même manière (jusqu'à l'ordre n).*

Exemple(s) 1.4.3 *Un exemple très utile pour la suite...*

Pour tout réel $a > 0$, la fonction $x \mapsto e^{-ax^2}$ est intégrable sur $\mathbb{R}$ et on a

$$\int_0^{+\infty} e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}}$$

Cette égalité est appelée formule de Poisson et elle est fondamentale en statistiques notamment (intégration de gaussiennes, loi normale, ...) Pour la preuve de ce résultat, voir les notes du cours Math0009.

On a en fait la généralisation suivante : *pour tout $a > 0$ et tout $b \in \mathbb{R}$, la fonction $x \mapsto \cos(bx) e^{-ax^2}$ est intégrable sur $\mathbb{R}$ et on a*

$$\boxed{\int_0^{+\infty} \cos(bx) e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}} e^{-b^2/(4a)}}$$

Puisque les valeurs de la fonction cosinus sont dans $[-1, 1]$ et que la fonction $x \mapsto e^{-ax^2}$ est intégrable sur $\mathbb{R}$ (quand $a > 0$), il est clair que $x \mapsto \cos(bx) e^{-ax^2}$ est intégrable sur $\mathbb{R}$ quels que soient $a > 0$ et $b \in \mathbb{R}$. Cela étant cette égalité peut être démontrée en ayant recours au théorème de dérivation des intégrales paramétriques.

On fixe $a > 0$ et on considère la fonction $(x, b) \mapsto \cos(bx) e^{-ax^2}$ avec $A = [0, +\infty[$, $\Lambda = \mathbb{R}$ et $(x, b) \in A \times \Lambda$. **Notes à compléter (voir cours enseigné).**

1.5 Transformation de Fourier des fonctions intégrables sur $\mathbb{R}$

1.5.1 Définition et interprétation

Donnons tout d'abord l'introduction suivante, récupérée dans wikipedia (extraits)

[...] La transformation de Fourier associe à une fonction intégrable sur $\mathbb{R}$ et à valeurs réelles ou complexes, une autre fonction sur $\mathbb{R}$ appelée transformée de Fourier dont la variable indépendante peut s'interpréter en physique comme la fréquence ou la pulsation.

[...] La transformée de Fourier représente une fonction par la densité spectrale dont elle provient, en tant que moyenne de fonctions trigonométriques de toutes fréquences.

Lorsqu'une fonction représente un phénomène physique, comme l'état du champ électromagnétique ou du champ acoustique en un point, on l'appelle signal et sa transformée de Fourier s'appelle son spectre.

Pour illustrer cette introduction, voir la fin de cette sous-section.

Définition 1.5.1 Soit f une fonction intégrable sur $\mathbb{R}$. La transformée de Fourier « négative » de f est la fonction

$$\mathcal{F}^- f : y \in \mathbb{R} \mapsto \int_{\mathbb{R}} e^{-ixy} f(x) dx$$

et la transformée de Fourier « positive » de f est la fonction

$$\mathcal{F}^+ f : y \in \mathbb{R} \mapsto \int_{\mathbb{R}} e^{ixy} f(x) dx.$$

Pour la valeur en y de ces fonctions, on utilise les notations

$$\mathcal{F}_y^- f, \quad \mathcal{F}_y^+ f$$

ou encore

$$(\mathcal{F}^- f)(y), \quad (\mathcal{F}^+ f)(y).$$

Notons que cette définition a bien un sens puisque pour tout y , la fonction $x \mapsto e^{ixy} f(x)$ (resp. $x \mapsto e^{-ixy} f(x)$) est intégrable sur $\mathbb{R}$ puisqu'en module c'est $|f|$.

Remarque 1.5.2 Pour toute fonction f intégrable sur $\mathbb{R}$, on a

$$\mathcal{F}_y^- f = \mathcal{F}_{-y}^+ f \quad \forall y \in \mathbb{R}.$$

Preuve. Il s'agit d'une simple écriture différente pour l'expression

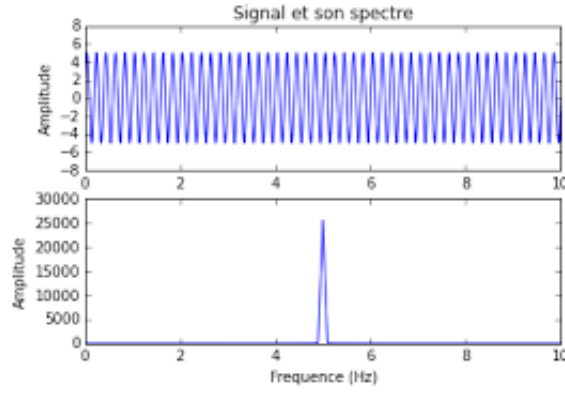
$$\int_{\mathbb{R}} e^{-ixy} f(x) dx = \int_{\mathbb{R}} e^{ix(-y)} f(x) dx.$$

□

Pour illustrer un peu l'introduction qui met l'accent sur l'aspect fréquentiel de la transformation de Fourier, notons que, pour tout $r > 0$ (indispensable de recouper le cosinus, lequel n'est pas intégrable sur $\mathbb{R}$)

$$\mathcal{F}_y^\pm (\cos \chi_{[-r,r]}) = \begin{cases} \frac{\sin(r(y+1))}{y+1} + \frac{\sin(r(y-1))}{y-1} & \text{si } y \neq 1, y \neq -1 \\ \frac{\sin(2r)}{2} + r & \text{si } y = 1 \text{ et si } y = -1 \end{cases}$$

La plus grande valeur de cette transformée de $x \mapsto \cos(\mathbf{1}x) = \cos(-\mathbf{1}x)$ est donc en -1 et 1 , c'est-à-dire qu'il y a un « pic » en ces points. En théorie des distributions où l'on peut vraiment prendre la transformée de Fourier de la distribution définie par le cosinus, on trouve effectivement les distributions de Dirac.



1.5.2 Exemples

Exemple(s) 1.5.3 Transformation de Fourier de $\chi_{[a,b]}$ ($a, b \in \mathbb{R}$) et de $e^{-a|\cdot|}$ ($a > 0$) :

$$\begin{aligned}\mathcal{F}_y^\pm \chi_{[a,b]} &= \begin{cases} \frac{e^{\pm iby} - e^{\pm iay}}{\pm iy} & \text{si } y \neq 0 \\ b - a & \text{si } y = 0 \end{cases} \\ \mathcal{F}_y^\pm e^{-a|\cdot|} &= \frac{2a}{y^2 + a^2}.\end{aligned}$$

En particulier, pour tout $r > 0$, on a

$$\mathcal{F}_y^\pm \chi_{[-r,r]} = \begin{cases} \frac{2 \sin(ry)}{y} & \text{si } y \neq 0 \\ 2r & \text{si } y = 0. \end{cases}$$

Preuve. Ce sont des calculs immédiats d'intégrales simples. $\square$

Le premier exemple montre que la transformée de Fourier d'une fonction intégrable n'est pas nécessairement intégrable. Lorsqu'on aura besoin de cette propriété, on devra donc donner des conditions suffisantes pour avoir cette intégrabilité. Dans la suite, nous donnons les deux résultats les plus courants, bien utiles plus loin dans le cours.

Par ailleurs, ces deux exemples (ainsi que celui qui suit) conduisent à des transformées qui sont continues sur $\mathbb{R}$ et ont une limite nulle à l'infini. Ces propriétés sont en fait vraies pour toute transformée de Fourier de fonction intégrable, comme démontré dans la suite.

Voici alors un exemple qui se révélera fondamental dans la preuve du théorème de Fourier.

