Analyse: cours général A

Yulia Kuznetsova

15 octobre 2025

Table des matières

Chapitre 1

Éléments de la géométrie différentielle

1.1 Rapports

Les leçons concernées (2025) sont :

214 Théorème d'inversion locale, théorème des fonctions implicites. Illustrations en analyse et en géométrie.

Rapport 2024 : Les deux théorèmes fondamentaux auxquels cette leçon est consacrée offrent une belle utilisation de la complétude, qu'il conviendra d'évoquer. La démonstration de l'un de ces deux théorèmes peut parfaitement faire l'objet d'un des deux développements. On pourra par exemple mettre en pratique, sur des exemples bien choisis, le théorème des fonctions implicites au moins dans le cas de deux variables réelles, pour enrichir le plan avec profit.

Des applications significatives aussi bien en analyse qu'en géométrie sont attendues : problèmes d'optimisation sous contraintes (inégalité de Hölder, inégalité d'Hadamard, etc), régularité des racines d'un polynôme en fonction des coefficients, etc.

La méthode des multiplicateurs de Lagrange a bien évidemment toute sa place dans cette leçon, à condition qu'elle soit illustrée par des exemples. L'interprétation de l'énoncé en termes d'espace tangent est visuellement éclairante et permet d'éviter les éventuelles confusions résultant de raisonnements purement matriciels. Les candidates et candidats solides peuvent s'intéresser à l'étude locale d'applications suffisamment régulières (submersions, immersions, théorème du rang constant, lemme de Morse), au lemme de Sard, ainsi qu'aux sous-variétés de $\mathbb{R}^n$.

215 Applications différentiables définies sur un ouvert de $\mathbb{R}^n$. Exemples et applications.

L'idée de départ de cette leçon est qu'une fonction suffisamment régulière se comporte localement comme une application linéaire. De nombreuses différentielles usuelles (notamment issues de l'algèbre linéaire) peuvent ainsi être obtenues en calculant directement un développement limité. Des exemples significatifs en dimension 2 et 3 pourront venir illustrer la différence fondamentale avec la dimension 1. Les dérivées partielles lorsqu'elles existent pourront clarifier l'expression de nombreuses différentielles ainsi que la règle de la chaîne.

Les candidates et candidats solides peuvent s'intéresser la notion de différentielle seconde pour les fonctions de classe C2, à la différentielle de l'exponentielle matricielle, ainsi qu'aux points où celle- ci est un difféomorphisme local, aux fonctions harmoniques et à leurs propriétés élémentaires, à la caractérisation des fonctions holomorphes et son interprétation géométrique.

Pour ce qui concerne les applications, de nombreux thèmes relatifs aux leçons 214 ou 219 sont ici appropriés.

219 Extremums : existence, caractérisation, recherche. Exemples et applications.

Cette leçon offre aux candidates et candidats une multitude d'approches possibles : utilisation de la topologie, du calcul différentiel, de la convexité (fonctions convexes, projection sur un convexe fermé et leurs multiples applications), de l'holomorphie.

Les candidates et candidats peuvent proposer des problèmes d'optimisation sous contraintes, si possible autres que la preuve de l'inégalité arithmético-géométrique. À ce sujet, une bonne compréhension de la méthode des multiplicateurs de Lagrange requiert celle de la notion d'espace tangent, qui en donne une justification beaucoup plus claire que certains raisonnements purement matriciels. Les algorithmes de recherche d'extremums ont également leur place dans cette leçon (méthode de Newton, du gradient à pas optimal, problème des moindres carrés, etc).

Les candidates et candidats solides peuvent s'intéresser aux diverses versions du principe du maximum (fonctions holomorphes ou harmoniques, équations aux dérivées partielles), au calcul des variations, ou réfléchir à l'unicité de la meilleure approximation dans divers espaces fonctionnels, à commencer par celle des fonctions continues sur un segment par des polynômes de degré au plus égal à d.

1.2 Théorème d'inversion locale

Théorème 1.2.1. *(*Inversion locale.*) Soit $U \subset \mathbb{R}^n$ un ouvert et soit $f\colon U \to \mathbb{R}^n$ une fonction de classe C^1. Soit $a \in U$ tel que $df(a)\colon \mathbb{R}^n \to \mathbb{R}^n$ est inversible. Alors il existe un ouvert $V \subset U$ contenant a et un ouvert $W \subset \mathbb{R}^n$ contenant $f(a)$ tels que $f(V) = W$ et la fonction induite $f\colon V \to W$ est un difféomorphisme.*

On dit dans ce cas que f est un difféomorphisme local en a.

Démonstration. 1 On va commencer par simplifier le contexte et montrer qu'on peut se placer dans le cas $a = f(a) = 0$, $df(a) = I$.

On pose $U' = U - a$ et $g\colon U' \to \mathbb{R}^n$ définie par

$$g(z) = df(a)^{-1}\big(f(a+z) - f(a)\big) \qquad \text{pour tout } z \in U'.$$

L'ouvert U' contient 0, $g(0) = 0$ et par composition, g est de classe C^1 et $dg(0) = I_n$.

Si on trouve deux ouverts $V' \ni 0$, $W' \ni 0$ tels que $g : V' \to W'$ est un difféomorphisme, alors l'égalité $f(x) = f(a) + df(a)g(x-a)$ implique que f est une bijection de $V = V' + a$ sur $W = f(a) + df(a)W'$, ce qui est un ouvert contenant $f(a)$.

Les fonctions inverses sont liées par $f^{-1}(y) = g^{-1}\Big([df(a)]^{-1}(y - f(a))\Big) + a$, ce qui montre que f^{-1} est aussi de classe C^1 et $f : V \to W$ un difféomorphisme.

2 On suppose désormais

$$a = 0,\ f(0) = 0, \ \text{ et } \ df(0) = I_n, \tag{1.1}$$

et on trouvera des ouverts $V \ni 0$ et $W \ni 0$ tels que

$$f(V) = W \qquad \text{et que} \qquad f\colon V \to W \text{ est bijective.} \tag{1.2}$$

Comme la fonction $df\colon U \to L(\mathbb{R}^n)$ est continue, il existe $r > 0$ tel que $\overline{B}(0,r) \subset U$ et

$$\forall\, x \in \overline{B}(0,r), \qquad \|df(x) - I_n\| \le \frac{1}{2}\,. \tag{1.3}$$

On pose $W = B(0, \frac{r}{2})$ et

$$V = f^{-1}\big(W\big) \cap B(0,r) \;=\; \big\{x \in \mathbb{R}^n \;:\; \|x\| < r \text{ et } \ \|f(x)\| < \tfrac{r}{2}\big\}.$$

Puisque f est continue, V est un ouvert ; il est clair que V contient 0 et est inclus dans U, et que $f(V) \subset B(0, \frac{r}{2})$. Montrons qu'en réalité, f est bijective de V sur W.

Pour cela, on fixe $y \in W = B(0, \frac{r}{2})$. La recherche de son inverse se fera par le théorème de point fixe. Associons à y la fonction $\varphi_y \colon U \to \mathbb{R}^n$ définie par

$$\varphi_y(x) = x - \big(f(x) - y\big).$$

On observe que $\varphi_y(x) = x$ si et seulement si $f(x) = y$.

La fonction φ_y est différentiable et par addition, $d\varphi_y(x) = I_n - df(x)$ pour tout $x \in U$. Donc par (1.3), $\|d\varphi_y(x)\| \leq \frac{1}{2}$ pour tout $x \in \overline{B}(0,r)$. La boule fermée $\overline{B}(0,r)$ est convexe donc d'après l'inégalité des accroissements finis, on a

$$\forall\, x, x' \in \overline{B}(0,r), \qquad \|\varphi_y(x) - \varphi_y(x')\| \leq \frac{1}{2}\,\|x - x'\|, \tag{1.4}$$

c'est-à-dire, φ_y est $\frac{1}{2}$-Lipschitzienne.

De plus, $f(0) = 0$ donc $\varphi_y(0) = y$, et en appliquant l'inégalité triangulaire, on obtient pour tout $x \in \overline{B}(0,r)$,

$$\|\varphi_y(x)\| \leqslant \|\varphi_y(x) - \varphi_y(0)\| + \|\varphi_y(0)\| \leq \frac{1}{2}\,\|x\| + \|y\| < \frac{r}{2} + \frac{r}{2} = r.$$

Donc $\varphi_y\big(\overline{B}(0,r)\big) \subset B(0,r) \subset \overline{B}(0,r)$. La boule fermée $\overline{B}(0,r)$ est un espace métrique complet donc d'après le Théorème du point fixe de Banach, il existe un unique $x \in \overline{B}(0,r)$ tel que $\varphi_y(x) = x$. Puisque φ_y envoie la boule fermée $\overline{B}(0,r)$ dans la boule ouverte $B(0,r)$, x appartient à cette dernière. On a donc $x \in V$ et $f(x) = y$. Comme le point fixe de φ_y sur $\overline{B}(0,r)$ est unique, x est l'unique élément de V tel que $f(x) = y$. Ceci étant vrai pour tout $y \in B(0, \frac{r}{2})$, on obtient la bijectivité de f.

3 Il reste maintenant à vérifier que l'application inverse

$$f^{-1} \colon W = B(0, \tfrac{r}{2}) \longrightarrow V$$

est de classe C^1. En sachant que sa différentielle est nécessairement $df^{-1}(y) = [df(f^{-1}(x))]^{-1}$, il suffit de montrer qu'elle est différentiable.

Montrons d'abord que f ne diminue pas beaucoup les distances. Pour tout $x, x' \in V$ et pour tout $y \in W$ on a, par (1.4),

$$\|\big(x - (f(x) - y)\big) - \big(x' - (f(x') - y)\big)\| = \|x - x' - \big(f(x) - f(x')\big)\| \leqslant \frac{1}{2}\,\|x - x'\|$$

(donc cette expression ne dépend pas de y), d'où par l'inégalité triangulaire

$$\|f(x) - f(x')\| \geqslant \frac{1}{2}\,\|x - x'\|. \tag{1.5}$$

Cela implique notamment que f^{-1} est 2-Lipschitzienne : $\|f^{-1}(y) - f^{-1}(y')\| \leqslant 2\|y - y'\|$ pour tous $y, y' \in W$.

Ensuite, par (1.3), $df(x)$ est inversible pour tout $x \in \overline{B}(0, r)$, avec l'inverse de norme $\|[df(x)]^{-1}\| \leqslant 2$:

$$\|[df(x)]^{-1}\| = \Big\| \sum_{k=0}^{\infty} (I - df(x))^k \Big\| \leq \sum_{k=0}^{\infty} \Big(\frac{1}{2}\Big)^k = 2.$$

On peut montrer maintenant par définition que pour $y \in W$, l'application $T = [df(f^{-1}(y))]^{-1}$ est la différentielle de f^{-1} en y. Soit $\alpha > 0$ tel que $B(y, \alpha) \subset W$. Pour tout $y' \in B(y, \alpha)$ on note $x = f^{-1}(y)$, $x' = f^{-1}(y')$, et on obtient :

$$\begin{aligned} \|f^{-1}(y') - f^{-1}(y) - T(y' - y)\| &= \|T\big(df(x)(x' - x) + f(x) - f(x')\big)\| \\ &\leqslant 2\, \|f(x') - [f(x) + df(x)(x' - x)]\|. \end{aligned} \tag{1.6}$$

Comme f est différentiable en x, pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour $x' \in B(x, \delta)$ (on peut supposer cette boule contenue dans V) on a

$$\|f(x') - [f(x) + df(x)(x' - x)]\| < \varepsilon \|x - x'\|.$$

Comme f^{-1} est 2-Lipschitzienne, il suffit de poser $\beta = \min(\alpha, \delta/2)$ pour garantir, pour tout $y' \in B(y, \beta)$, l'inclusion $x' \in B(x, \delta)$. On peut donc majorer l'expression (1.6) par

$$< 2\varepsilon \|x - x'\| \leqslant 4\varepsilon \|y - y'\|,$$

$y' \in B(y, \beta)$, ce qui montre la différentiabilité de f^{-1}. □

Évidemment, $df(a)$ est inversible si et seulement si $\det\big(J_f(a)\big) \neq 0$.

A priori, il n'y a aucune raison pour qu'un difféomorphisme local soit un difféomorphisme au sens global du terme. Nous verrons au Corollaire 1.2.4 comment déduire une version 'globale' du Théorème d'inversion locale.

Exemple 1.2.2. *Soit $f : M_n(\mathbb{R}) \to M_n(\mathbb{R})$, $A \mapsto A^2$. Pour $A, H \in M_n(\mathbb{R})$ on a*

$$f(A + H) - f(A) = (A + H)^2 - A^2 = AH + HA + H^2.$$

Comme $H^2 = o(\|H\|)$, $H \to 0$, l'application linéaire $L : H \mapsto AH + HA$ est la différentielle $df(A)$ de f en A.

Si $A = I$, on a $df(I)(H) = 2H$ et $df(I) = 2I_{M_n(\mathbb{R})}$. Cette application est inversible, donc f est un difféomorphisme local en I : il existe un ouvert W contenant I tel que la racine carrée f^{-1} est bien définie et de classe C^1 sur W.

