COURS DE M2, TECHNIQUES D'ANALYSE

Guy David, Université de Paris-Sud 11

Ce qui suit est une rédaction sans doute approximative, et livrée sans garantie, de mon cours de M2 à Orsay (initialement au premier semestre 2006-07, et modifié plusieurs fois ensuite, souvent par ajout sauvage d'un paragraphe). D'où parfois une certaine désorganisation dans l'ordre des chapitres et leurs numéros.

J'essaie de corriger relativement régulièrement des erreurs (le plus souvent, après le cours), mais il en restera toujours beaucoup.

Le thème général est un mélange d'analyse et de théorie géométrique de la mesure. Il y a quelques chapitres à la fin sur les fonctions holomorphes et les représentations conformes, mais c'est loin d'être le sujet principal des notes.

Une très courte bibliographie se trouve au début, mais le document essaie d'être à peu près autocontenu.

INTRODUCTION

Résumé officiel du cours l'année dernière:
Il s'agira d'une introduction certaines techniques fondamentales d'analyse classique et de théorie géométrique de la mesure. Les sujets traités devraient être parmi les suivants.

Une première partie traitera des propriétés de différentiabilité des fonctions de plusieurs variables: fonction maximale de Hardy-Littlewood, différentiation des mesures, interpolation réelle et complexe, espaces de Sobolev, inégalités de Poincaré et Sobolev, fonctions Lipschitziennes, différentiabilité presque-partout.

Je crois que la partie "Opérateurs d'intégrale singulière" n'existera pas ou peu cette année. Pas plus que la fin sur les fonctions holomorphes.

Une seconde partie abordera les bases de la théorie géométrique de la mesure: mesures de Hausdorff, changement de variables dans les intégrales, formule de la co-aire, formule de la divergence, fonctions à variation bornée, rectifiabilité.

Une courte troisième partie d'applications de la seconde pourrait exister, si le temps le permet.

Quelques ouvrages de référence que je risque de citer

L. Ambrosio, N. Fusco and D. Pallara, Functions of bounded variation and free discontinuity problems, Oxford Mathematical Monographs, Clarendon Press, Oxford 2000. (pour ce qui ressemble à BV utilisé pour le calcul des variations)
Y. Meyer, Ondelettes et Opérateurs (3 volumes), Actualités Mathématiques, Herman 1990-91. (à la fois opérateurs de Calderón-Zygmund et caractérisation des espaces fonctionnels)
Enrico Giusti, Minimal surfaces and functions of bounded variation, Monographs in Mathematics vol.80, Birkhäuser 1984. (tout sur BV)
K. Falconer, The geometry of fractal sets, Cambridge University Press 1984. (mesures, fractals, petit livre agréable)
P. Mattila, Geometry of sets and measures in Euclidean space, Cambridge Studies in Advanced Mathematics 44, Cambridge University Press 1995. (plus complet que le précédent, contient beaucoup de théorie géométrique de la mesure, sans être aussi difficile à lire que La Grande Référence).
F. Morgan, Geometric measure theory, A beginner's guide, Academic Press 1988. (pour une description rapide et vivante de très nombreux résultats de théorie géométrique de la mesure; mais souvent on ne comprend vraiment les arguments que quand on les connaît déjà. A l'air très bien aussi)
J. Heinonen, Lectures on Analysis on Metric Spaces, Universitext, Springer-Verlag, New York, 2001. x+140 pp. (pour le théorème de Rademacher-Calderón et les espaces métriques)
H. Federer, Geometric Measure Theory, Springer. (Contient presque tout, mais il faut être courageux et entrer dans les notations)

Pour la partie complexe,
W. Rudin, Real and complex analysis, third edition, McGraw-Hill Book Co., New York, 1987 (existe aussi en français).
E. M. Stein, Singular integrals and differentiability properties of functions, Princeton university press 1970.
A. Zygmund, Trigonometric series, Cambridge University Press 1968 (peut-être y a-t-il une nouvelle édition).
- W. Rudin, Real and complex Analysis, Mc Graw Hill, ou sa version française.
- Peter L. Duren, Univalent functions, Grundlehren 259, Springer
- C. Pommerenke, Boundary behaviour of conformal maps, Grundlehren 299, Springer

- Peter L. Duren, Theory of H^P spaces, Pure ane Applied Mathematics 38, Academic Press.
Newman, Elements of topology in the plane set of points, Cambridge 1964. [Pour la topologie du plan.]