Exemple(s) 1.5.4 (Le cas des fonctions gaussiennes) Pour tout $a > 0$ on définit la gaussienne g_a par

$$g_a(x) = e^{-ax^2}, \quad x \in \mathbb{R}.$$

Cette fonction est intégrable et on a

$$\mathcal{F}_x^\pm g_a = \left(\frac{\pi}{a}\right)^{1/2} g_{1/(4a)}(x) \quad \text{pour tout } x$$

On en déduit que

$$\mathcal{F}_x^\mp \mathcal{F}_x^\pm g_a = 2\pi g_a(x) \quad \text{pour tout } x$$

Preuve. L'intégrabilité des gaussiennes est claire. Cela étant, on a directement (on se réfère à une intégrale « remarquable » obtenue par le théorème de dérivation des intégrales paramétriques), quel que soit $y \in \mathbb{R}$,

$$\mathcal{F}_y^\pm g_a = \int_{\mathbb{R}} e^{\pm ixy} e^{-ax^2} dx$$

$$\begin{aligned}
&= 2 \int_0^{+\infty} \cos(xy) e^{-ax^2} dx \\
&= \sqrt{\frac{\pi}{a}} e^{-y^2/(4a)} \\
&= \sqrt{\frac{\pi}{a}} g_{(4a)^{-1}}(y).
\end{aligned}$$

On vient donc de voir que la transformée de Fourier d'une gaussienne est un multiple d'une gaussienne. Reprenons alors la transformée de Fourier de cette transformée de Fourier, en utilisant l'expression générale qui vient d'être trouvée. Quel que soit $x \in \mathbb{R}$, on a successivement ⁶

$$\mathcal{F}_x^\mp \mathcal{F}^\pm g_a = \sqrt{\frac{\pi}{a}} \mathcal{F}_x^\mp g_{(4a)^{-1}} = \sqrt{\frac{\pi}{a}} \sqrt{\frac{\pi}{(4a)^{-1}}} g_{(4(4a)^{-1})^{-1}}(x) = 2\pi g_a(x)$$

et on conclut. $\square$

1.5.3 Premières propriétés

Propriété(s) 1.5.5 (1) La transformation de Fourier est un opérateur linéaire.

(2) La transformée de Fourier d'une fonction intégrable f est une fonction bornée sur $\mathbb{R}$ car on a

$$|\mathcal{F}_y f| \leq \int_{\mathbb{R}} |f(x)| dx \quad \forall y \in \mathbb{R}$$

(3) La transformée de Fourier d'une fonction intégrable f est une fonction continue⁷ sur $\mathbb{R}$.

(4) (Riemann-Lebesgue) La transformée de Fourier d'une fonction intégrable f tend vers 0 à l'infini.

Preuve. (1) C'est immédiat vu la linéarité de l'intégration.

(2) C'est immédiat car le module d'une exponentielle imaginaire pur est égal à 1 :

$$|\mathcal{F}_y f| = \left| \int_{\mathbb{R}} e^{\pm ixy} f(x) dx \right| \leq \int_{\mathbb{R}} |e^{\pm ixy}| |f(x)| dx = \int_{\mathbb{R}} |f(x)| dx \quad \forall y \in \mathbb{R}$$

(3) Admis

(4) Admis $\square$

1.5.4 Dérivation et transformation de Fourier

Propriété(s) 1.5.6 (1) Si $f \in C_L(\mathbb{R})$ et si $D^\alpha f$ est intégrable sur $\mathbb{R}$ quel que soit α tel que $\alpha \leq L$ alors

$$\mathcal{F}_y^\pm D^\alpha f = (\mp iy)^\alpha \mathcal{F}_y^\pm f \quad \forall y \in \mathbb{R}$$

(2) Si les fonctions $x \mapsto x^\alpha f(x)$ ($\alpha \leq L$) sont intégrables, alors $\mathcal{F}^\pm f \in C_L(\mathbb{R})$ et, quel que soit α tel que $|\alpha| \leq L$, on a

$$D^\alpha \mathcal{F}_y^\pm f = (\pm i)^\alpha \int_{\mathbb{R}} x^\alpha e^{\pm ixy} f(x) dx \quad \forall y \in \mathbb{R}.$$

Preuve. (1) Pour $L = 1$, cela résulte directement d'une intégration par parties. De fait, les deux limites

$$\lim_{x \rightarrow +\infty} (e^{\pm ixy} f(x)), \quad \lim_{x \rightarrow -\infty} (e^{\pm ixy} f(x))$$

sont nulles car elles existent et les fonctions $x \mapsto e^{\pm ixy} f(x)$ sont intégrables à l'infini. Le cas général s'effectue de même, en répétant la manoeuvre précédente.

(2) L'expression

$$\mathcal{F}_y^\pm f = \int_{\mathbb{R}} e^{\pm ixy} f(x) dx$$

est une intégrale paramétrique. Dans le cas présent, vu les hypothèses données, celles du théorème de dérivation sont clairement satisfaites comme on le vérifie tout de suite et ainsi on obtient la dérivabilité et l'expression des dérivées en permutant dérivée et intégrale. $\square$.

6. Ici, les deux transformées de Fourier $\pm$ sont les mêmes car une gaussienne est paire ; cependant, pour le théorème de Fourier en toute généralité, cela ne sera pas toujours le cas et il importera de respecter l'alternance des « signes »

7. et même uniformément continue

1.5.5 Intégration et transformation de Fourier

On a vu précédemment que la transformée de Fourier d'une fonction intégrable n'est pas nécessairement intégrable. Donnons ici une condition suffisante pour qu'elle le soit.

Propriété(s) 1.5.7 Si $f \in C_2(\mathbb{R})$ est tel que f, Df et D^2f soient intégrables sur $\mathbb{R}$, alors $\mathcal{F}^\pm f$ est intégrable.

Preuve. Notons tout d'abord que l'on a la continuité de la fonction dont on veut montrer l'intégrabilité. Cela étant, on obtient l'intégrabilité tout de suite car

$$|x|^2 |\mathcal{F}_x^\pm f| = |x^2 \mathcal{F}_x^\pm f| = |\mathcal{F}_x^\pm D^2 f| \leq C \quad \forall x$$

donc, pour tout $x \neq 0$, on a

$$|\mathcal{F}_x^\pm f| \leq \frac{C}{x^2}$$

et on conclut étant donné l'intégrabilité de la fonction $x \mapsto 1/x^2$ à l'infini. $\square$

1.5.6 Théorème de transfert et de Fourier

Propriété(s) 1.5.8 (Transfert) Si f et g sont intégrables alors

$$\int_{\mathbb{R}} \mathcal{F}_x^\pm f g(x) dx = \int_{\mathbb{R}} f(x) \mathcal{F}_x^\pm g dx$$

Preuve. Il est clair que les deux membres de l'égalité ont un sens car le produit d'une fonction intégrable par une fonction bornée est intégrable.

Cela étant, une simple permutation de l'ordre d'intégration permet de conclure. De fait, la fonction $(x, y) \mapsto |f(y)| |g(x)|$ est intégrable sur $\mathbb{R}^2$ et on a

$$\begin{aligned} \int_{\mathbb{R}} \mathcal{F}_x^\pm f g(x) dx &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\pm ixy} f(y) dy \right) g(x) dx \\ &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} e^{\pm ixy} g(x) dx \right) f(y) dy \\ &= \int_{\mathbb{R}} \mathcal{F}_y^\pm g f(y) dy \end{aligned}$$

$\square$

Théorème 1.5.9 (Fourier) Si f est intégrable et de transformée de Fourier intégrable, on a

$$\mathcal{F}^\mp \mathcal{F}^\pm f = 2\pi f$$

en tout point où f est continu.

Preuve. Admis $\square$

NOTONS ICI QU'IL EST INDISPENSABLE D'ÊTRE CAPABLE D'EXPRIMER CETTE ÉGALITÉ EN UN RÉEL QUELCONQUE EN EXPLICITANT LES INTÉGRALES QUI « SE CACHENT » DANS LES TRANSFORMÉES DE FOURIER.

1.6 Produit de composition

Nous présentons ici l'essentiel, avec pour but la présentation d'un autre outil de l'analyse du signal que la transformation de Fourier ET un lien fondamental entre ces deux outils.