Exercice 1.2.3. *Obtenir la conclusion similaire pour* $\exp : X \mapsto e^X$ *sur $M_n(\mathbb{R})$.*

Corollaire 1.2.4. *(*Inversion globale.*) Soient $U \subset \mathbb{R}^n$ un ouvert et $f : U \to \mathbb{R}^n$ une fonction* <u>*injective*</u> *de classe C^1. Si, pour tout $x \in U$, l'application linéaire $df(x) : \mathbb{R}^n \to \mathbb{R}^n$ est inversible, alors l'image $f(U)$ est un ouvert de $\mathbb{R}^n$ et la fonction $f : U \to f(U)$ est un difféomorphisme.*

Démonstration. Puisque f est injective, la fonction induite $f : U \to f(U)$ est une bijection. On note $f^{-1} : f(U) \to U$ sa fonction réciproque. Soit $y \in f(U)$ et soit $x = f^{-1}(y)$. Par hypothèse, $df(x)$ est inversible donc par le Théorème 1.2.1, il existe un ouvert $V \subset U$ contenant x et un ouvert W contenant y tels que $f(V) = W$ et $f : V \to W$ est un difféomorphisme. Pour plus de clarté, on notera ici $f_{|V}$ cette fonction restreinte.

D'une part, l'ouvert W est inclus dans $f(U)$. Ainsi tout point de $f(U)$ admet un voisinage ouvert inclus dans $f(U)$, ce qui prouve que $f(U)$ est ouvert. D'autre part, la restriction de f^{-1} à W coincide avec $(f_{|V})^{-1}$ et cette fonction est de classe C^1. Donc f^{-1} est de classe C^1 au voisinage de y. Ceci étant vrai pour tout $y \in f(U)$, on obtient que la fonction f^{-1} est de classe C^1. □

Exercice 1.2.5.

1. *Montrer que l'application*

$$\varphi : (r, \theta) \mapsto (x, y) = (r \cos\theta, r \sin\theta)$$

est un C^1-difféomorphisme de l'ouvert $]0, \infty[\times] - \pi, \pi[$ sur le plan privé de la demi-droite $\mathbb{R}_-$. Si $f(x, y) = g(r, \theta)$, donner les formules de passage entre les dérivées partielles de f et celles de g.

2. *Soit U le plan privé de l'origine, et $f(x, y) = (x^2 - y^2, 2xy)$. Montrer que f est un difféomorphisme local à tout point de U mais n'est pas un difféomorphisme global.*

3. *Soit h l'application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ définie par $(x, y) \mapsto (e^x \cos y, e^x \sin y)$. Montrer que h est de classe C^1 dans $\mathbb{R}^2$; que $dh(x, y)$ est inversible pour tout (x, y) de $\mathbb{R}^2$; mais que h n'est pas un homéomorphisme de $\mathbb{R}^2$ sur $h(\mathbb{R}^2)$.*

1.3 Théorème des fonctions implicites

On suppose connu le théorème de difféomorphisme local.

Exemple : $x^2 + y^2 = R^2$ est le graphe d'une fonction $y = h(x)$ si $y \neq 0$, $x = h(y)$ sinon ; à remarquer que $2y = \dfrac{\partial}{\partial y}(x^2 + y^2 - R^2)$.

Théorème 1.3.1. *(*Fonctions implicites.*) Soit $U \subset \mathbb{R}^p \times \mathbb{R}^n$ un ouvert et soit $f\colon U \to \mathbb{R}^n$ une fonction de classe C^1. Soit $a = (x_0, y_0) \in U$ tel que $f(x_0, y_0) = 0$ et l'application linéaire $\partial_y f(x_0, y_0)\colon \mathbb{R}^n \to \mathbb{R}^n$ est inversible.*

Alors il existe un ouvert $V \subset \mathbb{R}^p$ contenant x_0, un ouvert $\Omega \subset U$ contenant (x_0, y_0) et une fonction $\varphi\colon V \to \mathbb{R}^n$ de classe C^1 tels que pour tout $(x, y) \in U$,

$$\big\{(x,y) \in \Omega \quad \textit{et} \quad f(x,y) = 0\big\} \iff \big\{x \in V \quad \textit{et} \quad y = \varphi(x)\big\}. \tag{1.7}$$

Notez que si dans cet énoncé on partait d'une fonction $f : U \to \mathbb{R}^m$, l'hypothèse d'inversiblité $\partial_y f(x_0, y_0)$ imposerait la condition $n = m$. C'est pourquoi on se place d'emblée dans ce cadre.

Démonstration. On définit $g\colon U \to \mathbb{R}^p \times \mathbb{R}^n$ en posant

$$g(x,y) = \big(x, f(x,y)\big) \qquad \text{pour tout } (x,y) \in U.$$

Puisque $f(x_0, y_0) = 0$, on a $g(x_0, y_0) = (x_0, 0)$. De plus la fonction g est de classe C^1 et

$$\forall\,(x,y) \in U,\ \forall\,(h,k) \in \mathbb{R}^p \times \mathbb{R}^n, \qquad \big[dg(x,y)\big](h,k) \;=\; \big(h, \big[df(x,y)\big](h,k)\big).$$

Soient $L = dg(x_0, y_0) \in L(\mathbb{R}^p \times \mathbb{R}^n)$, $A = \partial_x f(x_0, y_0) \in L(\mathbb{R}^p, \mathbb{R}^n)$ et $B = \partial_y f(x_0, y_0) \in L(\mathbb{R}^n)$. D'après ce qui précède, on a

$$L(h,k) = \big(h, A(h) + B(k)\big)$$

pour tout $(h, k) \in \mathbb{R}^p \times \mathbb{R}^n$. Par hypothèse, B est inversible, ce qui permet de définir une application $T \in L(\mathbb{R}^p \times \mathbb{R}^n)$ en posant

$$T(u,v) = \big(u, -B^{-1}A(u) + B^{-1}(v)\big) \qquad \text{pour tout } (u,v) \in \mathbb{R}^p \times \mathbb{R}^n.$$

On vérifie que $T \circ L$ et $L \circ T$ sont égaux à l'identité de $\mathbb{R}^p \times \mathbb{R}^n$, donc $L = dg(x_0, y_0)$ est inversible (d'inverse T).

Nous sommes donc en mesure d'appliquer le Théorème d'inversion locale à g au point (x_0, y_0). D'après celui-ci, il existe un ouvert $\Omega \subset U$ contenant (x_0, y_0) et un ouvert $\Omega' \subset \mathbb{R}^p \times \mathbb{R}^n$ contenant $(x_0, 0)$ tels que $g(\Omega) = \Omega'$ et $g\colon \Omega \to \Omega'$ est un difféomorphisme. Quitte à restreindre, on peut supposer que $\Omega' = V \times V'$, avec $V \subset \mathbb{R}^n$ ouvert contenant x_0 et $V' \subset \mathbb{R}^p$ ouvert contenant 0. On note $g^{-1}\colon V \times V' \to \Omega \subset \mathbb{R}^p \times \mathbb{R}^n$ la fonction inverse de la restriction de g à Ω. Compte tenu de la définition de g, il existe $\theta\colon V \times V' \to \mathbb{R}^n$ de classe C^1 tel que

$$g^{-1}(x,z) = (x, \theta(x,z)) \qquad \text{pour tout } \ (x,z) \in V \times V'.$$

Alors pour tous $x \in \mathbb{R}^p, y, z \in \mathbb{R}^n$, on a

$$\big\{(x,y) \in \Omega \text{ et } f(x,y) = z\big\} \iff \big\{(x,z) \in V \times V' \text{ et } g(x,z) = y\big\}$$

par définition de l'inverse. On définit alors $\varphi \colon V \to \mathbb{R}^n$ en posant $\varphi(x) = \theta(x,0)$ pour tout $x \in V$, et il est clair que cette fonction vérifie les conclusions du théorème. □

Le théorème dit donc que la surface $\{f = 0\}$ est localement le graphe d'une fonction φ.

On peut calculer la dérivée / différentielle de la fonction implicite : $F\big(x, \varphi(x)\big) \equiv 0$ dans V implique que pour tout i, j

$$0 = \frac{\partial F_i}{\partial x_j}\big(x, \varphi(x)\big) + \sum_{l=1}^{m} \frac{\partial F_i}{\partial y_l}\big(x, \varphi(x)\big) \frac{\partial \varphi_l}{\partial x_j}(x),$$

ou, en forme matricielle,

$$-d_x F\big(x, \varphi(x)\big) = d_y F\big(x, \varphi(x)\big) d\varphi(x),$$

d'où

$$d\varphi(x) = -\Big(d_y F\big(x, \varphi(x)\big)\Big)^{-1} d_x F\big(x, \varphi(x)\big).$$

Si $x = x_0$, on a $(x, \varphi(x)) = a$ et alors

$$d\varphi(x_0) = -\big(d_y F(a)\big)^{-1} d_x F(a).$$

Il est donc possible de calculer la différentielle de φ en a sans connaitre la fonction φ elle-même.

Si f est de classe $C^{(k)}$, $2 \leqslant k \leqslant \infty$, alors φ l'est aussi : on le montre par récurrence à partir de la formule ci-dessus.

Exemple 1.3.2. *L'équation $u^3 - 2u^2x + uxy - 2 = 0$ définit une fonction implicite $u(x,y)$ dans un voisinage de $(1,1)$ avec $u(1,1) = 2$, car $\frac{\partial}{\partial u}(u^3 - 2u^2x + uxy - 2) = 3u^2 - 4ux + xy$ vaut $5 \neq 0$ en $(1,1,2)$. Les dérivées partielles de cette fonction sont :*

$$\frac{\partial u}{\partial x}(1,1) = -\Big(\frac{\partial F}{\partial u}(1,1,2)\Big)^{-1} \frac{\partial F}{\partial x}(1,1,2) = -\frac{1}{5}(-2u^2 + uy)|_{(1,1,2)} = \frac{6}{5},$$

$$\frac{\partial u}{\partial y}(1,1) = -\Big(\frac{\partial F}{\partial u}(1,1,2)\Big)^{-1} \frac{\partial F}{\partial y}(1,1,2) = -\frac{1}{5}(ux)|_{(1,1,2)} = -\frac{2}{5}.$$

On peut donc écrire sa forme approchée, par la formule de Taylor :

$$u(x,y) = 2 + \frac{6}{5}(x-1) - \frac{2}{5}(y-1) + o\big(\|(x-1, y-1)\|\big).$$

Exercice 1.3.3. 1 *Soit* $f(x,y,z) = z^3 - 2z^2x + xyz - 2$. *Montrer que dans un voisinage de* $(1,1,2)$ *la fonction implicite* $z = \varphi(x,y)$ *existe. Trouver ses dérivées partielles et le développement limité du premier ordre en* $(1,1)$.

2 *Soit*

$$F(x,y,u,v) = \begin{pmatrix} xu + yv - u^3 \\ x + y + u + v \end{pmatrix}.$$

Montrer que dans un voisinage de $(1,0,1,-2)$ *on peut exprimer* (u,v) *en tant que fonction de* (x,y).
Écrire sa différentielle en $(1,0)$.

3 *Soit* $f(x,y,z) = (x^2 - y^2 + z^2 - 1, xyz - 1)$. *Supposons que* $f(x_0, y_0, z_0) = (0,0)$. *Montrer qu'il existe un ouvert* I *de* $\mathbb{R}$ *contenant* x_0 *et une application* $\varphi : I \to \mathbb{R}^2$ *telle que* $\varphi(x_0) = (y_0, z_0)$ *et* $f(x, \varphi(x)) = (0,0)$ *pour tout* $x \in I$.

Exemple 1.3.4. Les racines d'un polynôme. *Soit* $P(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ *un polynôme à coefficients réels scindé à racines simples. Montrer qu'il existe un voisinage* V *de* $A = (a_{n-1}, .., a_1, a_0)$ *tel que pour tout* $(\alpha_{n-1}, \ldots, \alpha_1, \alpha_0) \in V$, *le polynôme* $Q_\alpha(x) = x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0$ *est scindé à racines simples et ses racines s'expriment comme fonctions* C^∞ *de* $(\alpha_{n-1}, \ldots, \alpha_0)$.

Solution : pour chaque racine x_k *de* P, $k = 1, \ldots, n$, *on applique le théorème de fonctions implicites en* $(a_0, \ldots, a_{n-1}, x_k)$ *à*

$$F(\alpha_0, \ldots, \alpha_{n-1}, x) = x^n + \alpha_{n-1}x^{n-1} + \cdots + \alpha_1 x + \alpha_0.$$

Les racines de P *sont supposées simples, donc*

$$\frac{\partial F}{\partial x}(a_0, \ldots, a_{n-1}, x_k) = P'(x_k) \neq 0,$$

ce qui implique l'existence d'une fonction implicite φ_k *dans un voisinage* V *de* A *telle que* $\varphi(A) = x_k$, *et d'un ouvert* $\Omega_k \ni (A, x_k)$ *tel que* $\{F = 0\} \cap \Omega_k = \{(\alpha, \varphi_k(\alpha)) : \alpha \in V\}$. *On peut supposer le voisinage* V *le même pour tout* k *; en le diminuant si besoin on peut garantir que les ouverts* Ω_k *sont disjoints. Pour* $\alpha \in V$ *on aura donc* $F(\alpha, \varphi_1(\alpha)) = \cdots = F(\alpha, \varphi_n(\alpha)) = 0$, *c'est-à-dire,* $\varphi_k(\alpha)$, $k = 1, \ldots, n$, *sont les racines de* Q_α. *Comme il y en a* n *et elles sont distinctes, il n'y en a pas d'autres.*

Exercice 1.3.5. *Application : montrer qu'au voisinage d'une matrice réelle* $A \in M_n(\mathbb{R})$, *possédant* n *valeurs propres réelles distinctes, les matrices* M *gardent* n *valeurs propres distinctes et que ces dernières peuvent s'exprimer comme des fonctions* C^∞ *de* M.