Les premiers cours devraient donner une idée (un peu flatteuse?) de la suite: mélange de géométrie élémentaire et d'analyse. J'ai mis (l'an dernier) l'analyse complexe à la fin, sans autre forme de procès ni d'ailleurs renuméroter

TABLE DES MATIÈRES

1. LEMMES DE RECOUVREMENT

On va en voir deux, en commençant par le plus simple qui est d'usage très courant. Les applications viendront après.

Les deux sont très bien faits dans Mattila, et le premier est aussi dans Stein.

Notation standard: $B(x,r)$ est la boule ouverte de centre x et de rayon r. On sera dans $\mathbb{R}^n$.

Lemme (de type Vitali). *Soit $\{B_i\}_{i\in I}$ une famille de boules ouvertes dans $\mathbb{R}^n$. On pose $B_i = B(x_i, r_i)$, et on suppose que $\sup_i r_i < +\infty$. Alors il existe $J \subset I$, au plus dénombrable, tel que:*

les boules B_j, $j \in J$, sont disjointes;

$$\bigcup_{i\in I} B_i \subset \bigcup_{j\in J} B(x_j, 5r_j).$$

"Disjointes" signifie "deux à deux disjointes". Et 5 pourrait être remplacé par n'importe quoi de strictement plus grand que 3.
Contrexemple quand les rayons ne sont pas bornés: les boules $B(0,N)$ qui recouvrent $\mathbb{R}^n$.
Souvent on ne se donne pas les boules B_i directement, mais un ensemble E qu'on recouvre par des B_i et qu'on veut recouvrir de manière plus performante. Mais bien sûr cela revient au même, puisqu'alors un bon recouvrement des B_i recouvrira E aussi.
Démonstration. On pose $M = \sup_i r_i < +\infty$ et, pour $k \geq 0$, $I_k = \{i \in I\,;\, 2^{-k-1}M < r_i \leq 2^{-k}M\}$ (donc I est l'union des I_k).

On prend une partie maximale J_0 de I_0 telle que les B_j, $j \in J_0$, soient disjointes. C'est facile de voir qu'elle est au plus dénombrable, puisque chaque boule contient des rationnels. Et l'existence peut aussi se faire sans utiliser l'axiome du choix (prendre une collection maximale dans $B(0,1)$, puis dans $B(0,2)$, etc.).

Par récurrence, on prend une partie maximale J_k dans I_k, telle que les B_j, $j \in J_k$ soient disjointes deux à deux et aussi des B_i, $i \in J_0 \cup \cdots J_{k-1}$.

On prend $J = \cup_{k\geq 0} J_k$, et on va vérifier que ça marche. Les boules sont disjointes par construction, donc il suffit de voir que chaque B_i est contenue dans $5B_j$ pour un $j \in J_k$.

Fixons i. Soit k tel que $i \in I_k$. Si $i \in J_k$, pas de problème. Sinon, par maximalité, B_i rencontre B_j pour un $j \in J_0 \cup \cdots \cup J_k$. Noter que $r_i \leq 2r_j$. Alors $B_i \subset 5B_j =: B(x_j, 5r_j)$ (faire un dessin). $\square$

Commentaire: c'est un résultat facile et utile, qui se généralise facilement à des espaces métriques. C'est surtout utile s'ils sont munis d'une mesure doublante (voir peut-être plus loin), mais ça reste quand même assez général. Et on peut se permettre des formes de boules assez différentes, le point étant de pouvoir montrer que $B_i \subset CB_j$ quand B_i rencontre B_j et que $r_i \leq 2r_j$ (où r_j joue le rôle d'un rayon pour B_j).

Exercice recommandé: le lemme de presque-recouvrement de Vitali qui se trouve dans Mattila, page 24.

Il est difficile de résister à énoncer le lemme de recouvrement de Besicovitch, qui est plus délicat (et par exemple ne marcherait pas dans le groupe de Heisenberg), mais permet de s'en sortir sans doublement des boules.