Définition 1.6.1 Soient f et g deux fonctions définies sur $\mathbb{R}$; si $y \in \mathbb{R}$ et si la fonction $x \mapsto f(x) g(y-x)$ est intégrable, son intégrale est notée

$$(f * g)(y) = \int_{\mathbb{R}} f(x) g(y-x) dx.$$

Lorsque la fonction $x \mapsto f(x) g(y-x)$ est intégrable pour tout y , le produit de convolution (ou de composition) de f et g est la fonction

$$y \mapsto (f * g)(y).$$

Propriété(s) 1.6.2 *Le produit de composition est commutatif et linéaire sur chacun des facteurs.*

Preuve. C'est immédiat par changement de variable (linéaire) : on a

$$(f * g)(y) = \int_{\mathbb{R}} f(x) g(y-x) dx = \int_{\mathbb{R}} f(y-t) g(t) dt = (g * f)(y).$$

La linéarité est immédiate étant donné la linéarité de l'intégrale. $\square$

Interprétation 1.6.3 Voir cours et aussi le descriptif par exemple à l'adresse

https://fr.wikipedia.org/wiki/Produit_de_convolution

Propriété(s) 1.6.4 (Produit de composition et transformée de Fourier) *Si f et g sont intégrables alors*

$$\mathcal{F}^{\pm}(f * g) = \mathcal{F}^{\pm} f \mathcal{F}^{\pm} g$$

c'est-à-dire que la transformée de Fourier du produit de composition de deux fonctions intégrables est égale au produit des transformées de Fourier des deux fonctions.

Preuve. Montrons que le produit de composition existe et définit une fonction intégrable. Comme rien n'indique que le produit de deux fonctions intégrables soit intégrable, il n'est pas évident a priori que la fonction $x \mapsto f(x) g(y-x)$ soit intégrable. Mais si on s'y prend autrement, l'intégrabilité va apparaître tout de suite. De fait, pour x fixé, la fonction

$$y \mapsto |f(x)| |g(y-x)|$$

est intégrable puisque g l'est ; ensuite, la fonction

$$x \mapsto \int_{\mathbb{R}} |f(x)| |g(y-x)| dy = |f(x)| \int_{\mathbb{R}} |g(y-x)| dy = |f(x)| \int_{\mathbb{R}} |g(t)| dt$$

est intégrable puisque f l'est. Il s'ensuit que la fonction de deux variables

$$(x, y) \mapsto f(x) g(y-x)$$

est intégrable et on conclut (cf le chapitre sur l'intégration à plusieurs variables du cours MATH1009).

On a alors directement, quel que soit $x \in \mathbb{R}$,

$$\begin{aligned} \mathcal{F}_x^{\pm}(f * g) &= \int_{\mathbb{R}} e^{\pm ixy} (f * g)(y) dy \\ &= \int_{\mathbb{R}} e^{\pm ixy} \left(\int_{\mathbb{R}} f(t) g(y-t) dt \right) dy \\ &= \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} e^{\pm ixy} g(y-t) dy \right) dt \\ &= \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} e^{\pm ix(t+u)} g(u) du \right) dt \\ &= \int_{\mathbb{R}} e^{\pm ixt} f(t) \left(\int_{\mathbb{R}} e^{\pm ixu} g(u) du \right) dt \\ &= \left(\int_{\mathbb{R}} e^{\pm ixu} g(u) du \right) \times \left(\int_{\mathbb{R}} e^{\pm ixt} f(t) dt \right) \\ &= \mathcal{F}_x^{\pm} g \times \mathcal{F}_x^{\pm} f. \end{aligned}$$

$\square$

Chapitre 2

Seconde partie

2.1 Le développement limité de Taylor

Commençons par présenter les développements limités de Taylor à plusieurs variables. Pour cela *il est essentiel de bien se rappeler l'énoncé dans le cas d'une variable.*

Théorème 2.1.1 (Développement limité de Taylor, ouvert quelconque) .

Cas de l'ordre 1. Soit Ω un ouvert de $\mathbb{R}^2$ et soit $f \in C_1(\Omega)$, f à valeurs réels. Pour tous points (x_0, y_0) , (x, y) de cet ouvert qui sont tels que les points du segment qui les relie appartiennent encore à l'ouvert, il existe $t_0 \in [0, 1]$ tel que

$$f(x, y) = f(x_0, y_0) + (x - x_0)D_1f(u_0, v_0) + (y - y_0)D_2f(u_0, v_0)$$

où on a posé $u_0 = x_0 + t_0(x - x_0)$ et $v_0 = y_0 + t_0(y - y_0)$.

Cas de l'ordre 2. Soit Ω un ouvert de $\mathbb{R}^2$ et soit $f \in C_2(\Omega)$, f à valeurs réels. Pour tous points (x_0, y_0) , (x, y) de cet ouvert qui sont tels que les points du segment qui les relie appartiennent encore à l'ouvert, il existe $t_0 \in]0, 1[$ tel que

$$\begin{aligned} f(x, y) = & f(x_0, y_0) + (x - x_0)D_1f(x_0, y_0) + (y - y_0)D_2f(x_0, y_0) \\ & + \frac{(x - x_0)^2}{2}D_1^2f(u_0, v_0) + (x - x_0)(y - y_0)D_1D_2f(u_0, v_0) + \frac{(y - y_0)^2}{2}D_2^2f(u_0, v_0) \end{aligned}$$

où on a posé $u_0 = x_0 + t_0(x - x_0)$ et $v_0 = y_0 + t_0(y - y_0)$.

Preuve. Effectuons la démonstration dans le cas de l'ordre 1. Pour l'ordre 2, la procédure est la même mais fait intervenir plus de dérivées partielles et est donc plus lourde.

La droite passant par les points de coordonnées (x_0, y_0) et (x, y) a pour équation paramétrique

$$\begin{cases} X = x_0 + t(x - x_0) \\ Y = y_0 + t(y - y_0) \end{cases} \quad t \in \mathbb{R};$$

dès lors les points du segment joignant (x_0, y_0) et (x, y) ont pour coordonnées

$$(x_0 + t(x - x_0), y_0 + t(y - y_0)), \quad t \in [0, 1].$$

Comme les points de coordonnées (x_0, y_0) et (x, y) appartiennent à un ouvert (Ω) , il existe une boule centrée en chacun d'eux qui reste dans l'ouvert. (Introduire un dessin) Par conséquent, il existe $r > 0$ tel que les points de coordonnées

$$(x_0 + t(x - x_0), y_0 + t(y - y_0)), \quad t \in]-r, 1 + r[$$

appartiennent tous à l'ouvert Ω . Comme la fonction f appartient à $C_1(\Omega)$ et que les fonctions $t \mapsto x_0 + t(x - x_0)$ et $t \mapsto y_0 + t(y - y_0)$ appartiennent à $C_\infty(]-r, 1 + r[)$, la fonction de fonction F définie par

$$F(t) = f(x_0 + t(x - x_0), y_0 + t(y - y_0))$$

appartient à $C_1[] - r, 1 + r[]$ et pour tout $t \in] - r, 1 + r[$, on a

$$DF(t) = (D_1f)(X, Y)(x - x_0) + (D_2f)(X, Y)(y - y_0)$$

avec

$$X = x_0 + t(x - x_0), \quad Y = y_0 + t(y - y_0).$$

Cela étant, comme $0, 1 \in] - r, 1 + r[$, le théorème des accroissements finis dans l'intervalle $] - r, 1 + r[$ donne l'existence de $t_0 \in]0, 1[$ tel que

$$F(1) = F(0) + DF(t_0).$$

Comme on a $F(1) = f(x, y)$, $F(0) = f(x_0, y_0)$ on obtient

$$f(x, y) = f(x_0, y_0) + (x - x_0)(D_1f)(u_0, v_0) + (y - y_0)(D_2f)(u_0, v_0)$$

avec $u_0 = x_0 + t_0(x - x_0)$ et $v_0 = y_0 + t_0(y - y_0)$. $\square$

2.2 Extrema : définitions et première propriété

Commençons par donner quelques définitions et une propriété importante, tout cela étant une simple généralisation du cas d'une variable. Dans les présentes notes, nous ne présenterons que le cas de deux variables.

Définition 2.2.1 Soit f une fonction de deux variables réelles définie dans $A \subset \mathbb{R}^2$, à valeurs réelles.