1.4 Extréma liés

Théorème 1.4.1. *Soit $U \subset \mathbb{R}^n$ un ouvert, $f : U \to \mathbb{R}$, $G : U \to \mathbb{R}^m$ de classe C^1, $m \leqslant n$. Soit $M = \{x \in U : G(x) = 0\}$. Si $a \in M$ est un extrémum local de $f|_M$ et $dG(a)$ a le rang maximal (= m), alors il existent $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ dits les multiplicateurs de Lagrange tels que*

$$\nabla f(a) = \lambda_1 \nabla G_1(a) + \cdots + \lambda_m \nabla G_m(a). \tag{1.8}$$

Démonstration. Si $n = m$, l'hypothèse dit que $dG(a)$ est inversible, donc les colonnes de cette matrice engendrent linéairement $\mathbb{R}^n \ni \nabla f(a)$. On suppose donc par la suite $m < n$.

Soit $p = \operatorname{rang} dG(a)$. Il existent donc des lignes libres $\nabla G_{i_k}(a)$, $k = 1, \dots, p$. Quitte à changer la numérotation, on peut supposer que ce sont $\nabla G_k(a)$, $k = n-p+1, \dots, n$. Notons dans la suite $x = (z, y) \in \mathbb{R}^n$ avec $z \in \mathbb{R}^p$ et $y \in \mathbb{R}^{n-p}$. Nous sommes dans le hypothèses du théorème de fonction implicite : il existe un ouvert V contenant $z_0 = (a_1, \dots, a_{n-p})$, un ouvert $\Omega \subset U$ contenant a et une fonction $\varphi : V \to \Omega$ de classe C^1 tels que

$$M \cap \Omega = \{(z, \varphi(z)) : z \in V\}.$$

Maintenant la fonction $h : V \to \mathbb{R}$, $z \mapsto f(z, \varphi(z))$ admet un extrémum local en z_0, donc sa différentielle est nulle. On calcule

$$dh(z_0) = \partial_z f(a) + \partial_y f(a) d\varphi(z_0) = \partial_z f(a) - \partial_y f(a) \big[\partial_y G(a)\big]^{-1} \partial_z G(a) = 0$$

(ce sont des vecteurs lignes de $\mathbb{R}^p$).

La matrice $\partial_y G(a)$ est inversible, ses lignes forment donc une base de $\mathbb{R}^m$, et il existent $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ tels que

$$\partial_y f(a) = \lambda_1 \partial_y G_1(a) + \cdots + \lambda_m \partial_y G_m(a) = (\lambda_1, \dots, \lambda_m) \partial_y G(a).$$

En le mettant dans la formule de $dh(z_0)$, on obtient

$$\partial_z f(a) - (\lambda_1, \dots, \lambda_m) \partial_y G(a) \big[\partial_y G(a)\big]^{-1} \partial_z G(a) = 0,$$

d'où $\partial_z f(a) = (\lambda_1, \dots, \lambda_m) \partial_z G(a)$. On a donc l'égalité en chaque coordonnée, ce qui vérifie (1.8). $\square$

Géométriquement : la surface M est orthogonale aux gradients des fonctions G_i, $i = 1, \dots, m$. Si f a un extrémum, sa dérivée «en toute direction sur la surface» est nulle, donc son gradient y est orthogonal lui aussi, d'où la décomposition (**??**).

En pratique : on obtient n équations, a priori non-linéaires, en n variables x, paramétrées par $\lambda \in \mathbb{R}^m$.

Exemple 1.4.2. *Trouver les extréma de $f(x,y) = xy$ sur le cercle $S^1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.*

On a $\nabla f(x,y) = (y,x)$, remarquons donc que f (définie sur $\mathbb{R}^2 \setminus \{(0,0)\}$) n'a pas d'extrémum local. Mais sur le compact S^1 elle aura bien un maximum et un minimum ; trouvons-les.

Le cercle est l'ensemble de zéros de $g(x,y) = x^2+y^2-1$, à gradient $\nabla g(x,y) = (2x, 2y)$ qui ne s'annule pas sur S^1. Dans un extrémum lié il doit exister alors $\lambda \in \mathbb{R}$ tel que $\nabla f = \lambda \nabla g$, ce qu'on explicite :

$$\begin{cases} y = 2\lambda x, \\ x = 2\lambda y. \end{cases}$$

Donc $x = 4\lambda^2 x$. Comme $x = 0$ implique $y = 0$ et ne donne pas de points du cercle, on obtient $4\lambda^2 = 1$ et $\lambda = \pm 1/2$. Par conséquence, $x = y$ ou $x = -y$. Le premier cas donne deux points $\pm(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$, le deuxième deux autres points $\pm(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}})$.

La valeur de f en ses points est respectivement $\frac{1}{2}$ et $-\frac{1}{2}$, ce sont donc deux maximums et deux minimums.

Exemple 1.4.3. *Si $dG(a)$ n'a pas le rang maximal, alors la méthode ne permet pas toujours de trouver les extrémums, même s'ils existent. Soit $g(x,y) = y^2 - x^3$, $f(x,y) = (x+1)^2 + y^2 = d\big((x,y),(-1,0)\big)^2$. La distance entre le point $P = (-1,0)$ et un ponit de la courbe $y^2 = x^3$ est minimisé évidemment en $(x,y) = (0,0)$. Calculons les gradients : $\nabla f(x,y) = (2(x+1), 2y)$, $\nabla f(0,0) = (2,0)$, $\nabla g(x,y) = (-3x^2, 2y)$, $\nabla g(0,0) = (0,0)$. Les hypothèses du théorème de sont pas vérifiées en ce minimum lié, et la conclusion non plus : un multiplicateur de Lagrange n'existe pas.*

Il faut donc vérifier à part les points où le rang n'est pas maximal.

Exercice 1.4.4. *Trouver les extréma de $f(x,y,z) = x+y+z$ sur l'ellipsoïde*

$$E = \{(x,y,z) : \frac{x^2}{2} + \frac{y^2}{4} + \frac{z^2}{6} = 1\}.$$

Exemple 1.4.5 (Inégalité de Hölder)**.** *Soient $p, q > 1$ des réels conjugués : $\frac{1}{p} + \frac{1}{q} = 1$. Alors pour tout $n \geqslant 1$ et tous réels positifs a_k, b_k, $k = 1, \ldots, n$ on a*

$$\sum_{k=1}^{n} a_k b_k \leqslant \Big(\sum_{k=1}^{n} a_k^p \Big)^{1/p} \Big(\sum_{k=1}^{n} b_k^q \Big)^{1/q}.$$

L'inégalité est bien sûr nontriviale que dans le cas $n \geqslant 2$, ce qu'on suppose dans la suite. On suppose aussi $A = \left(\sum_{k=1}^n a_k^p\right)^{1/p}$ non-nul, sinon $a_k = 0$ pour tout k l'inégalité est vérifiée. Quitte à diviser par A, on peut donc supposer $A = 1$.

On considère alors les b_k des paramètres fixes, et on cherche à maximiser $f(a) = \sum_k a_k b_k$ sous la contrainte $g(a) = \sum_{k=1}^n a_k^p = 1$. Les gradients sont :

$$\nabla f(a) = (b_1, \ldots, b_n), \qquad \nabla g(a) = p\,(a_1^{p-1}, \ldots, a_n^{p-1}),$$

et celui de g ne s'annule pas sur la surface en question. Dans les points extrémaux il existe alors $\lambda \in \mathbb{R}$ tel que $b_k = \lambda\, p\, a_k^{p-1}$ pour tout k. La contrainte donne

$$\sum_{k=1}^n a_k^p = \sum_{k=1}^n \left[\frac{b_k}{\lambda p}\right]^{p/(p-1)} = (\lambda p)^{-q} \sum_{k=1}^n b_k^q = 1,$$

car $\frac{p}{p-1} = (1 - \frac{1}{p})^{-1} = q$. En notant $\sum_{k=1}^n b_k^q = B^q$, on obtient $\lambda p = B$, donc l'extrémum est unique avec $a_k = \left(\frac{b_k}{\lambda p}\right)^{1/(p-1)}$. La valeur de f en ce point est

$$f(a) = \sum_{k=1}^n b_k^{\frac{1}{p-1}+1} (\lambda p)^{-\frac{1}{p-1}} = \sum_{k=1}^n b_k^{\frac{p}{p-1}} B^{-\frac{1}{p-1}} = B^{q-\frac{1}{p-1}} = B.$$

Il reste à montrer que c'est bien un maximum. Si $B = 0$, alors f est identiquement nulle. Sinon, en posant $a_k = 1$, $a_j = 0$ pour $j \neq k$, on obtient $b_k \leqslant B$ en tant que les valeurs de f. Or, si b_k sont tous égaux à B, vu que $n \geqslant 2$, on aurait la contradiction $nB^q = B^q$, ce qui montre qu'on a $b_k < B$ pour au moins un indice k, donc B est un maximum. On en déduit $f(a) \leqslant B$ pour tout a sous la contrainte $A = 1$, d'où l'inégalité.

Exercice 1.4.6 (Moyenne arithmétique et moyenne géométrique)**.** *Soient $(a_i)_{1 \leqslant i \leqslant n}$ des nombres réels positifs. Montrer que*

$$\frac{a_1 + \cdots + a_n}{n} \geqslant \sqrt[n]{a_1 \cdots a_n}.$$

Exercice 1.4.7. $f(x, y) = x + 2y$ *sous la contrainte* $g(x, y) = x^2 + xy + y^2 + y - \frac{13}{9} = 0$, $U = \mathbb{R}^2$.

Exercice 1.4.8. $f(x, y, z) = x^2 + y^2 + z^2$ *sous contrainte*

$$\begin{cases} \frac{x^2}{4} + y^2 + z^2 = 1 \\ x + y + z = 0 \end{cases}$$

Exercice 1.4.9. $f(x, y, z) = (x-1)^2 + (y-2)^2 + (z-3)^2$ *sous contrainte*

$$\begin{cases} x^2 + y^2 + z^2 = 21 \\ 3x + 2y + z = 0 \end{cases}$$

Exercice 1.4.10. *Trouvez le volume maximum d'une boite parallélépipédique rectangle dont la surface est égale à* S.

Exercice 1.4.11. *Optimiser* $f(x, y, z) = yz + xy$ *sous les contraintes* $xy = 1$, $y^2 + z^2 = 1$.

Exercice 1.4.12. *Minimiser* $f(x, y, z) = x^2 + y^2 + z^2$ *si* $x + y + z = 9$ *et* $x + 2y + 3z = 20$.

1.5 Submersion, immersions

Définition 1.5.1. *Soit* $U \subset \mathbb{R}^m$ *un ouvert. Une application différentiable* $f : U \to \mathbb{R}^n$ *est dite submersion si le rang de* $df(x)$ *est* n *pour tout* $x \in U$ *(il en suit* $n \leqslant m$*). On dit que* f *est une immersion si* $df(x)$ *est injective pour tout* $x \in U$ *(donc nécessairement* $n \leqslant m$*).*

Exemple 1.5.2. *Les fonctions suivantes sont-elles des immersions ? des submersions ?*

1. $f : \mathbb{R}^2 \to \mathbb{R}^3$, $(x, y) \mapsto (x, y, 0)$
2. $f : \mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (y, z)$
3. $f : \mathbb{R}^3 \to \mathbb{R}$, $(x, y, z) \mapsto xy + 2yz + 3xz$
4. $f : \mathbb{R} \to \mathbb{R}^2$, $t \mapsto (\sin(2t), \sin(3t))$
5. $f : \mathbb{R}^3 \to \mathbb{R}^2$, $(x, y, z) \mapsto (x^2 + y^2 + z^2, xy)$
6. $f : \mathbb{R}^2 \to \mathbb{R}^3$, $(x, y) \mapsto (e^x, \cos y, \sin y)$

1.6 Sous-variétés de $\mathbb{R}^n$

Les énoncés et les définitions de cette section suivent [2]. Dans la définition et dans le théorème il faut comprendre «difféomorphisme» et «lisse» dans le sens «de classe $C^{(k)}$», où k (qui peut être infini) est le même dans toutes les assertions. Donc strictement dit, on parle toujours d'une «sous-variété de classe $C^{(k)}$» en précisant le k si besoin.

Définition 1.6.1. *Une partie* $M \subset \mathbb{R}^n$ *est une sous-variété de dimension* p *de* $\mathbb{R}^n$ *si pour tout* $a \in M$ *il existe des voisinages ouverts* U *de* a *et* V *de* 0 *et un difféomorphisme* $f : U \to V$ *tels que* $f(U \cap M) = V \cap (\mathbb{R}^p \times \{0\})$.

Cette définition n'est pas la plus intuitive dans le cas de $\mathbb{R}^n$, mais elle correspond à la notion générale d'une sous-variété différentielle.

Exemple 1.6.2. *Si* $M = \{(x, y) \in \mathbb{R}^2 : x^2+y^2 = 1\}$*, sur* $U = \{(x, y) : y > 0\}$ *on peut poser* $f(x, y) = (x, x^2 + y^2 - 1)$*. On obtient* $V = \{(x, z) \in \mathbb{R}^2 : z > x^2 - 1\}$*, ce qui se vérifie par exemple en définissant la fonction inverse* $f^{-1}(x, z) = (x, \sqrt{z + 1 - x^2})$*. Ces* U, V, f *valent pour* $y > 0$ *; pour les autres points de* M *on construit des fonctions similaires.*

Théorème 1.6.3. *Soit* M *une partie de* $\mathbb{R}^n$*. Les propriétés suivantes sont équivalentes :*

1. M *est une sous-variété de dimension* p *de* $\mathbb{R}^n$ *;*
2. *Pour tout* $a \in M$ *il existe un ouvert* U *contenant* a *et une submersion* $g : U \to \mathbb{R}^{n-p}$ *(c.à.d. le rang de* $dg(u)$ *est* $n - p$ *pour tout* $u \in U$*) telle que* $U \cap M = g^{-1}(0)$ *;*
3. *Pour tout* $a \in M$ *il existe une ouvert* U *contenant* a*, un ouvert* $\Omega \subset \mathbb{R}^p$ *contenant 0 et une application* $h : \Omega \to \mathbb{R}^n$ *qui est à la fois immersion (c.à.d.* $dh(x)$ *est injective pour tout* $x \in \Omega$*) et un homéomorphisme de* Ω *sur* $U \cap M$ *;*
4. *Pour tout* $a \in M$ *il existe un ouvert* U *contenant* a*, un ouvert* $V \subset \mathbb{R}^p$ *contenant* $(a_1, \ldots, a_p)$ *et une application lisse* $G : V \to \mathbb{R}^{n-p}$ *tels que, après permutation éventuelle des coordonnées,* $U \cap M$ *soit égal au graphe de* G*.*

Démonstration. On note $m = n - p$.

1 $\Longrightarrow$ 2 Pour $a \in M$ soient U, V, f donnés par (1). On garde l'ouvert U et on pose $g(u) = \big(f^{(p+1)}(u), \ldots, f^{(n)}(u)\big) \in \mathbb{R}^m$ pour $u \in U$. Évidemment $g(u) = 0$ si $u \in U \cap M$. Réciproquement, si $g(u) = 0$, alors $f(u) \in V \cap (\mathbb{R}^p \times \{0\})$, donc il existe $v \in U \cap M$ tel que $f(u) = f(v)$. Par l'hypothèse f est un difféomorphisme donc $u = v$, et on conclut que $g^{-1}(0) = U \cap M$. La matrice de dg est constituée par les dernières m colonnes de la matrice de df ; elles sont donc linéairement indépendantes, ce qui montre que g est bien une submersion.

2 $\Longrightarrow$ 4 Les colonnes de la matrice $dg(a)$ sont $\dfrac{\partial g}{\partial u_j}$, $j = 1, \ldots, n$. Par l'hypothèse, parmi ces colonnes il y a m libres. Quitte à permuter les coordonnées, on peut admettre que ce sont celles avec $j = p + 1, \ldots, n$. Pour $u \in \mathbb{R}^n$ on note, comme dans le théorème des fonctions implicites, $u = (x, y)$ où $x \in \mathbb{R}^p$ et $y \in \mathbb{R}^m$. On peut appliquer ce théorème à la fonction g car $d_y g(a)$ est inversible, par le choix des coordonnées y. Il existent alors un ouvert $W \subset U$, un ouvert $V \subset \mathbb{R}^p$ et une fonction lisse $G : V \to \mathbb{R}^m$ tels que $W = \{(x, G(x)) : x \in V\}$.

4 $\Longrightarrow$ 3 On pose $\Omega = V$ et $h(x) = \big(x, G(x)\big)$. Par l'hypothèse, $h(\Omega) =$

$U \cap M$. La fonction h est continue, son inverse aussi car c'est la projection sur les premières p coordonnées, donc h est un homéomorphisme. Enfin, $dh(x) = \begin{pmatrix} I_p \\ dG(x) \end{pmatrix}$ est évidemment injective de $\mathbb{R}^p$ dans $\mathbb{R}^n$.

3 $\Longrightarrow$ 1 Par l'hypothèse, il y a p indépendantes parmi les lignes de la matrice $dh(a)$. En permutant les coordonnées si nécessaire, on peut admettre que ce soient le premières p. Soit p_x la projection sur ces coordonnées. Par Corollaire, il existent un ouvert W contenant $p_x a$ et un ouvert $V \subset \Omega$ contenant 0 tels que $g := p_x h : V \to W$ est un difféomorphisme.

On pose $f(x,y) = (x, y - p_y h(g^{-1}(x))$ sur $U_1 = U \cap (W \times \mathbb{R}^m)$, et on note $V = f(U_1)$. Pour tout $v \in V$ on pose $x = p_x v$ et $y = v + p_y h(g^{-1}(x))$ pour obtenir une fonction inverse. Il est évident que f et f^{-1} sont de classe $C^{(k)}$, f est alors un difféomorphisme de U_1 sur V.Enfin, si $(x,y) \in U_1 \cap M$, alors $(x,y) = h(z)$ avec $z \in \Omega$, donc $x = g(z)$ et $y = p_y h(z)$ d'où $f(x,y) \in \mathbb{R}^p \times \{0\}$. $\square$

Exemple 1.6.4. *1 Tout ouvert de $\mathbb{R}^n$ est une sous-variété, $f = \mathrm{Id}$ dans la définition. En particulier : $GL_n(\mathbb{R})$ dans $\mathbb{R}^{n^2}$.*

2 $M = SL_n(\mathbb{R})$ dans $\mathbb{R}^{n^2}$ est une sous-variété de dimension $n^2 - 1$: dans le Théorème (2), poser $g(A) = \det A - 1$. En tout point A, la dérivée partielle $\dfrac{\partial g}{\partial A_{ij}}$ est le mineur correspondant et il en existe au moins un de non-nul.

3 La sphère S^n est une sous-variété de dimension n dans $\mathbb{R}^{n+1}$. En (2) du Théorème, poser $g(x) = x_1^1 + \cdots + x_{n+1}^2 - 1$. Le gradient $dg(x) = 2x$ est non-nul en tout $x \in S^n$.

4 Le produit $M = M_1 \times M_2$ de sous-variétés de $\mathbb{R}^{n_j}$ de dimension p_j, $j = 1, 2$, est une sous-variété de $\mathbb{R}^{n_1+n_2}$ de dimension $p_1 + p_2$. Dans Théorème (4), poser $U = U_1 \times U_2$, $V = V_1 \times V_2$ et $G(x_1, x_2) = (G_1(x_1), G_2(x_2))$. Le même pour un produit de tout nombre fini de facteurs.
En particulier, le tore $\mathbb{T}^n = (S^1)^n$ est une variété de dimension n (qui peut être réalisé en tant que sous-variété de $\mathbb{R}^{2n}$, mais éventuellement de $\mathbb{R}^m$ avec m plus petit).

5 $O_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) : AA^t = I\}$, alors $O_n(\mathbb{R}) = F^{-1}(0)$ où $F(A) = AA^t - I$. On a $F(A+H) + I = (A+H)(A+H)^t = AA^t + AH^t + HA^t + HH^t$ donc $dF(A)(H) = AH^t + HA^t$. L'image de $dF(A)$ est contenu dans l'espace des matrices symétriques $Sym_n(\mathbb{R})$, mais aussi toute matrice symétrique S y est contenue car on peut poser $H = \frac{1}{2}SA$. La dimension de $Sym_n(\mathbb{R})$ est $n + (n-1) + \cdots + 1 = \dfrac{n(n+1)}{2}$, et s'est le rang de $dF(A)$ pour toute A orthogonale. Soit P la projection de $M_n(\mathbb{R})$ sur $Sym_n(\mathbb{R})$ (qui ne garde que les coordonnées *sur la diagonale et dessus), alors $P \circ F$ est une submersion. La dimension de $O_n(\mathbb{R})$ est donc $n^2 - \dfrac{n(n+1)}{2} = \dfrac{n(n-1)}{2}$.*

À noter, la dimension de $SO_n(\mathbb{R}) = O_n(\mathbb{R}) \cap \det^{-1}(]0, +\infty[)$ est la même.

6 *Le cone $M = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 - z^2 = 0, z \geqslant 0\}$ n'est pas une sous-variété de $\mathbb{R}^3$. Supposons, en vue d'une contradiction, qu'il l'est, alors par (3) du Théorème il existe un ouvert U contenant $0 \in M$, un ouvert $\Omega \subset \mathbb{R}^2$ contenant 0 (en considérant les autres points de M on obtient nécessairement $p = 2$) et une immersion $h : \Omega \to \mathbb{R}^3$ réalisant un homéomorphisme entre Ω et $U \cap M$. En particulier, la fonction $f : t \mapsto h(t, 0)$, définie dans un intervalle autour de 0, est dérivable, de dérivée non-nulle car $df(0) = dh(0,0)\begin{pmatrix}1\\0\end{pmatrix}$ et $dh(0,0)$ est injective. On a donc $f(t) \sim f'(0)t$, $t \to 0$ et $\|f(t)\| \sim C|t|$ avec $C = \|f'(0)\| \neq 0$. Au même temps, $f_3(t)$ atteint son minimum global en 0, donc $f_3'(0) = 0$ et $f_3(t) = o(t)$, $t \to 0$. Les valeurs de f étant dans M, on doit pourtant avoir $f_3(t)^2 = f_1(t)^2 + f_2(t)^2$ ce qui implique $\|f(t)\|^2 = o(t^2)$, contradiction avec l'équivalence $\|f(t)\| \sim C|t|$, $C \neq 0$, obtenue avant.*

1.7 Le plan tangent

Soit M une sous-variété de $\mathbb{R}^n$, qu'on suppose pour simplicité égal à

$$M = F^{-1}(0) = \{x, \varphi(x) : x \in V\},$$

où $F : U \to \mathbb{R}^m$, $m < n$, est de classe C^1 avec $dF(a)$ de rang m en tout $a \in M$, et $V \subset \mathbb{R}^p$ est un ouvert, $p = n - m$, et φ de classe C^1.

Soit $a \in M$, et soit $\gamma : [-\varepsilon, \varepsilon] \to M$ (avec $\varepsilon > 0$) de classe C^1, telle que $\gamma(0) = a$. Sa dérivée $\gamma'(0)$ en 0 est un vecteur de $\mathbb{R}^n$; on définit T_aM comme l'ensemble de tels vecteurs tangents.

Il faut vérifier que T_aM est un sous-espace vectoriel de $\mathbb{R}^n$. Pour le voir, on se rappelle que $\gamma(t) = (x(t), y(t)) \in M$ donc $y(t) = \varphi(x(t))$ pour tout t, alors

$$\begin{aligned}\gamma'(t) &= \Big(x_1'(t), \ldots, x_p'(t), \sum_{k=1}^{p} \frac{\partial \varphi_1}{\partial x_k}(\gamma(t))x_k'(t), \ldots, \sum_{k=1}^{p} \frac{\partial \varphi_m}{\partial x_k}(\gamma(t))x_k'(t)\Big) \\ &= (x'(t), d\varphi(\gamma(t))[x'(t)]),\end{aligned}$$

en particulier, $\gamma'(0) = (x'(0), d\varphi(a)[x'(0)])$. On peut avoir $h \in \mathbb{R}^p$ quelconque en tant que $x'(0)$: il suffit de poser $\gamma(t) = (x_0 + th, \varphi(x_0 + th))$ si $a = (x_0, y_0)$. Donc

$$T_aM = \{(h, d\varphi(a)[h]) : h \in \mathbb{R}^p\},$$

c'est un sous-espace vectoriel de dimension p.

Exemple 1.7.1. *Soit $\varphi(x,y) = \sqrt{R^2 - x^2 + y^2}$, alors M est une demi-sphère (on suppose $x^2 + y^2 < R^2$). En tout point $a = (x,y,z)$, on calcule*

$$d\varphi(x,y) = \begin{pmatrix} -\frac{x}{z} \\ -\frac{y}{z} \end{pmatrix}$$

donc

$$T_aM = \{(h,k,-(hx+ky)/z : h,k \in \mathbb{R}\}.$$

On peut remarquer que

$$xh + yk - (hx+ky) = \langle (x,y,z),(h,k,-(hx+ky)/z\rangle = 0$$

pour tout h,k, de sorte que T_aM est l'orthogonal de (x,y,z). C'est un fait général, et c'est vrai puisque $(x,y,z) = \frac{1}{2}\nabla F$, où $F(x,y,z) = x^2+y^2+z^2-R^2$ est la fonction telle que $M = F^{-1}(0)$ (restreint sur $U = \{(x,y,z) : z > 0\}$).