- La fonction f admet un minimum (resp. maximum) local en $(x_0, y_0) \in A$ s'il existe $\varepsilon > 0$ tel que

$$f(x, y) \geq f(x_0, y_0) \text{ (resp. } f(x, y) \leq f(x_0, y_0)) \quad \forall (x, y) \in A : |x - x_0| \leq \varepsilon \text{ et } |y - y_0| \leq \varepsilon.$$

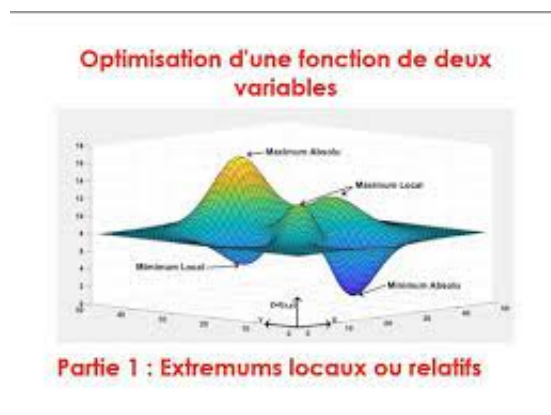
- La fonction f admet une minimum (resp. maximum) global en $(x_0, y_0) \in A$ si

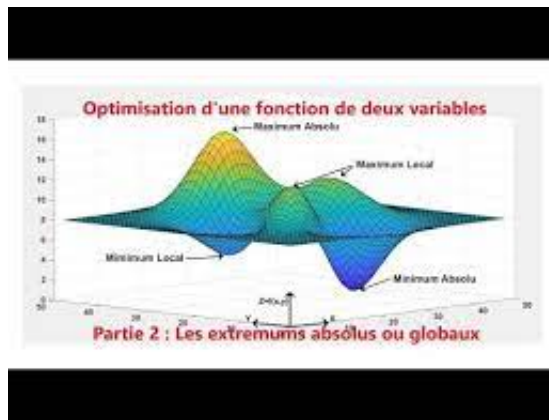
$$f(x, y) \geq f(x_0, y_0) \text{ (resp. } f(x, y) \leq f(x_0, y_0)) \quad \forall (x, y) \in A.$$

- Lorsque les inégalités entre les valeurs de la fonction sont strictes sauf en (x_0, y_0) bien sûr, on parle de minimum et de maximum strict (local ou global).

- Si A est ouvert et si f est dérivable dans A , on appelle point stationnaire de f dans A tout point de A qui annule les deux dérivées partielles de f .

Comme dans le cas d'une variable, on appelle *extremum* un minimum ou un maximum.





Et voici la propriété qui montre que dans la recherche des extrema d'une fonction dérivable, on ne doit s'intéresser qu'aux points stationnaires. La preuve est tout à fait analogue au cas d'une variable.

Propriété(s) 2.2.2 *Tout extremum local d'une fonction dérivable dans un ouvert est un point stationnaire pour cette fonction.*

Cela étant si la recherche d'extrema d'une fonction d'une variable qui est deux fois dérivable est directe par étude des deux premières dérivées, il n'en est plus de même dans le cas d'une fonction de deux variables. De fait, le développement limité de Taylor d'une fonction deux fois continûment dérivable est plus complexe bien que toujours très naturel : il fait intervenir **toutes** les dérivées partielles secondes, donc ici la somme de trois termes et non plus un seul terme comme dans le cas d'une variable. L'étude du signe de cette somme de trois termes est donc plus ardue. Cette étude fait en fait appel à du calcul matriciel, et plus précisément aux propriétés des matrices symétriques réelles.

2.3 Compléments de calcul matriciel

Soit une matrice réelle symétrique

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

(puisque la matrice est réelle, on a donc $a, b, c \in \mathbb{R}$). On a les résultats suivants.

- Propriété(s) 2.3.1**
1. Les valeurs propres λ_1, λ_2 de A sont réelles.
 2. La matrice A est toujours diagonalisable par une matrice S telle que $\tilde{S}S = I$ (I désigne la matrice identité).
 3. Les deux valeurs propres sont strictement positives (resp. négatives) si et seulement si $\tilde{X}AX > 0$ (resp. $\tilde{X}AX < 0$) pour tout vecteur colonne non nul X .
Si $\lambda_1\lambda_2 < 0$ alors $\tilde{X}AX$ peut changer de signe en fonction de X .
 4. Les deux valeurs propres sont strictement positives (resp. négatives) si et seulement si le déterminant de A est strictement positif et $a > 0$ (resp. $a < 0$).
On a $\lambda_1\lambda_2 < 0$ si et seulement si le déterminant de A est strictement négatif.

Preuve (1) C'est immédiat par simple calcul : les valeurs propres d'une matrice sont les zéros de son polynôme caractéristique, lequel est ici

$$\lambda \mapsto \det \begin{pmatrix} a - \lambda & c \\ c & b - \lambda \end{pmatrix}$$

ou encore, en développant

$$\begin{aligned} \det \begin{pmatrix} a - \lambda & c \\ c & b - \lambda \end{pmatrix} &= (a - \lambda)(b - \lambda) - c^2 \\ &= \lambda^2 - (a + b)\lambda + ab - c^2. \end{aligned}$$

Le réalisant (discriminant) de ce polynôme est égal à

$$(a+b)^2 - 4(ab - c^2) = a^2 + 2ab + b^2 - 4ab + 4c^2 = a^2 - 2ab + b^2 + 4c^2 = (a-b)^2 + 4c^2;$$

comme a, b, c sont des réels et que le carré d'un réel est toujours positif, ce réalisant est un réel positif. On peut donc conclure que le polynôme a deux zéros réels.

(2) Soit

$$X_1 = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix}$$

un vecteur propre réel de valeur propre λ_1 . Quitte à le multiplier par une constante non nulle, on peut supposer que

$$x_1^2 + y_1^2 = \widetilde{X}_1 X_1 = 1.$$

Cela étant, définissons S par

$$S = \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix}.$$

On obtient immédiatement

$$\widetilde{S}S = \begin{pmatrix} x_1 & y_1 \\ -y_1 & x_1 \end{pmatrix} \begin{pmatrix} x_1 & -y_1 \\ y_1 & x_1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I.$$

Montrons alors que

$$S^{-1}AS = \widetilde{S}AS$$

est une matrice diagonale. La première colonne de S est le vecteur propre X_1 ; notons Z la seconde colonne de S et remarquons que X_1 et Z sont des vecteurs colonnes orthogonaux, c'est-à-dire $\widetilde{X}_1 Z = 0 = \widetilde{Z} X_1$. Cela étant, les colonnes de la matrice AS sont les vecteurs AX_1 et AZ ; on peut donc écrire

$$AS = (AX_1 \ AZ).$$

On obtient ainsi

$$\widetilde{S}AS = \begin{pmatrix} \widetilde{X}_1 \\ \widetilde{Z} \end{pmatrix} (AX_1 \ AZ) = \begin{pmatrix} \widetilde{X}_1 AX_1 & \widetilde{X}_1 AZ \\ \widetilde{Z} AX_1 & \widetilde{Z} AZ \end{pmatrix}.$$

Comme X_1 est vecteur propre de A de valeur propre λ_1 et tel que $\widetilde{X}_1 X_1 = 1$, on a $AX_1 = \lambda_1 X_1$ et

$$\widetilde{S}AS = \begin{pmatrix} \lambda_1 & \widetilde{X}_1 AZ \\ \lambda_1 \widetilde{Z} X_1 & \widetilde{Z} AZ \end{pmatrix}.$$

Mais $\widetilde{Z} X_1 = 0$ donc

$$\widetilde{S}AS = \begin{pmatrix} \lambda_1 & \widetilde{X}_1 AZ \\ 0 & \widetilde{Z} AZ \end{pmatrix}.$$

On sait aussi que A est symétrique, c'est-à-dire $\widetilde{A} = A$. Ainsi,

$$\widetilde{\widetilde{S}AS} = \widetilde{S}\widetilde{A}S = \widetilde{S}AS.$$

La matrice

$$\widetilde{S}AS = \begin{pmatrix} \lambda_1 & \widetilde{X}_1 AZ \\ 0 & \widetilde{Z} AZ \end{pmatrix}$$

est donc symétrique, c'est-à-dire que ses éléments situés de part et d'autre de la diagonale principale sont égaux. Il s'ensuit que $\widetilde{X}_1 AZ = 0$ donc que $\widetilde{S}AS$ est une matrice diagonale.

Pour (3) : **A compléter (voir cours enseigné).**

(4) Le polynôme caractéristique de A est le polynôme

$$\lambda \mapsto \lambda^2 - (a+b)\lambda + ab - c^2 = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1 \lambda_2.$$

On en déduit que

$$\begin{cases} a+b = \lambda_1 + \lambda_2 \\ ab - c^2 = \det(A) = \lambda_1 \lambda_2 \end{cases}$$

On obtient donc tout de suite (seconde égalité) que le déterminant de A est strictement négatif si et seulement si $\lambda_1 \lambda_2 < 0$ c'est-à-dire si et seulement si les valeurs propres sont de signe différent.