En général donc : par le théorème des fonctions implicites,

$$d\varphi(x_0) = -\big(\partial_y F(a)\big)^{-1}\partial_x F(a),$$

et pour $h \in \mathbb{R}^p$ et $k = 1,\dots,m$ on a

$$\langle (h, d\varphi(a)[h]), \nabla F_k(a)\rangle = \sum_{j=1}^{p} h_j \frac{\partial F_k}{\partial x_j}(a) + \sum_{j=1}^{m} (d\varphi(a)[h])_j \frac{\partial F_k}{\partial y_j}(a);$$

c'est la ligne k de

$$dF(a)[h, d\varphi(a)[h]] = \partial_x F(a)[h] - \partial_y F(a)\big(\partial_y F(a)\big)^{-1}\partial_x F(a)[h] = 0.$$

Tout vecteur de T_aM est donc orthogonal à chaque $\nabla F_k(a)$, $k = 1,\dots,m$. Comme les dimensions de ces espaces sont p et $m = n-p$ respectivement, nous avons l'égalité

$$T_aM = \langle \nabla F_k(a),\ k = 1,\dots,m\rangle,$$

qui donne une description alternative de l'espace tangent.

Exemple 1.7.2. *Soit $M = SL_n(\mathbb{R}) = \{A \in M_n(\mathbb{R}) \simeq \mathbb{R}^{n^2} : \det A = 1\}$. On a $F(A) = \det A - 1$, $m = 1$, et*

$$\frac{\partial F}{\partial A_{ij}}(A) = C_{ij}$$

(les cofacteurs de A). En $A = I$, on a $C_{ij} = \delta_{ij}$, donc

$$T_IM = \{H \in M_n(\mathbb{R}) : \sum_{i=1}^{n} H_{ii} = 0\} = \{H \in M_n(\mathbb{R}) : \mathrm{Tr} H = 0\}.$$

On note habituellement cet espace $\mathfrak{sl}_n(\mathbb{R})$.

Chapitre 2

Intégration

2.1 Rapports

228 : Continuité, dérivabilité des fonctions réelles d'une variable réelle. Exemples et applications. Au delà des définitions et premiers théorèmes, le programme offre de nombreuses pistes aux candidates et candidats pour élaborer leur plan : recherche d'extrema, utilisations de la continuité uniforme, fonctions convexes et leur régularité, approximation par des fonctions régulières, utilisations des formules de Taylor, liens entre caractère C^∞ et analycité, etc.

Des exemples explicites de fonctions continues et nulle part dérivables, de fonctions continues et croissantes à dérivée nulle presque partout, de fonctions C^∞ à dérivées en un point prescrites, etc. sont les bienvenus dans cette leçon.

Les candidates et candidats solides peuvent s'intéresser à la dérivabilité des fonctions monotones ou lipschitziennes ou à celle de l'intégrale indéfinie d'une fonction intégrable, proposer diverses applications du théorème de Baire (continuité d'une limite simple de fonctions continues, points de continuité d'une dérivée, généricité des fonctions nulle part dérivables parmi les fonctions continues ou des fonctions nulle part analytiques parmi les fonctions C^∞, etc.)

229 : Fonctions monotones. Fonctions convexes. Exemples et applications. Les définitions et premières propriétés liées à ces notions doivent bien sûr être présentées pour pouvoir aborder les questions de limites et de continuité de ces fonctions et leurs caractérisations à l'aide de leurs dérivées. Il convient d'illustrer son exposé par de nombreux dessins.

La convexité est une source inépuisable d'inégalités, dans divers domaines y compris les probabilités. Dans ce même domaine, l'étude des fonctions de répartition de variables aléatoires réelles, fonctions croissantes s'il en est, est une piste intéressante.

Au delà de la dimension 1, les fonctions convexes définies sur une partie convexe de $\mathbb{R}^n$ font partie de cette leçon. La recherche de leurs extrema constitue une thèmatique riche d'exemples. Les candidates et candidats solides peuvent s'intéresser à des questions de dérivabilité des fonctions monotones, ou de continuité des fonctions convexes définies sur un ouvert convexe de $\mathbb{R}^n$.

2.2 Limites inférieures et supérieures

Définition 2.2.1. *Soit $I \subset \mathbb{R}$ un intervalle, $f : I \to \mathbb{R}$. Si x_0 est contenu dans I avec un voisinage droit, on dit que*

$$\overline{\lim_{x \to x_0+}} f(x) = \ell$$

si

$$\forall \varepsilon > 0 \; \exists \delta > 0 : \; \forall x \in]x_0, x_0 + \delta[\cap I \;\; f(x) < \ell + \varepsilon;$$

$$\forall \varepsilon > 0 \; \forall \delta > 0 \; \exists x \in]x_0, x_0 + \delta[\cap I : \;\; f(x) > \ell - \varepsilon.$$

On définit de façon similaire la limite inférieure à droite et les limites à gauche.

Exercice 2.2.2. *Montrer que*

$$\overline{\lim_{x \to x_0+}} f(x) = \inf_{\delta > 0} \sup_{x_0 < x < x_0 + \delta} f(x).$$

Exercice 2.2.3. *Montrer qu'on a toujours* $\underline{\lim}_{x \to x_0+} f(x) \leqslant \overline{\lim}_{x \to x_0+} f(x)$.

Exercice 2.2.4. *Soit*

$$f(x) = \begin{cases} \sin(\frac{1}{x}), & x \neq 0 \\ 0, & x = 0. \end{cases}$$

Trouver les quatre limites de f en 0.

Exercice 2.2.5. *Montrer que f admet une limite en x_0 si et seulement si*

$$\overline{\lim_{x \to x_0+}} f(x) = \underline{\lim_{x \to x_0+}} f(x) = \overline{\lim_{x \to x_0-}} f(x) = \underline{\lim_{x \to x_0-}} f(x).$$

Définition 2.2.6. *Soit $I \subset \mathbb{R}$ un intervalle, $f : I \to \mathbb{R}$. Si x_0 est contenu dans I avec un voisinage droit, on définit la dérivée supérieure de f à droite :*

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

De façon similaire, sont définies la dérivée inférieure à droite $\underline{f}'_d(x_0)$ et les dérivées à gauche.

Exercice 2.2.7. *Montrer que f admet une dérivée en x_0 si et seulement si*

$$\bar{f}'_d(x_0) = \underline{f}'_d(x_0) = \bar{f}'_g(x_0) = \underline{f}'_g(x_0).$$

Exercice 2.2.8. *Soit $g(x) = -f(-x)$. Montrer que $\underline{g}'_g(x_0) = \underline{f}'_d(x_0)$ et $\bar{g}'_d(x_0) = \bar{f}'_g(x_0)$.*

2.3 Fonctions monotones

Exercice 2.3.1. *Soit f croissante sur $]a,b[$, $x_0 \in]a,b[$. Montrer que f admet des limites à droite et à gauche, et*

$$\lim_{x\to x_0+} f(x) = \inf_{x>x_0} f(x), \qquad \lim_{x\to x_0-} f(x) = \sup_{x<x_0} f(x).$$

Solution : soit $\ell_+ = \inf_{x>x_0} f(x)$. Pour tout $\varepsilon > 0$ il existe $x_1 > x_0$ tel que $f(x_1) < \ell_+ + \varepsilon$. Mais alors $f(x) < \ell_+ + \varepsilon$ pour tout $x \in]x_0, x_1[$; et au même temps, évidemment $f(x) \geqslant \ell_+$. On a montré alors que $\ell_+ = \lim_{x\to x_0+} f(x)$. Pour la limite à gauche la preuve est similaire.

Un fonction monotone f est donc continue en x_0 ssi ses limites à droite et à gauche sont égales. Un point de discontinuité est dit un saut de f.

Exercice 2.3.2. *Soit f croissante sur $[a,b]$, et soit S l'ensemble de ses points de saut. Montrer que S est dénombrable et*

$$\sum_{s\in S} [f(s+) - f(s-)] \leqslant f(b) - f(a).$$

Soit f continue sur $[a,b]$. Disons que $x \in [a,b]$ est *invisible à droite* s'il existe $y \in]x,b]$ tel que $f(x) < f(y)$.

Lemme 2.3.3 (de Riesz)**.** *Soit f continue sur $[a,b]$. L'ensemble U des points invisible à droite est alors ouvert, et en le représentant $U = \cup_n]a_n, b_n[$ en réunion d'intervalles disjoints, on a $f(a_n) \leqslant f(b_n)$ pour tout n.*

Démonstration. Si x est invisible à droite, alors par continuité tout point assez proche de x l'est aussi, alors U est ouvert. Dans la représentation donnée ci-dessus, supposons que pour un certain n, on a $f(a_n) > f(b_n)$. Il existe alors $x \in]a_n, b_n[$ tel que $f(x) > f(b_n)$; soit $x_0 = \sup\{y \in [x, b_k] : f(y) = f(x)\}$. On a évidemment $x_0 \in]a_n, b_n[$.

Comme $x_0 \in U$, il existe $y \in]x_0, b]$ tel que $f(x_0) < f(y)$, et on ne peut pas avoir $y < b_n$ car sinon il existerait $z \in [y, b_n]$ avec $f(z) = f(x_0)$. Mais alors $f(b_n) < f(y)$ et $y \in]b_n, b]$, ce qui implique $b_n \in U$; cela contredit l'hypothèse que b_n est une borne de décomposition de U en composantes disjointes. □

De manière identique, on montre que l'ensemble des points invisibles à gauche vérifie les conclusions symétriques, avec $f(a_n) \geqslant f(b_n)$.

Lemme 2.3.4. *Soit I un intervalle ouvert borné, et soit $A \subset I$ tel que avec un certain $\rho \in]0,1[$, pour tout $[\alpha, \beta] \subset I$ on a*

$$\lambda(A \cap]\alpha, \beta[) < \rho(\beta - \alpha).$$

Alors $\lambda(A) = 0$.

Démonstration. La mesure de Lebesgue est régulière, alors pour tout $\varepsilon > 0$ il existe un ouvert $U \subset I$ tel que $A \subset U$ et $\lambda(U) < \lambda(A) + \varepsilon$. On décompose $U = \cup_n]a_n, b_n[$ en réunion disjointe. Par l'hypothèse, on a $\lambda(A \cap]a_n, b_n[) < \rho(b_n - a_n)$ pour tout n, d'où

$$\lambda(A) \leqslant \sum_n \lambda(A \cap]a_n, b_n[) < \rho \sum_n (b_n - a_n) = \rho\lambda(U),$$

donc $\lambda(U) < \rho\lambda(U) + \varepsilon$ d'òu $\lambda(U) < \varepsilon/(1-\rho)$. On en déduit immédiatement $\lambda(A) = 0$.

□

Théorème 2.3.5 (Lebesgue)**.** *Toute fonction monotone sur un intervalle est dérivable presque partout.*

Démonstration. Introduisons les notations : $I \subset \mathbb{R}$ un intervalle, $f : I \to \mathbb{R}$. Sans perte de généralité nous pouvons supposer f croissante et $I = [a, b]$, $a < b$. Nous allons montrer que presque partout

$$(1)\ \bar{f}'_d(x) < \infty, \qquad (2)\ \underline{f}'_g(x) \geqslant \bar{f}'_d(x). \tag{2.1}$$

1 Soit $I_{sup,d} = \{x \in I : \bar{f}'_d(x) = \infty\}$. Supposons que $x_0 \in I_{sup,d}$. Alors

$$\inf_{\delta > 0} \sup_{x_0 < x < x_0 + \delta} \frac{f(x) - f(x_0)}{x - x_0} = +\infty,$$

et pour tout $C > 0$ il existe $x > x_0$ dans I tel que $\dfrac{f(x) - f(x_0)}{x - x_0} > C$, d'où $f(x) - Cx > f(x_0) - Cx_0$. Le point x_0 est donc invisible à droite pour la fonction $g(x) = f(x) - Cx$. Par le lemme de Riesz, l'ensemble U de tels points est ouvert et se décompose en réunion disjointe $U = \cup_n]a_n, b_n[$ telle que $g(a_n) \leqslant g(b_n)$ pour tout n. On a donc $f(a_n) - Ca_n \leqslant f(b_n) - Cb_n$ et

$$\sum_n (b_n - a_n) \leqslant \sum_n \frac{f(b_n) - f(a_n)}{C} \leqslant \frac{f(b) - f(a)}{C}.$$

La mesure de $I_{sup,d} \subset U$ est donc majorée par $\frac{f(b)-f(a)}{C}$, avec $C > 0$ arbitraire ; cela montre que sa mesure est nulle.

(2) Notons $I_{ig<sd} = \{x : \underline{f}'_g(x) < \bar{f}'_d(x)\}$. Si $x_0 \in I_{ig<sd}$, alors il existent C_1, C_2 tels que

$$\underline{f}'_g(x_0) < C_1 < C_2 < \bar{f}'_d(x_0). \tag{2.2}$$

On appliquera le lemme de Riesz deux fois sur des sous-intervalles.