Il est tout aussi évident vu la seconde égalité que le déterminant de A est strictement positif si et seulement si $\lambda_1 \lambda_2 > 0$ c'est-à-dire si et seulement si les deux valeurs propres sont de même signe.

On peut alors conclure. De fait

-si les deux valeurs propres sont strictement positives (resp. négatives) alors le déterminant est strictement positif et $ab = c^2 + \det(A) > 0$ donc a et b ont le même signe. Comme $a + b = \lambda_1 + \lambda_2$, on obtient que $a + b > 0$ (resp. $a + b < 0$) donc a et b sont strictement positifs (resp. négatifs) puisque de même signe et de somme strictement positive (resp. négative).

-réciproquement, si $a > 0$ (resp. $a < 0$) et $\det(A) > 0$, alors les deux valeurs propres sont de même signe ; comme $ab = c^2 + \det(A) > 0$, a et b sont aussi de même signe, c'est-à-dire positif (resp. négatif). Une fois encore, de l'égalité $a + b = \lambda_1 + \lambda_2$, on en déduit que si $a > 0$ (resp. $a < 0$) alors les valeurs propres sont strictement positives (resp. négatives).

□

Remarque : une matrice S telle que $\tilde{S}S = I$ est appelée *matrice orthogonale*. On a alors $S^{-1} = \tilde{S}$ et $S\tilde{S} = I$. Les égalités $\tilde{S}S = I$ et $S\tilde{S} = I$ signifient que les colonnes de S ont des vecteurs orthonormés, de même que les lignes.

2.4 Retour aux extrema : cas des extrema dits « libres »

Grâce aux résultats concernant les matrices réelles symétriques qui viennent s'ajouter au développement limité de Taylor, nous avons les outils pour démontrer un résultat donnant des conditions suffisantes pour qu'un point stationnaire soit un extremum.

Propriété(s) 2.4.1 Soient un ouvert Ω de $\mathbb{R}^2$, une fonction $f \in C_2(\Omega)$, à valeurs réelles et un point (x_0, y_0) appartenant à Ω , stationnaire pour f .

(1) Si $\det H_f(x_0, y_0) > 0$ et si $D_1^2 f(x_0, y_0)$ est un réel strictement positif, alors le point (x_0, y_0) est un minimum local strict de f dans Ω .

(2) Si $\det H_f(x_0, y_0) > 0$ et si $D_1^2 f(x_0, y_0)$ est un réel strictement négatif, alors le point (x_0, y_0) est un maximum local strict de f dans Ω .

(3) Si $\det H_f(x_0, y_0) < 0$ alors le point (x_0, y_0) n'est pas un extremum de f dans Ω .

Preuve. Pour démontrer ce résultat, il suffit de reprendre le développement limité de Taylor de f à l'ordre 2 au point stationnaire (x_0, y_0) , d'utiliser une propriété de la continuité qui permet d'avoir une inégalité autour d'un point lorsque on l'a seulement au point et d'utiliser les résultats de la section 2.3.

Comme (x_0, y_0) appartient à un ouvert, on sait qu'il existe $\varepsilon > 0$ tel que si (x, y) est tel que $|x - x_0| < \varepsilon$ et $|y - y_0| < \varepsilon$, alors $(x, y) \in \Omega$ (cela signifie que si (x, y) est voisin du point (x_0, y_0) de l'ouvert alors il est aussi dans cet ouvert). De plus, un tel point est donc dans un carré centré en (x_0, y_0) , donc le segment qui le relie à (x_0, y_0) est aussi inclus dans le carré, donc dans l'ouvert.

Cela étant, par le développement limité de Taylor, on sait alors qu'il existe un point (u_0, v_0) sur le segment qui relie (x_0, y_0) et (x, y) tel que

$$f(x, y) = f(x_0, y_0) + \frac{(x - x_0)^2}{2} D_1^2 f(u_0, v_0) + (x - x_0)(y - y_0) D_1 D_2 f(u_0, v_0) + \frac{(y - y_0)^2}{2} D_2^2 f(u_0, v_0)$$

(les termes avec les dérivées premières sont nuls puisque (x_0, y_0) est un point stationnaire). Notons X le vecteur colonne

$$\begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}$$

On obtient ainsi

$$f(x, y) = f(x_0, y_0) + \frac{1}{2} \tilde{X} H_f(u_0, v_0) X$$

où H_f est la matrice hessienne de f .

Maintenant, on voudrait bien sûr utiliser les hypothèses sur la matrice hessienne, lesquelles font intervenir le point stationnaire et non (u_0, v_0) ! C'est donc le moment de se rappeler une propriété fort importante des fonctions continues. Comme f est deux fois continûment dérivable, les dérivées secondes sont continues, donc aussi le déterminant de sa matrice hessienne.

Les hypothèses de (1) impliquent donc l'existence de $\eta > 0$, que l'on peut bien sûr prendre plus petit que ε , tel que l'on ait encore $\det H_f(s, t) > 0$ et $D_1^2 f(s, t) > 0$ pour autant que $|s - x_0| < \eta$ et $|t - y_0| < \eta$.

On va alors conclure facilement en utilisant les point (3) et (4) de la Propriété 2.3.1. De fait, en prenant (x, y) tel que $|x - x_0| < \eta$ et $|y - y_0| < \eta$, on a aussi $|u_0 - x_0| < \eta$ et $|v_0 - y_0| < \eta$. Dès lors $\det H_f(u_0, v_0) > 0$ et $D_1^2 f(u_0, v_0) > 0$, donc les deux valeurs propres de la matrice $H_f(u_0, v_0)$ sont strictement positives et on en déduit finalement

$$\tilde{X}H_f(u_0, v_0)X > 0$$

pour autant que $(x, y) \neq (x_0, y_0)$. On obtient ainsi

$$f(x, y) = f(x_0, y_0) + \frac{1}{2}\tilde{X}H_f(u_0, v_0)X > f(x_0, y_0)$$

et on en conclut que (x_0, y_0) est bien un minimum local strict de f dans Ω .

Bien sûr la preuve s'adapte tout de suite au cas des hypothèses (2).

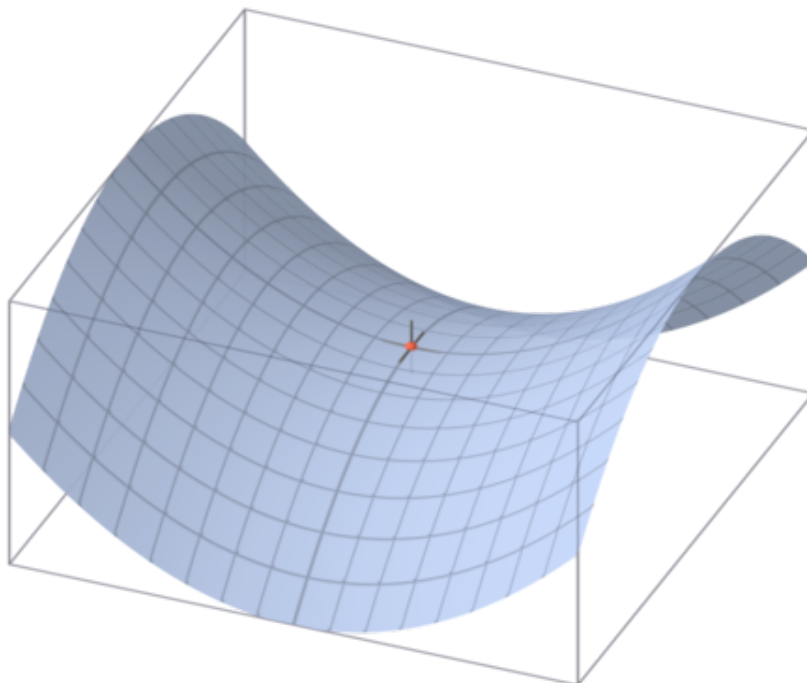
Enfin, si le déterminant de la matrice hessienne de f en (x_0, y_0) est négatif, il le reste dans un voisinage de ce point, les valeurs propres de $H_f(u_0, v_0)$ sont donc de signe opposé et l'on applique à nouveau (3) et (4) de la propriété 2.3.1 pour finalement conclure que comme $\tilde{X}H_f(u_0, v_0)X$ peut changer de signe, le point (x_0, y_0) n'est pas un extremum de f dans Ω . $\square$

Considérons quelques exemples.

- La fonction $f(x, y) = x^2 - y^2$, $(x, y) \in \mathbb{R}^2$ est indéfiniment continûment dérivable sur $\mathbb{R}^2$ et est à valeurs réelles. Sa représentation graphique est la surface d'équation cartésienne $z = x^2 - y^2$; il s'agit donc d'un paraboloïde hyperbolique (voir la partie consacrée aux quadriques). Le seul point stationnaire de f est l'origine; la matrice hessienne de f est constante (c'est bien sûr le cas pour toutes les fonctions polynomiales de degré 2) et est

$$\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Le déterminant de cette matrice est égal à -4 et le point stationnaire n'est donc pas un extremum. On parle de « point selle », car la forme de la surface autour de ce point ressemble à une selle de cheval. On aurait d'ailleurs pu voir tout de suite (à savoir sans passer par la matrice hessienne) que l'origine n'est pas un extremum. En effet, on a $f(x, 0) = x^2 \geq 0 = f(0, 0)$ pour tout x et $f(0, y) = -y^2 \leq 0 = f(0, 0)$ pour tout y .



• La fonction $f(x, y) = x^2 - 2xy + y^3/3 - 3y$, $(x, y) \in \mathbb{R}^2$ est indéfiniment continûment dérivable sur $\mathbb{R}^2$ et est à valeurs réelles. Cherchons-en les points stationnaires. On a

$$D_1f(x, y) = 2x - 2y \quad \text{et} \quad D_2f(x, y) = -2x + y^2 - 3$$

donc

$$\begin{aligned} \begin{cases} D_1f(x, y) = 0 \\ D_2f(x, y) = 0 \end{cases} &\Leftrightarrow \begin{cases} x = y \\ y^2 - 2y - 3 = 0 \end{cases} \\ &\Leftrightarrow \begin{cases} x = y \\ (y+1)(y-3) = 0 \end{cases} \\ &\Leftrightarrow x = y = -1 \quad \text{ou} \quad x = y = 3. \end{aligned}$$

Cela étant, la matrice hessienne de f est la matrice

$$\begin{pmatrix} D_1^2f(x, y) & D_1D_2f(x, y) \\ D_1D_2f(x, y) & D_2^2f(x, y) \end{pmatrix} = \begin{pmatrix} 2 & -2 \\ -2 & 2y \end{pmatrix}.$$

Au point stationnaire $(-1, -1)$, elle est égale à

$$\begin{pmatrix} 2 & -2 \\ -2 & -2 \end{pmatrix},$$

matrice dont le déterminant est égal à -8 . Ce point stationnaire n'est donc pas extremum.

Au point stationnaire $(3, 3)$, elle est égale à

$$\begin{pmatrix} 2 & -2 \\ -2 & 6 \end{pmatrix},$$

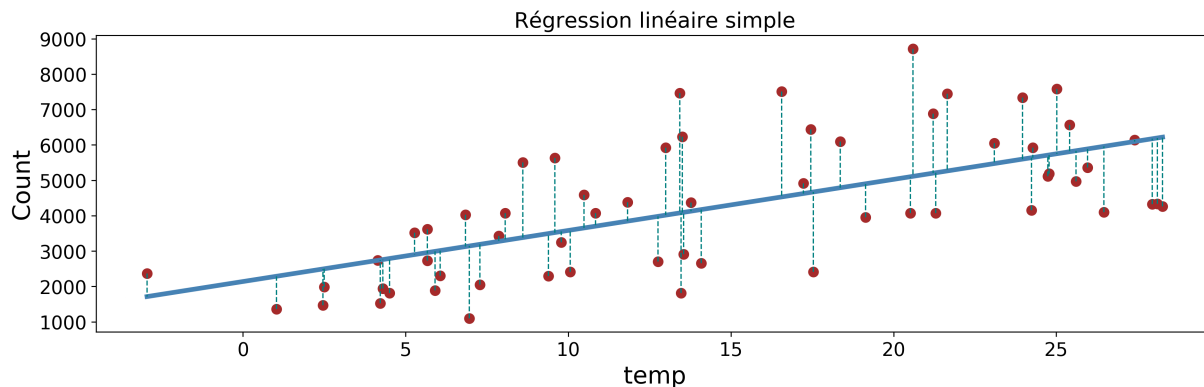
matrice dont le déterminant est égal à 8 . Comme la dérivée seconde de f par rapport à la première variable vaut 2 (en tout point), la fonction a un minimum local strict au point stationnaire $(3, 3)$. Regardons alors si ce minimum est global. On a $f(3, 3) = -9$; comme $f(0, y) = y^3/3 - 3y$, on a $\lim_{y \rightarrow -\infty} f(0, y) = -\infty$; dès lors le minimum est seulement local.

• Le résultat théorique (conditions suffisantes pour avoir un extremum) ne permet pas toujours de conclure, parce que les hypothèses ne sont pas vérifiées. Par exemple il se peut que la fonction ne soit pas deux fois dérivable. Il se peut aussi que le déterminant de la matrice hessienne soit nul, ou que la dérivée seconde en la première variable soit nulle. Dans ce cas, il faut revenir à la définition des extrema.

Par exemple, pour $f(x, y) = x^4 + y^4$ ($x, y) \in \mathbb{R}^2$ et $g(x, y) = x^3 + y^3$ ($x, y) \in \mathbb{R}^2$, le seul point stationnaire est l'origine. Mais en ce point, les dérivées secondes de f et de g sont nulles. Dès lors la matrice hessienne est nulle et on ne peut pas conclure en utilisant le résultat théorique précédent. Cependant, il est clair que f admet un minimum global strict à l'origine. Et pour g , comme on a $g(x, 0) = x^3$ et $g(0, 0) = 0$, l'origine n'est pas extremum étant donné que l'on a $g(x, 0) = x^3 > 0$ si x est positif et $g(x, 0) = x^3 < 0$ si x est négatif.

2.5 Application à la régression linéaire

Nous renvoyons ici à des pages extraites d'un ancien syllabus, en annexe.



On a

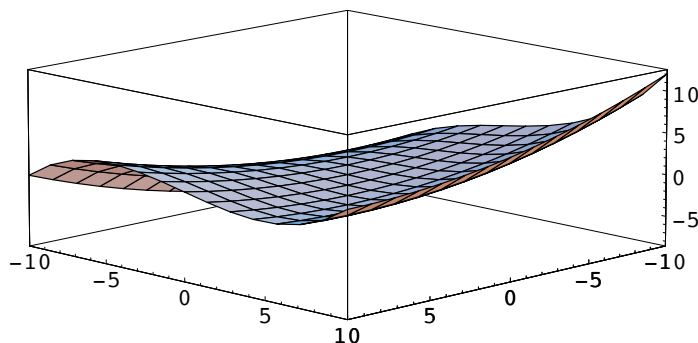
$$\det(H_f(3, 3)) = 8, \quad D_1^2 f(3, 3) = 2$$

donc $(3, 3)$ correspond à un minimum local strict.

On a

$$\det(H_f(-1, -1)) = -8$$

donc $(-1, -1)$ n'est pas un extremum.



5.3 Application: la régression linéaire

Il s'agit d'une application de l'étude des extrema des fonctions dites "quadratiques".

5.3.1 Généralités

Une fonction de deux variables réelles² du type

$$f(x, y) = a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + c$$

où les coefficients des variables x, y sont des réels et où les a_{ij} ne sont pas tous nuls est appelée *fonction quadratique*. On les a déjà rencontrées précédemment dans le cadre de ce cours car elles sont intimement liées aux coniques et aux quadriques: dans le plan muni d'un repère, une équation du type $f(x, y) = 0$ est l'équation cartésienne d'une conique (en toute généralité) et dans l'espace muni d'un repère, une équation du type $z = f(x, y)$ est l'équation d'une quadrique³.