Soit $[\alpha, \beta] \subset I$, et supposons que $x_0 \in I_{ig<sd} \cap]\alpha, \beta[$. Par (2.2), on trouve $x < x_0$ tel que $x > \alpha$ et

$$\frac{f(x) - f(x_0)}{x - x_0} < C_1,$$

d'où $f(x) - f(x_0) > C_1(x - x_0)$. Ceci implique que x_0 est invisible à gauche pour $g(x) = f(x) - C_1 x$. Par le lemme de Riesz (sur $]\alpha, \beta[$), on décompose l'ensemble de tels points $U_g = \cup_n]a_n, b_n[$, en obtenant $g(a_n) = f(a_n) - C_1 a_n \geqslant g(b_n) = f(b_n) - C_1 b_n$ donc

$$\sum_n (b_n - a_n) \geqslant \sum_n \frac{f(b_n) - f(a_n)}{C_1}.$$

Revenons à $I_{ig<sd}$. Pour $x_0 \in I_{ig<sd} \cap]\alpha, \beta[\subset U_g$, soit n tel que $x_0 \in]a_n, b_n[$. Par (2.2), on trouve $x > x_0$ qu'on peut choisir dans $]a_n, b_n[$ tel que

$$\frac{f(x) - f(x_0)}{x - x_0} > C_2,$$

donc $f(x) - f(x_0) > C_2(x - x_0)$ et x_0 est invisible à droite pour $h(x) = f(x) - C_2 x$, sur l'intervalle $]a_n, b_n[$. On applique le lemme de Riesz comme dans la partie (1), en obtenant

$$\sum_k (b_{kn} - a_{kn}) \leqslant \sum_k \frac{f(b_{kn}) - f(a_{kn})}{C_2} \leqslant \frac{f(b_n) - f(a_n)}{C_2}.$$

Mais alors, en notant $\rho = \frac{C_1}{C_2}$,

$$\lambda\big(I_{ig<sd} \cap]\alpha, \beta[\big) \leqslant \sum_n \frac{f(b_n) - f(a_n)}{C_2} \leqslant \frac{C_1}{C_2} \sum_n (b_n - a_n) \leqslant \rho(\beta - \alpha).$$

Lemme 2.3.4 implique maintenant que $\lambda(I_{ig<sd}) = 0$.

(3) En appliquant le résultat obtenu (2.1) à $g(x) = -f(-x)$, nous obtenons que presque partout

$$\underline{g}'_g(x) = \underline{f}'_d(x) \geqslant \bar{g}'_d(x) = \bar{f}'_g(x),$$

donc presque partout, avec (2.1),

$$\underline{f}'_g(x) \leqslant \bar{f}'_g(x) \leqslant \underline{f}'_d(x) \leqslant \bar{f}'_d(x) \leqslant \underline{f}'_g(x),$$

ce qui montre que ces quatre nombres sont égaux et f est dérivable en x. □

La dérivée peut parfois surprendre :

Exercice 2.3.6. *Montrer que l'escalier de Cantor a presque partout la dérivée nulle.*

Corollaire 2.3.7. *Soit f Lebesgue intégrable sur $[a,b]$, et soit $F(x) = \int_a^x f(t)dt$ son intégrale indéfinie. Alors F est presque partout dérivable.*

2.4 La dérivée de l'intégrale

Soit f intégrable sur $[a,b]$. On peut définir son intégrale indéfinie qui est une fonction F sur $[a,b]$:

$$F(x) = \int_a^x f d\lambda.$$

On ne peut pas s'attendre à ce que F soit dérivable partout : il suffit de considérer $[a,b] = [-1,1]$ et

$$f(x) = \begin{cases} 1, & x \geqslant 0 \\ 0, & x < 0. \end{cases}$$

Mais on verra que F est dérivable *presque partout*.

Il est clair que F est la différence de deux fonctions monotones :

$$F(x) = \int_a^x f_+ d\lambda - \int_a^x f_- d\lambda.$$

Théorème 2.4.1. *Soit f intégrable sur $[a,b]$, et soit F son intégrale indéfinie. Alors F' existe presque partout et $F' \underset{p.p.}{=} f$.*

Démonstration. L'existence presque partout suit du théorème de Lebesgue, et le reste de la preuve est pour montrer l'égalité $F' \underset{p.p.}{=} f$. On va démontrer d'abord que $f \underset{p.p.}{\geqslant} F'$.

Si x est tel que $f(x) < F'(x)$, alors ils existent $q, r \in \mathbb{Q}$ tels que $f(x) < q < r < F'(x)$. Soit

$$E_{q,r} = \{x : f(x) < q < r < F'(x)\}.$$

On vérifie qu'en fait, l'ensemble $E = \{x : f(x) < F'(x)\}$ est la réunion

$$E = \bigcup_{q,r \in \mathbb{Q}} E_{q,r}.$$

Le but est de montrer que $\lambda(E_{q,r}) = 0$ pour tout q, r ; comme la réunion est dénombrable, il en suivra $\lambda(E) = 0$ et donc $f \underset{p.p.}{\geqslant} F'$.

Avec q, r fixes, on peut ajouter une constante à f — ce qui ajoute la même constante à F' — et admettre $q \geqslant 0$.

Fixons $\varepsilon > 0$. Par la continuité absolue, il existe $\delta > 0$ tel que $|\int_I f d\lambda| < \varepsilon$ pour tout intervalle I de longueur $|I| = \lambda(I) < \delta$. On peut choisir $\delta < \varepsilon/r$. Par la construction de la mesure de Lebesgue (à rappeler, la mesure extérieure est la borne inférieure par des recouvrements par des intervalles, et ceux-là on peut choisir ouverts) il existe un ouvert V tel que $E_{qr} \subset V$ et $\lambda(V) < \lambda(E_{qr}) + \delta$.

Comme $F'(x) > r$ pour $x \in E_{qr}$, il existe un voisinage assez petit $]a_x, b_x[$ de x (qu'on peut supposer contenu dans V) tel que dans ce voisinage,

$$\frac{F(y) - F(x)}{y - x} > r. \tag{2.3}$$

Si on note $G(y) = F(y) - ry$, on remarque, en transformant (2.3), que

$$G(y) > G(x) \text{ si } y > x \text{ et } G(y) < G(x) \text{ si } y < x, \tag{2.4}$$

pour $y \in]a_x, b_x[$.

Le point x est donc invisible à droite pour G. Soit U l'ensemble des points invisibles à droite pour G, et soit $U = \bigcup_n]\alpha_n, \beta_n[$ sa décomposition en réunion dénombrable d'intervalles ouverts disjoints. Par le lemme de Riesz, on a $G(\alpha_n) \leqslant G(\beta_n)$ pour tout n, et on rappelle que $E_{qr} \subset U$.

L'inégalité obtenue $G(\alpha_n) \leqslant G(\beta_n)$
est équivalente à

$$F(\beta_n) - F(\alpha_n) \geqslant r(\beta_n - \alpha_n)$$

pour tout n.

On aura alors :

$$\int_U f d\lambda = \sum_{k=0}^{\infty} \int_{\alpha_k}^{\beta_k} f d\lambda = \sum_{k=0}^{\infty} [F(\beta_k) - F(\alpha_k)] \geqslant r \sum_{k=0}^{\infty} (\beta_k - \alpha_k) = r\lambda(U),$$

et au même temps

$$\int_U f d\lambda = \int_{E_{qr}} f d\lambda + \int_{U \setminus E_{qr}} f d\lambda < q\lambda(E_{qr}) + \varepsilon$$

car $\lambda(U \setminus E_{qr}) \leqslant \lambda(V \setminus E_{qr}) < \delta$. On obtient $q\lambda(E_{qr}) + \varepsilon > r\lambda(U) \geqslant r\lambda(E_{qr})$, et

$$\lambda(E_{qr}) < \frac{\varepsilon}{r - q}.$$

En choisissant ε arbitrairement petit, on montre alors que $\lambda(E_{qr}) = 0$. Cela implique $f \underset{p.p.}{\geqslant} F'(x)$.

En passant à $-f$ et donc à $-F$, on montre que $-f \underset{p.p.}{\geqslant} -F'(x)$ et $f \underset{p.p.}{\leqslant} F'(x)$, ce qui termine la preuve. □

(fin du cours 5)

2.5 Continuité absolue et l'intégrale indéfinie

Théorème 2.5.1 (Continuité absolue de l'intégrale)**.** *Soit f intégrable sur X. Alors pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que si $\mu(A) < \delta$, alors*

$$\left|\int_A f d\mu\right| \leqslant \int_A |f| d\mu < \varepsilon.$$

Démonstration. En utilisant l'inégalité $\left|\int f d\mu\right| \leqslant \int |f| d\mu$, on peut supposer f positive. Selon la définition de l'intégrale, il existe une fonction étagée g telle que $0 \leqslant g \leqslant f$ et

$$\int_X g d\mu > \int_X f d\mu - \frac{\varepsilon}{2},$$

alors

$$0 \leqslant \int_X (f-g) d\mu < \frac{\varepsilon}{2}.$$

On peut supposer f non-nulle (sinon l'assertion est triviale), alors on pourra toujours choisir g non-nulle. Comme g est bornée, pour tout $A \in \mathcal{F}$

$$\int_A g d\mu \leqslant \mu(A) \|g\|_\infty \neq 0.$$

On pose $\delta = \dfrac{\varepsilon}{2\|g\|_\infty}$, alors pour tout ensemble A de mesure $\mu(A) < \delta$ on obtient

$$0 \leqslant \int_A f d\mu = \int_A g d\mu + \int_A (f-g) d\mu < \mu(A)\|g\|_\infty + \frac{\varepsilon}{2} < \varepsilon.$$

□

Définition 2.5.2. *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré. Une autre mesure ν sur $(X, \mathcal{F})$ est dite absolument continue par rapport à μ si $\mu(A) = 0$ implique $\nu(A) = 0$. Notation : $\nu \ll \mu$.*

Exercice 2.5.3. *Soit μ et ν des mesures finies sur $(X, \mathcal{F})$ telles que $\nu \ll \mu$. Montrer que pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que $\mu(A) < \delta$ implique $\nu(A) < \varepsilon$.*

[On suppose le contraire et on choisit $A_n \in \mathcal{F}$ tels que $\mu(A_n) < 2^{-n}$, $\nu(A_n) \geqslant \varepsilon_0 > 0$ pour tout n. En posant $B_n = \cup_{k \geqslant n} A_k$, on obtient une suite décroissante (B_n) de mesure $\mu(B_n) \leqslant 2^{1-n}$, tant que $\nu(B_n) \geqslant \nu(A_n) \geqslant \varepsilon_0$. Pour $B = \cap_n B_n$ on aura alors $\mu(B) = 0$ mais $\nu(B) \geqslant \varepsilon_0$ par continuité, ce qui contredit la continuité absolue.]

Ce résultat n'est pas vrai si les mesures ne sont pas finies : sur $X = \mathbb{R}_+$, considérer $\nu(A) = \int_A \frac{1}{x} dx$ (par rapport à la mesure de Lebesgue).

Remarque 2.5.4. *Toute mesure à densité est évidemment absolument continue par rapport à μ. On peut montrer la reciproque : si ν est absolument continue,* *alors il existe f telle que $\nu = \mu_f$ est une mesure à densité f. Ce résultat s'appelle le théorème de Radon-Nikodym (donné plus tard).*

Cette notion est liée à la classe suivante de fonctions :

Définition 2.5.5. *On dit que $F : [a, b] \to \mathbb{R}$ est absolument continue si pour tout $\varepsilon > 0$ il existe $\delta > 0$ tel que pour toute famille d'intervalles disjoints $\sqcup_k (a_k, b_k) \subset [a, b]$*

$$\sum_k (b_k - a_k) < \delta \quad \textit{implique} \quad \sum_k |F(b_k) - F(a_k)| < \varepsilon.$$

Remarque 2.5.6. *Toute fonction absolument continue est continue en tout point $x \in [a, b]$: si ε, δ sont comme dans la définition et $|y - x| < \delta$, alors $|F(y) - F(x)| < \varepsilon$.*

[Toute fonction absolument continue est à variation bornée.]

Proposition 2.5.7. *Si F est l'intégrale indéfinie d'une fonction f intégrable sur $[a, b]$, alors F est absolument continue.*

Démonstration. Suit directement du Théorème 2.5.1 : en posant $I = \sqcup_k (a_k, b_k)$, on a $\lambda(I) = \sum_k (b_k - a_k) < \delta$, alors

$$\begin{aligned}
\sum_k |F(b_k) - F(a_k)| &= \sum_k \Big| \int_a^{b_k} f d\lambda - \int_a^{a_k} f d\lambda \Big| \\
&= \sum_k \Big| \int_{[a_k, b_k]} f d\lambda \Big| \leqslant \sum_k \int_{[a_k, b_k]} |f| d\lambda = \int_I |f| d\lambda < \varepsilon.
\end{aligned}$$

□

Exercice 2.5.8. *Montrer que toute fonction lipschitzienne est absolument continue.*

Exercice 2.5.9. *Montrer que la fonction hölderienne x^α, $0 < \alpha$, est absolument continue sur $[0, 1]$.*

Exercice 2.5.10. *Montrer que l'escalier de Cantor (a) n'est pas absolument continu ; (b) vérifie la condition de Höldér avec $\alpha = \dfrac{\ln 2}{\ln 3}$.*

2.6 Théorème fondamental de l'analyse

On appelle ainsi la formule de Newton–Leibnitz :

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

Il faut préciser que même dans le sens de Lebesgue, l'égalité n'est pas toujours vérifié : pour l'«escalier de Cantor» F où on a $F(1) - F(0) = 1$ mais $F' \underset{p.p.}{=} 0$. On verra que la formule est vraie pour les fonctions absolument continues ; et que c'est exactement la classe des intégrales indéfinies de fonctions intégrables.