Dans ce cas simple, le système permettant de déterminer les points stationnaires de f s'écrit

$$\begin{cases} D_x f(x, y) = 2a_{11}x + 2a_{12}y + 2b_1 = 0 \\ D_y f(x, y) = 2a_{22}y + 2a_{12}x + 2b_2 = 0 \end{cases}$$

et la matrice hessienne

$$H_f = \begin{pmatrix} D_x^2 f(x, y) & D_x D_y f(x, y) \\ D_x D_y f(x, y) & D_y^2 f(x, y) \end{pmatrix} = 2 \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$$

est constante. En un point stationnaire (x_0, y_0) , on a ainsi

$$f(x, y) = f(x_0, y_0) + \frac{1}{2} \begin{pmatrix} x - x_0 & y - y_0 \end{pmatrix} H_f \begin{pmatrix} x - x_0 \\ y - y_0 \end{pmatrix}, \quad \forall x, y \in \mathbb{R}.$$

On en déduit donc la propriété suivante (dans CE cas de fonction f).

² cela s'étend à plus de deux variables réelles

³ de type 'parabolôïde'. Il s'agit en effet d'un cas particulier d'une équation du type $F(x, y, z) = 0$ où F est une fonction quadratique de trois variables réelles.

Propriété 5.3.1 a) Si le déterminant de la matrice hessienne est non nul, alors il existe un point stationnaire unique.

b) Un extremum est toujours global.

□

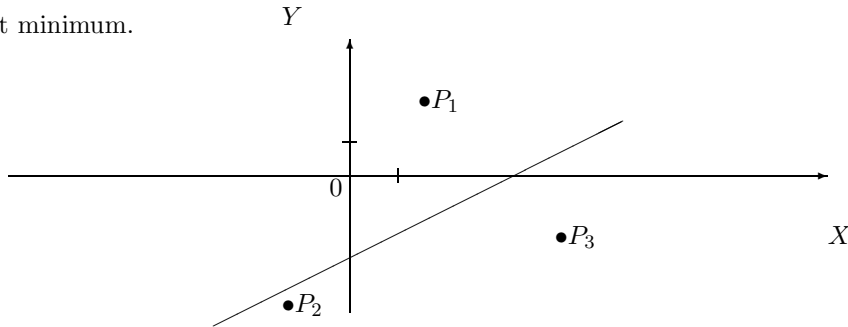
5.3.2 La régression linéaire

En guise d'application, examinons le problème de la *régression linéaire*.

Dans le plan muni d'un repère orthonormé, on donne un nombre fini de points distincts $P_j (j = 1, \dots, J)$ par leurs coordonnées cartésiennes $(x_j, y_j) (j = 1, \dots, J)$. Dans le cas où les x_j ne sont pas tous égaux, on demande de chercher la droite⁴ d (dite *droite de régression*) d'équation cartésienne $y = mx + b$ telle que

$$E(m, b) = \sum_{j=1}^J (y_j - (mx_j + b))^2$$

soit minimum.



Comme $|y_j - (mx_j + b)|$ est la valeur absolue de la différence entre l'ordonnée du point P_j et celle du point de d dont l'abscisse est x_j , l'expression $E(m, b)$ mesure d'une certaine manière la façon dont les points P_j sont répartis autour de⁵ d . La fonction E définie ci-dessus s'appelle *l'erreur quadratique totale*.

Le problème est donc celui de la recherche des extrema (libres) éventuels d'une fonction quadratique, notée ici E et dont les variables sont notées m et b .

Dans ce cas particulier⁶, l'expression de E permet de conclure immédiatement à l'existence d'un minimum (que l'on sait donc aussi être global). La recherche de l'équation cartésienne de d (correspondant à la recherche du minimum de E) consiste donc en la détermination de ou des solutions du système caractérisant les points stationnaires.

En utilisant les notations de la section précédente, les coefficients de $m^2, b^2, 2mb$ dans l'expression de E , sont respectivement

$$a_{11} = \sum_{j=1}^J x_j^2, \quad a_{22} = J, \quad a_{12} = \sum_{j=1}^J x_j$$

⁴existence et unicité aussi

⁵remarquons que dans le cas où les points ont tous la même abscisse r , la droite verticale d'équation $x = r$ contient tous les points P_j

⁶bien sûr si l'on s'attache à la méthode générale, c'est l'examen de la matrice hessienne qui permet de conclure au fait que l'on a affaire à un minimum

ce qui correspond à la matrice hessienne

$$H_E = 2 \begin{pmatrix} \sum_{j=1}^J x_j^2 & \sum_{j=1}^J x_j \\ \sum_{j=1}^J x_j & J \end{pmatrix}.$$

En utilisant les hypothèses sur les abscisses x_j et la propriété ci-dessous, on en déduit *qu'il existe un point stationnaire unique et que celui-ci donne lieu à un minimum strict global*.

Propriété. Par récurrence (sur J), montrons que, quel que soit $J \in \mathbb{N}_0$ et quels que soient les réels x_j ($j = 1, \dots, J$), on a

$$\left(\sum_{j=1}^J x_j \right)^2 \leq J \sum_{j=1}^J x_j^2.$$

De plus, on a l'égalité si et seulement si tous les x_j sont égaux.

De fait, pour $J = 1$ et aussi pour $J = 2$ c'est clair. Supposons à présent la propriété correcte pour J . On obtient successivement

$$\begin{aligned} \left(\sum_{j=1}^{J+1} x_j \right)^2 &= \left(\sum_{j=1}^J x_j + x_{J+1} \right)^2 = \left(\sum_{j=1}^J x_j \right)^2 + x_{J+1}^2 + 2x_{J+1} \left(\sum_{j=1}^J x_j \right) \\ &\leq J \sum_{j=1}^J x_j^2 + x_{J+1}^2 + 2x_{J+1} \left(\sum_{j=1}^J x_j \right) \\ &\leq J \sum_{j=1}^J x_j^2 + x_{J+1}^2 + \left(\sum_{j=1}^J (x_j^2 + x_{J+1}^2) \right) \\ &= (J+1) \sum_{j=1}^{J+1} x_j^2 \end{aligned}$$

Le cas de l'égalité se démontre en affinant un peu la preuve ci-dessus. $\square$

5.4 Extrema dans des ensembles bornés fermés

5.5 Extrema liés: la méthode des multiplicateurs de Lagrange

2.6 Les extrema liés (extrema « sous contrainte »)

En plus du cours enseigné, en guise d'introduction, on peut aller voir les pages en annexe (obtenues via wikipedia).

Considérons tout d'abord le cas d'une fonction f à valeurs réelles, de deux variables, définie et continûment dérivable dans un ouvert Ω pour laquelle on doit rechercher les éventuels extrema, mais pas dans Ω tout entier, mais seulement parmi les points de Ω qui sont situés sur une courbe, par exemple d'équation $g(x, y) = 0$, où g est une fonction continûment dérivable dans Ω , à valeurs réelles. On dit que l'on cherche à extrémiser f sous la contrainte $g = 0$.

On va montrer le résultat suivant.

Propriété(s) 2.6.1 *Si le point (x_0, y_0) de Ω est un extremum local de f sous la contrainte $g = 0$ alors il existe $\lambda \in \mathbb{R}$ tel que (x_0, y_0, λ) vérifie*

$$\begin{cases} D_1 f(x_0, y_0) = \lambda D_1 g(x_0, y_0) \\ D_2 f(x_0, y_0) = \lambda D_2 g(x_0, y_0) \\ g(x_0, y_0) = 0 \end{cases}$$

Preuve. Voir plus loin $\square$

Remarquer que le point (x_0, y_0, λ) est alors un point stationnaire pour la fonction (appelée « Lagrangien »)

$$L : (x, y, \lambda) \mapsto f(x, y) - \lambda g(x, y),$$

car

$$\begin{cases} D_1 L(x_0, y_0, \lambda) = D_1 f(x_0, y_0) - \lambda D_1 g(x_0, y_0) = 0 \\ D_2 L(x_0, y_0, \lambda) = D_2 f(x_0, y_0) - \lambda D_2 g(x_0, y_0) = 0 \\ D_3 L(x_0, y_0, \lambda) = -g(x_0, y_0) = 0 \end{cases}$$

Rechercher un extremum sous contrainte consiste donc à tout d'abord s'assurer qu'il existe, ensuite à résoudre le système d'équations ci-dessus et ensuite à chercher le ou les extrema en évaluant la fonction en chacune des solutions trouvées. Rappelons que si l'ensemble $\{(x, y) \in \Omega : g(x, y) = 0\}$ est fermé borné et que f est continu sur cet ensemble, alors les extrema existent.

Venons-en alors à la justification de la propriété 2.6.1. Un résultat dont la preuve sort du cadre de ce cours affirme ce qui suit (on peut se convaincre que cela peut être vrai en prenant des exemples concrets, cf cours).

Proposition 2.6.2 *Soit $g \in C_1(\Omega)$, à valeurs réelles et soit $(x_0, y_0) \in \Omega$ tel que $(D_1 g(x_0, y_0), D_2 g(x_0, y_0)) \neq (0, 0)$. Alors il existe un intervalle ouvert I de $\mathbb{R}$, un réel $t_0 \in I$, des fonctions g_1, g_2 , à valeurs réelles, appartenant à $C_1(I)$, un voisinage V de (x_0, y_0) tels que $g_1(t_0) = x_0$, $g_2(t_0) = y_0$ et*

$$g(x, y) = 0 \quad (x, y) \in V \quad \Leftrightarrow \quad \exists t \in I : \begin{cases} x = g_1(t) \\ y = g_2(t) \end{cases}$$

De plus $(Dg_1(t_0), Dg_2(t_0)) \neq (0, 0)$.

Cela étant, rechercher les extrema locaux de f sous la contrainte $g = 0$ revient donc à chercher les extrema de la fonction

$$F(t) = t \mapsto f(g_1(t), g_2(t)), \quad t \in I$$

On retombe donc dans le cas d'extrema libres pour une fonction d'une seule variable. Si maintenant (x_0, y_0) est extremum de f sous la contrainte $g = 0$, alors t_0 est extremum de F donc est un point stationnaire pour celle-ci car F est continûment dérivable dans I (comme composée de fonction continûment dérivables). Comme

$$DF(t) = (D_1 f)(g_1(t), g_2(t)) Dg_1(t) + (D_2 f)(g_1(t), g_2(t)) Dg_2(t)$$

on obtient l'égalité (*)

$$0 = DF(t_0) = (D_1f)(x_0, y_0) Dg_1(t_0) + (D_2f)(x_0, y_0) Dg_2(t_0).$$

Mais on a aussi

$$g(g_1(t), g_2(t)) = 0 \quad \forall t \in I;$$

on en déduit que (**)

$$0 = (D_1g)(x_0, y_0) Dg_1(t_0) + (D_2g)(x_0, y_0) Dg_2(t_0).$$

Les deux égalités (*) et (**) peuvent se réécrire

$$\begin{cases} (D_1f)(x_0, y_0) Dg_1(t_0) + (D_2f)(x_0, y_0) Dg_2(t_0) = 0 \\ (D_1g)(x_0, y_0) Dg_1(t_0) + (D_2g)(x_0, y_0) Dg_2(t_0) = 0 \end{cases}$$

ou encore, sous forme matricielle

$$\begin{pmatrix} D_1f(x_0, y_0) & D_2f(x_0, y_0) \\ D_1g(x_0, y_0) & D_2g(x_0, y_0) \end{pmatrix} \begin{pmatrix} Dg_1(t_0) \\ Dg_2(t_0) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Cela étant, comme $(Dg_1(t_0), Dg_2(t_0)) \neq (0, 0)$, on en déduit que la matrice

$$\begin{pmatrix} D_1f(x_0, y_0) & D_2f(x_0, y_0) \\ D_1g(x_0, y_0) & D_2g(x_0, y_0) \end{pmatrix}$$

n'est pas inversible ou, de manière équivalente, que son déterminant est nul. Maintenant, comme la seconde ligne de la matrice n'est pas nulle, la première en est un multiple donc il existe $\lambda \in \mathbb{R}$ tel que

$$\begin{cases} D_1f(x_0, y_0) = \lambda D_1g(x_0, y_0) \\ D_2f(x_0, y_0) = \lambda D_2g(x_0, y_0) \end{cases}$$

Comme $g(x_0, y_0) = 0$, on a donc bien obtenu le système annoncé dans la propriété 2.6.1

2.7 Les extrema sur des ensembles fermés bornés

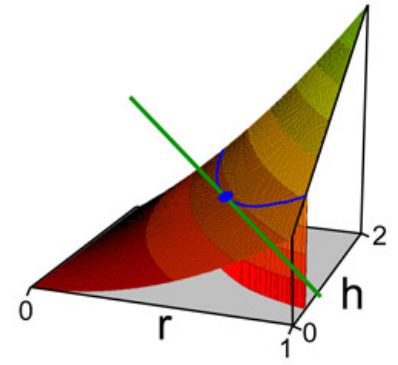
Voir cours et répétitions.

Exemple introductif

Soit v_0 un nombre strictement positif. L'objectif est de trouver la portion de cylindre de rayon r et de hauteur h de surface minimale (couvercles compris) et de volume v_0 . Pour cela on définit deux fonctions, v et s qui à (r, h) associent respectivement le volume et la surface de la portion de cylindre. On dispose des égalités

$$\forall r, h \in \mathbb{R}_+ \quad v(r, h) = \pi r^2 h \quad \text{et} \quad s(r, h) = 2\pi r(r + h).$$

La figure de droite représente la fonction s , qui à r et h associe la surface. La ligne bleue correspond aux points de volume égal à 1. L'objectif est de trouver le point bleu, de plus petite surface pour un volume égal à 1. La fonction s n'est autre que la fonction φ du préambule. La fonction ψ et la fonction L sont définies par :



La nappe correspond à la surface du cylindre, la courbe bleue aux points de volume égal à v_0 , choisi dans la représentation égal à 1.

$$\forall r, h \in \mathbb{R}_+ \quad \psi(r, h) = v(r, h) - v_0 \quad \text{et} \quad \forall \lambda \in \mathbb{R} \quad L(r, h, \lambda) = s(r, h) + \lambda \psi(r, h).$$

La méthode de Lagrange consiste à rechercher un point tel que la différentielle de L soit nulle. Sur un tel point, la dérivée partielle par rapport à λ est nulle, ce qui signifie que la fonction ψ est nulle, ou encore que la contrainte est respectée. Si l'on identifie s avec son approximation linéaire tangente, son comportement sur la contrainte, aussi identifiée à son approximation linéaire tangente, est aussi nécessairement nulle à partir de l'ordre 1. Ce comportement est illustré par la droite en vert sur la figure. Le long de cette droite, la fonction ψ est nulle, et le terme d'ordre 1 de la fonction s l'est alors nécessairement.

Il suffit, en conséquence, de calculer la différentielle de L , et plus précisément ses trois dérivées partielles, pour l'exemple choisi :

$$\frac{\partial L}{\partial r} = 2\pi(h + 2r + \lambda hr) = 0, \quad \frac{\partial L}{\partial h} = \pi(2r + \lambda r^2) = 0, \quad \frac{\partial L}{\partial \lambda} = \pi r^2 h - v_0 = 0.$$

On trouve les valeurs suivantes :

$$r = -\frac{2}{\lambda} = \left(\frac{v_0}{2\pi}\right)^{1/3}, \quad h = -\frac{4}{\lambda} = 2\left(\frac{v_0}{2\pi}\right)^{1/3} \quad \text{et} \quad \lambda = -2\left(\frac{2\pi}{v_0}\right)^{1/3}.$$

Autrement dit :

$$h = 2r \quad \text{et} \quad v_0 = \pi r^2 h, \text{ d'où } s = 6\pi r^2.$$

Cet exemple possède l'avantage d'une représentation graphique simple, guidant l'intuition. En revanche, la méthode du multiplicateur de Lagrange n'est pas nécessaire dans ce cas : on peut simplement exprimer la valeur de h pour que le volume du cylindre respecte la contrainte imposée au volume v_0 . On trouve :

$$h = \frac{v_0}{\pi r^2}.$$

En injectant cette contrainte dans l'équation décrivant l'aire, il vient :

$$s = 2\pi r^2 + \frac{2v_0}{r}$$

et il suffit de trouver la valeur de r minimisant cette fonction pour trouver la solution. De même qu'avec le multiplicateur de Lagrange, on trouve :

$$h = 2r.$$