Théorème 2.6.1. *Si F est absolument continue sur $[a, b]$, alors elle est presque partout dérivable et*

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

Suivra directement du théorème de Radon-Nikodym.

Pour le moment, on démontre le cas particulier suivant :

Théorème 2.6.2. *Si F est continue et dérivable sur $[a, b]$, avec la dérivée F' bornée sur $[a, b]$, alors F' est intégrable et*

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

Démonstration. On peut prolonger F sur $[b, b+1]$ en posant $F(x + b) = F(b) + F'(b)x$. Pour $n \in \mathbb{N}$, $n > 0$, on peut alors définir sur $[a, b]$ les fonctions

$$F_n(x) = \frac{F(x + \frac{1}{n}) - F(x)}{1/n} = n\big[F(x + \frac{1}{n}) - F(x)\big],$$

$x \in [a, b]$. Par l'hypothèse, $F_n \to F'$ sur $[a, b]$. Il en suit déjà que F' est mesurable, et en étant bornée elle est intégrable. Par le théorème d'accroissement finis pour chaque $x \in [a, b]$ et chaque n il existe $y \in [x, x + \frac{1}{n}]$ tel que $F_n(x) = F'(y)$, donc

$$|F_n(x)| \leqslant C = \sup\{|F'(x)| : x \in [a, b]\}.$$

Par le théorème de convergence dominée,

$$\int_a^b F' d\lambda = \lim_{n\to\infty} \int_a^b F_n d\lambda = \lim_{n\to\infty} \int_a^b n\big[F(x + \frac{1}{n}) - F(x)\big]d\lambda(x) =$$
$$= \lim_{n\to\infty} n\Big[\int_b^{b+1/n} F d\lambda - \int_a^{a+1/n} F d\lambda\Big]$$

(On peut faire le changement de variable $x \mapsto x - 1/n$ car les fonctions sous les intégrales sont continues, alors on peut calculer les intégrales au sens de Riemann.) Encore par le théorème d'accroissement finis mais appliqué à l'intégrale indéfinie de F, ces dernières intégrales sont égales à $F(b_n)/n$ et $F(a_n)/n$ respectivement, où $b_n \in [b, b + 1/n]$ et $a_n \in [a, a + 1/n]$. On conclut que

$$\int_a^b F' d\lambda = \lim_{n\to\infty} \big[F(b_n) - F(a_n)\big] = F(b) - F(a).$$

□

2.7 Décomposition de Hahn

Définition 2.7.1. *Soit $(X, \mathcal{F})$ un espace mesurable. On appelle une mesure à valeurs réelles toute application $\mu : \mathcal{F} \to \mathbb{R}$ σ-additive telle que $\mu(\emptyset) = 0$.*

On dit que $A \in \mathcal{F}$ est positif pour μ si $\mu(A \cap B) \geqslant 0$ pour tout $B \in \mathcal{F}$; autrement dit, si $\mu|_A$ est une mesure positive. De la même façon on définit des parties négatives.

Exercice 2.7.2. *Montrer qu'on a toujours*

$$\sup\{|\mu(A)| : \ A \in \mathcal{F}\} < \infty.$$

Théorème 2.7.3 (Décomposition de Hahn)**.** *Soit $(X, \mathcal{F})$ un espace mesurable et μ une mesure à valeurs réelles sur $(X, \mathcal{F})$. Il existe alors $A_+ \in \mathcal{F}$ tel que A_+ est positif pour μ et $A_- = X \setminus A_+$ est négatif.*

Démonstration. L'ensemble vide est considéré négatif pour μ ; soit

$$m = \inf\{\, \mu(A) : \ A \text{ est négatif pour } \mu\}.$$

On choisit une suite d'ensembles $(A_n) \subset \mathcal{F}$ telle que $\mu(A_n) \to m$, $n \to \infty$, et on pose pour chaque n

$$A'_n = \cup_{k \leqslant n} A_k.$$

On vérifie que A'_n est négatif, et alors

$$m \leqslant \mu(A'_n) = \mu(A_n) + \mu(A'_n \setminus A_n) \leqslant \mu(A_n),$$

d'où $\mu(A'_n) \to m$. On pose $A = \cup_n A'_n = \cup_n A_n$. On a alors

$$\mu(A_-) = \mu\big(A'_0 \cup \cup_n (A'_{n+1} \setminus A'_n)\big) = \mu(A'_0) + \sum_n \mu(A'_{n+1} \setminus A'_n) = \lim_n \mu(A'_n) = m.$$

L'ensemble A_- est négatif : pour $B \in \mathcal{F}$

$$\mu(A_- \cap B) = \mu(A'_0 \cap B) + \sum_n \mu(A'_{n+1} \cap (B \setminus A'_n)) \leqslant 0.$$

Montrons maintenant que $A_+ = X \setminus A_-$ est positif.

Si ce n'est pas le cas, alors il existe $B \subset A_+$ tel que $\mu(B) < 0$. Si l'ensemble B était négatif, alors $C = A_- \cup B$ le serait aussi, avec $\mu(C) < m$, ce qui est impossible. Il en suit que B n'est pas négatif. Pour les mêmes raisons, aucune de ses parties à mesure négative n'est pas négative.

Comme $0 < M = \sup\{\mu(C) : C \subset B\} < \infty$, il existe $B_1 \subset B$ tel que $\mu(B_1) > M - \varepsilon_1$, où $\varepsilon_1 = \min(1, M/2)$. On définit ensuite $M_1 = \sup\{\mu(C) :$

$C \subset B \setminus B_1\} < \infty$ et on choisit $B_2 \subset B \setminus B_1$ tel que $\mu(B_2) > M_1 - \varepsilon_2$, avec $\varepsilon_2 = \min(1/2, M_1/2) > 0$. En continuant ainsi, on obtient une suite (B_n) de parties disjointes de B, telles que pour tout n, on a $\mu(B_{n+1}) > M_n - \varepsilon_n$, où

$$M_n = \sup\{\mu(A) : A \subset B \setminus \cup_{k \leqslant n} B_k\} > 0$$

et $\varepsilon_n = \min(1/2^n, M_n/2)$. On pose $C = B \setminus \cup_n B_n$. Pour tout $A \subset C$ on doit avoir $\mu(A) \leqslant 0$, sinon il existerait $n \in \mathbb{N}$ tel que $\mu(A) > 1/2^n$; mais alors $A \cap B_{n+1} = \emptyset$, $A \cup B_{n+1} \subset B \setminus \cup_{k \leqslant n} B_k$ et

$$\mu(A \cup B_{n+1}) = \mu(A) + \mu(B_{n+1}) > \frac{1}{2^n} + M_n - \varepsilon_n \geqslant M_n,$$

ce qui est une contradiction. L'ensemble C est donc négatif ; mais dans ce cas, $C \cup A_-$ est négatif, avec

$$\mu(C \cup A_-) = m + \mu(C) = m + \mu(B) - \sum_n \mu(B_n) < m + \mu(B) < m,$$

ce qui est une contradiction encore. On conclut que A_+ est positif. □

Corollaire 2.7.4. *Soit $(X, \mathcal{F})$ un espace mesurable et μ une mesure à valeurs réelles sur $(X, \mathcal{F})$. Il existent alors deux mesures positives μ_+, μ_- sur $(X, \mathcal{F})$ telles que $\mu = \mu_+ - \mu_-$.*

Démonstration. Pour $A \in \mathcal{F}$, on pose $\mu_+(A) = \mu(A \cap A_+)$ et $\mu_-(A) = \mu(A \cap A_-)$. □

2.8 Théorème de Radon-Nikodym

Définition 2.8.1. *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré et ν une mesure sur $(X, \mathcal{F})$. On dit que ν est absolument continue par rapport à μ, noté $\nu \ll \mu$, si $\mu(A) = 0$ implique $\nu(A) = 0$.*

Théorème 2.8.2. *Soit $(X, \mathcal{F}, \mu)$ un espace mesuré et ν une mesure finie absolument continue par rapport à μ. Alors il existe $f : X \to \mathbb{R}$ mesurable positive telle que pour tout $A \in \mathcal{F}$*

$$\nu(A) = \int_A f d\mu.$$

La fonction f est unique (à l'égalité presque partout près) ; on l'appelle la dérivée de Radon-Nikodym de ν par rapport à μ.

Démonstration. Notons

$$E = \{g : X \to \mathbb{R}_+ \text{ mesurable } : \int_A g d\mu \leqslant \nu(A) \quad \forall A \in \mathcal{F}\}.$$

On a évidemment $0 \in E$ et $\int_X g d\mu \leqslant \nu(X)$ pour toute $g \in E$. Soit

$$M = \sup \Big\{ \int_X g d\mu : \ g \in E \Big\}.$$

On choisit une suite $(g_n) \subset E$ telle que $\int_X g_n d\mu \to M$. En passant à $f_n = \max(g_1, \dots, g_n)$, on obtient une suite croissante ; montrons que $f_n \in E$. Pour $A \in \mathcal{F}$ et $1 \leqslant k \leqslant n$, soit $A_k = \{x \in A : f_n = g_k\}$. Ces ensembles sont mesurables, et

$$\int_A f_n d\mu = \sum_k \int_{A_k} g_k d\mu \leqslant \sum_k \nu(A_k) = \nu(A),$$

donc on a bien $f_n \in E$. Notons $f = \sup_n f_n = \lim_n f_n$: ce sera la fonction recherchée. Par le lemme de Beppo Lévi on a pour tout $A \in \mathcal{F}$

$$\int_A f d\mu = \lim \int_A f_n d\mu \leqslant \nu(A),$$

donc $f \in E$; et en particulier, $\int_X f d\mu = M$.

Montrons maintenant que $\int_A f d\mu = \nu(A)$ pour tout $A \in \mathcal{F}$, en supposant le contraire : il existe $A_0 \in \mathcal{F}$ tel que

$$\int_{A_0} f d\mu < \nu(A_0);$$

l'inégalité opposée est impossible car $f \in E$.

La mesure $\nu' : A \mapsto \nu(A) - \int_A f d\mu$ est nonnégative, absolument continue par rapport à μ et $\nu'(A_0) > 0$.

Pour tout n, la mesure à valeurs réelles $\nu_n = \nu' - \mu/n$ admet une décomposition de Hahn ; soit A_n^+ son ensemble maximal positif, et soit $A^+ = \cup_n A_n^+$. Si $\mu(A_n^+) = 0$ pour tout n, alors $\mu(A^+) = 0$ d'où $\nu'(A^+) = 0$; dans ce cas on aurait $\nu'(A_0 \cap A^+) \leqslant \nu'(A^+) = 0$ et $\nu_n(A_0 \setminus A^+) \leqslant 0$, donc

$$\nu'(A_0 \setminus A^+) \leqslant \frac{1}{n} \mu(A_0 \setminus A^+)$$

pour tout n, d'où $\nu'(A_0 \setminus A^+) = 0$ et $\nu'(A_0) = 0$ contrairement au choix de A_0.

Il existe donc n tel que $\mu(A_n^+) > 0$. Posons $h = f + \frac{1}{n} I_{A_n^+}$. Pour tout $A \in \mathcal{F}$ on a

$$0 \leqslant \nu_n(A \cap A_n^+) = \nu'(A \cap A_n^+) - \frac{1}{n}\mu(A \cap A_n^+),$$

donc

$$\begin{aligned} \int_A h d\mu = \int_A f d\mu + \frac{1}{n}\mu(A \cap A_n^+) &\leqslant \int_A f d\mu + \nu'(A \cap A_n^+) \\ &\leqslant \int_A f d\mu + \nu'(A) = \nu(A). \end{aligned}$$

La fonction h est donc dans E, mais

$$\text{ínt}_X h d\mu = \int_X f d\mu + \frac{1}{n}\mu(A_n^+) > M,$$

contrairement au choix de f.

L'existence de A_0 est donc impossible, et nous avons pour tout A l'égalité recherchée. □

Démonstration. Supposons μ finie et posons $\lambda = \mu + \nu$. On a $\nu \leqslant \lambda$ donc $f \mapsto \int_X f d\nu$ est une fonctionnelle linéaire bornée sur $L^2(X, \lambda)$:

$$\Big|\int_X f d\nu\Big| \leqslant \int_X |f| d\nu \leqslant \int_X |f| d\lambda \leqslant \lambda(X)^{1/2}\|f\|_{2,\lambda}.$$

Par le théorème de Riesz, il existe $g \in L^2(X, \lambda)$ telle que $\int f d\nu = \int f g d\lambda$ pour toute $f \in L^2(\lambda)$, donc

$$\int f(1-g) d\nu = \int f g d\mu.$$

Soit $A_0 = \{x \in X : g(x) < 0$ et $A_1 = \{x \in X : g(x) \geqslant 1\}$. En posant $f = I_{A_j}$, on obtient :

$$\nu(A_0) \leqslant \int_{A_0} (1-g) d\nu = \int_{A_0} g d\mu \leqslant 0$$

d'où $\nu(A_0) = 0$, mais alors de la même inégalité $\mu(A_0) = 0$. Ensuite,

$$0 \geqslant \int_{A_1} (1-g) d\nu = \int_{A_1} g d\mu \geqslant \mu(A_1)$$

d'où $\mu(A_1) = 0$ et alors $\nu(A_1) = 0$. Il en suit $0 \leqslant g < 1$ l-presque partout, et la changeant si nécessaire, on peut avoir $0 \leqslant g < 1$ partout. On pose $\varphi = \dfrac{g}{1-g}$ et on a pour tout A

$$\int_A \varphi d\mu = \int_X I_A \frac{g}{1-g} d\mu = \int_X I_A d\nu = \nu(A),$$

car $I_A \in L^2(\lambda)$ (toutes les mesures sont finies). □

2.9 La dualité des L^p

On suppose fixé un espace mesuré $(X, \mathcal{F}, \mu)$. On fixe ensuite $p, q \in [1, +\infty]$ conjugués : $\dfrac{1}{p} + \dfrac{1}{q} = 1$.

Rappelons que si $\mu(A) < \infty$, alors $L^p(A, \mu) \subset L^1(A, \mu)$.

Lemme 2.9.1. *Pour $g \in L^q(X)$, soit $\varphi_g : L^p(X) \to \mathbb{C}$ défini par*

$$\varphi_g(f) = \int_X f g d\mu.$$

Alors φ_g est bien défini, linéaire, et bornée de norme $\|\varphi_g\| = \|g\|_q$.

Démonstration. Le fait que l'intégrale converge et est bornée par $\|f\|_p\|g\|_q$ est l'inégalité de Hölder qu'on suppose connue (et qui ne nécessite aucune hypothèse particulière sur μ).

Par définition de l'intégrale, il existe une suite (g_n) croissante de fonctions étagées telles que $0 \leqslant g_n \leqslant |g|^q$ et

$$\|g\|_q^q = \int_X |g|^q d\mu = \lim_{n\to\infty} \int_X g_n d\mu.$$

En posant $f_n = g_n^{1/p}$, on obtient $f_n \in L^p(X)$ de norme $\|f_n\|_p^p = \int_X g_n d\mu$, et (comme $|g| \geqslant g_n^{1/q}$)

$$\int_X f_n g d\mu \geqslant \int_X g_n^{1/q} g_n^{1/p} d\mu = \int_X g_n d\mu,$$

d'où

$$\frac{|\varphi_g(f_n)|}{\|f_n\|_p} \geqslant \Big(\int_X g_n d\mu\Big)^{1-1/p} \to \|g\|_q,\ n \to \infty,$$

ce qui montre l'égalité $\|\varphi_g\| = \|g\|_q$. □

Lemme 2.9.2. *Supposons que μ est σ-finie. Soit $g : X \to \mathbb{C}$ mesurable, et telle que pour tout $f \in L^p(\mu)$, on a $fg \in L^1(\mu)$, et il existe une constante $C > 0$ telle que pour tout $f \in L^p(\mu)$*

$$\Big|\int_X fg d\mu\Big| \leqslant C\|f\|_p.$$

Alors $g \in L^q(\mu)$ et $\|g\|_q \leqslant C$.

Démonstration. Soit $\varphi : X \to \mathbb{C}$ «l'argument de g» :

$$\varphi(x) = \begin{cases} \dfrac{|g(x)|}{g(x)}, & g(x) \neq 0, \\ 1, g(x) = 0. \end{cases}$$

C'est une fonction mesurable, $|\varphi| = 1$ et $\varphi g = |g|$. Si $f \in L^p(X)$, alors $f\varphi \in L^p(X)$ et

$$\Big|\int_X f\varphi g d\mu\Big| \leqslant C\|f\varphi\|_p = C\|f\|_p,$$

d'où l'on voit que l'hypothèse vaut aussi pour $|g|$, et on note que $g \in L^q(X)$ ssi $|g| \in L^q(X)$. On supposera alors dans la suite $g \geqslant 0$.

Supposons aussi $p > 1$. Soit $g_n \to g$ une suite croissante de fonctions étagées telle que $0 \leqslant g_n \leqslant g$. Pour tout $A \in \mathcal{F}$ de mesure finie, on pose $f_n = I_A g_n^{q-1}$, alors f_n est bornée donc dans $L^p(A)$ et, comme $p(q-1) = pq/p = q$,

$$0 \leqslant \int_A f_n g_n d\mu = \int_A g_n^q d\mu \leqslant \int_A f_n g d\mu \leqslant C\|f_n\|_p = C\Big(\int_A g_n^q d\mu\Big)^{1/p},$$

d'où $\Big(\int_A g_n^q d\mu\Big)^{1-1/p} = \|g_n I_A\|_q \leqslant C$. Il en suit tout de suite que $\|gI_A\|_q \leqslant C$, et par définition de l'intégrale, $\|g\|_q \leqslant C$.

[Rappel : μ est σ-finie, donc il existe une suite croissante (X_m) de parties de X de mesure finie telle que $X = \cup_m X_m$; par définition, $g \in L^q(X)$ si $g \in L^q(X_m)$ pour tout m, et $\|g\|_q = \sup_m \|gI_{X_m}\|_q$.]

Soit enfin $p = 1$ et $q = \infty$. On pose $A = \{x \in X : g(x) > C\}$. Si $\mu(A) > 0$, on peut trouver (encore utilisant le fait que μ est σ-finie) $B \subset A$ de mesure finie tel que $\mu(B) > 0$; pour $f = I_B$ on a par l'hypothèse

$$0 \leqslant \int_X fg d\mu = \int_B g d\mu \leqslant C\mu(B),$$

mais directement on obtient $\int_B g d\mu > C\mu(B)$. Cette contradiction montre que $\mu(A) = 0$. □

Théorème 2.9.3. *Supposons que μ est σ-finie. L'application $g \mapsto \varphi_g$, $L^q(X) \to L^p(X)^*$, est une isométrie bijective.*

Démonstration. Isométrie étant déjà démontrée, il nous faut montrer que pour $\varphi \in L^p(X)^*$ il existe $g \in L^q(X)$ telle que $\varphi = \varphi_g$.

Supposons d'abord que μ est finie. On passe par les mesures : pour $E \in \mathcal{F}$ on pose $\nu(E) = \varphi(I_E)$; c'est bien défini car μ est finie et $I_E \in L^p(X)$. On obtient une mesure à valeurs réelles (pas nécessairement positive) ν, qui est absolument continue par rapport à μ : si $\mu(E) = 0$ alors $I_E = 0$ dans $L^p(X)$, donc $\nu(E) = 0$. Elle est σ-additive : si $E = \cup_n E_n$ et les $E_n \in \mathcal{F}$ sont disjoints, alors la série $\sum_n I_{E_n}$ converge en norme $\|\cdot\|_p$: pour tout $N \in \mathbb{N}$

$$\|I_E - \sum_{n=1}^{N} I_{E_n}\|_p = \mu\Big(E \setminus \cup_{n=1}^{N} E_n\Big)^{1/p} \to 0,\ n \to \infty.$$

Il en suit

$$\nu(E) = \varphi(I_E) = \lim_{N\to\infty} \varphi\Big(\sum_{n=1}^{N} I_{E_n}\Big) = \sum_n \nu(E_n).$$

Par le théorème de Radon-Nikodym, il existe $g \in L^1(\mu)$, la densité de ν par rapport à μ, telle que $\nu(E) = \int_X g d\mu = \int_X I_E\, g\, d\mu$ pour tout $E \in \mathcal{F}$. L'égalité $\varphi(f) = \int_X fg$ étant vérifiée pour toutes les indicatrices, elle vaut par linéarité pour toutes les fonctions étagées donc par continuité sur $L^p(X)$ tout entier, et nous avons bien $\varphi = \varphi_g$. Par le lemme, $g \in L^q(\mu)$.

Soit maintenant μ σ-finie, et $X = \cup_m X_m$ où la suite (X_m) est croissante et chaque X_m de mesure finie. Pour chaque m, on peut identifier $L^p(X_m, \mu)$ avec l'espace des fonctions dans $L^p(X)$ qui s'annulent hors X_m. La restriction de φ sur cet espace est alors de forme $f \mapsto \int_{X_m} fg_m d\mu$ avec $g_m \in L^q(X_m) \subset L^q(X)$ de norme $\|g_m\|_q \leqslant \|\varphi\|$. Si $m < l$, les fonctionnelles $\varphi_{g_m} = \varphi_{g_l} = \varphi$ sur $L^p(X_m)$

sont les mêmes, donc $g_m = g_l$ presque partout. L'ensemble

$$A = \cup_{m,l}\{x \in X : g_m(x) \neq g_l(x)\}$$

est de mesure nulle, alors on peut définir la fonction $g : X \to \mathbb{C}$ en posant $g = g_m$ sur $X_m \setminus A$ et $g = 0$ sur A, pour obtenir $\varphi_g = \varphi_{g_m}$ sur chaque $L^p(X_m)$. Par construction, $g \in L^q(X)$ et $\|g\|_q \leqslant \|\varphi\|$. Enfin, si $f \in L^p(X)$, alors

$$\int_X fgd\mu = \lim_m \int_{X_m} fgd\mu = \lim_m \varphi(fI_{X_m}) = \varphi(f).$$

□

Corollaire 2.9.4. *Les espaces $L^p(X)$ sont réflexifs si $1 < p < \infty$ (et μ σ-finie).*

Exemple 2.9.5. *Sur $\ell^\infty \equiv \ell^\infty(\mathbb{N})$, il existe une fonctionnelle φ qui n'a pas la forme φ_y avec aucun $y \in \ell^1$. Sur l'espace $c \subset \ell^\infty$ des suites convergentes, on pose $\varphi(x) = \lim x_n$; on le prolonge par Hahn-Banach. Si on avait $\varphi(x) = \sum x_n y_n$ avec $y \in \ell^1$, on aurait $y_n = \varphi(e_n) = 0$ pour toute indicatrice e_n du singleton $\{n\}$; mais alors $y = 0$ tant que $\varphi \neq 0$.*

Exemple 2.9.6. *En plongeant ℓ^∞ isométriquement dans $L^\infty(\mathbb{R})$ par*

$$F(x)(t) = \begin{cases} 1, & t \in [n, n+1[,\ n \in \mathbb{N} \\ 0, & t < 0 \end{cases}$$

pour $x \in \ell^\infty$, on peut définir une fonctionnelle $\psi(f) = \varphi(F^{-1}f)$ sur $F(\ell^\infty)$, de norme 1, et la prolonger sur $L^\infty(\mathbb{R})$ par Hahn-Banach. Si ce prolongement étant l'intégrale avec une fonction $g \in L^1(\mathbb{R})$, on aurait pour tout n

$$\int_n^{n+1} g = \psi(I_{[n,n+1[}) = \varphi(e_n) = 0,$$

mais

$$\int_0^\infty g = \psi(I_{[0<+\infty[}) = \varphi(1) = 1,$$

ce qui est impossible.

Exemple 2.9.7. *Si μ n'est pas σ-finie, l'assertion reste vraie pour $1 < p < \infty$. Si $p = 1$, fixons d'abord un exemple d'une mesure non-σ-finie : $X = \mathbb{R}$, μ la mesure de comptage (qui est infinie sur toutes les parties infinies),*

$$\mathcal{F} = \{A \subset \mathbb{R} : A \text{ ou } \mathbb{R} \setminus A \text{ est dénombrable}\}.$$

Soit $B \subset \mathbb{R}$ non-dénombrable de complément $\mathbb{R}\setminus B$ non-dénombrable (par exemple, $B = [0, 1]$), et soit $\varphi \in L^1(\mathbb{R}, \mu)^*$ définie par

$$\varphi(f) = \sum_{x \in B} f(x), \quad f \in L^1(\mathbb{R}, \mu).$$

Alors φ est bien définie, linéaire, et bornée de norme 1. Pourtant, si on avait $\varphi = \varphi_g$ avec $g \in L^\infty(\mu)$, on aurait du avoir $g = I_B$, ce qui la fait non-mesurable.

Bibliographie

[1] Kolmogorov, Fomine, Eléments de la théorie de fonctions et de l'analyse fonctionnelle, Ellipses, 1994 (théorèmes de dérivation).

[2] J. Lafontaine, Introduction aux variétés différentielles, Grenoble Sciences, 2010 (variétés).

[3] G. Folland, Real Analysis, Wiley, 1999 (Radon-Nikodym, espaces L^p).