Lemme (Besicovitch). *Soit $\Omega \subset \mathbb{R}^n$ un ensemble borné. Pour chaque $x \in \Omega$, on se donne une boule B_x centrée en x. Alors:*
1. Il existe une famille (au plus dénombrable) $\{x_i\}$, $i \in I$, de points de Ω telle que

$$\Omega \subset \bigcup_{i \in I} B_{x_i} \quad \text{et} \quad \sum_{i \in I} \mathbf{1}_{B_{x_i}}(z) \leq C_n \text{ pour tout } z \in \mathbb{R}^n.$$

2. On peut trouver un nombre $\leq N$ de familles $\{x_i\}$, $i \in I_k$ (et donc $1 \leq k \leq N$) telles que les boules B_{x_i}, $i \in I_k$ avec k donné, sont disjointes, et $\Omega \subset \bigcup_k \bigcup_{i \in I_k} B_{x_i}$.
Les constantes C_n et N ne dépendent que de la dimension n.

Remarques. Contrexemple si Ω n'est pas borné: prendre les boules $B(n, 2^n)$, $n \in \mathbb{N}$.
On peut remplacer "Ω est borné" par l'hypothèse que $M = \sup_x r_x < +\infty$. Voir plus bas.
On pourrait s'être donné plus d'une boule pour chaque x, et aussi certaines à l'extérieur de Ω. L'essentiel est que pour chaque $x \in \Omega$ il y ait au moins une boule centrée en x.
Les boules peuvent être ouvertes ou fermées, ça ne change rien. De même, elles pourraient être cubiques; mais tout n'est pas autorisé, il y a quand même un peu de géométrie.

Enfin, la seconde partie est plus forte que la première, mais la première suffira souvent.

Démonstration de la première partie. On se donne Ω et les B_x. On note r_x le rayon de B_x. On peut supposer que $M = \sup_x r_x < +\infty$ (sinon, une boule suffit pour recouvrir Ω qui est borné).

On pose $E_k = \{x \in \Omega;\ 2^{-k-1}M < r_x \leq 2^{-k}M\}$. On choisit $x_1, x_2, \cdots x_k$ dans E_0 tels que x_i n'appartient à aucune des boules B_{x_i} précédentes. Cela ne peut pas durer éternellement: chaque centre est à distance au moins $M/2$ des précédents, et il ne peut pas y avoir une infinité de points dans Ω borné qui soient à distances mutuelles $\geq M/2$.

Quand il n'est plus possible de trouver des points, on passe à E_1 et on cherche $x_{k+1}, \cdots, x_{k_1}$ dans E_1, tels que chacun soit hors des boules B_{x_i} précédentes. Ca ne peut durer éternellement. Quand il n'y a plus de points possibles, on recommence avec E_2, et ainsi de suite.

Notre première famille I est donc composée de tous les points x_i qu'on a trouvés ci-dessus. "Au plus dénombrable" est clair (et fini peut arriver). Le fait que les boules recouvrent Ω est aussi clair: si on n'a pas choisi x au moment de sa génération, c'est qu'il se trouvait dans une des boules précédentes.

Il reste à voir pourquoi un même point y ne peut se trouver dans plus de C_n boules B_{x_i}. On peut d'ailleurs supposer $y = 0$. Notons r_i le rayon de B_{x_i}.

Pour chaque $r > 0$ il ne peut y avoir plus de C_1 boules B_{x_i} telles que $r \leq r_i \leq 10r$, disons. En effet, par construction chaque x_i est à distance au moins r des autres centres (parce que celui des deux centres qui arrive en second dans la construction n'est pas dans la première des deux boules), et par ailleurs y est dans toutes les $B(x_i, 10r)$. Pour estimer C_1, on peut remarquer que toutes les boules $B(x_i, r/2)$ sont disjointes et contenues dans $B(y, 20r)$, comme chacune a une mesure de Lebesgue au moins $c_n(r/2)^n$ et que la mesure totale est au plus $c_n(20r)^n$, ça fait au plus 400^n boules et on peut prendre $C_1 = 40^n$.

Il suffira donc de prouver que l'on ne peut pas trouver plus de C_2 points x_i tels que $y \in B_{x_i}$ pour tout i, et aussi tels que pour $i \neq j$, ou bien $r_i > 10r_j$ ou bien $r_j > 10r_i$.

Ou encore que, si $y \in B_{x_i} \cap B_{x_j}$, et $r_i > 10r_j$, alors l'angle entre x_i et x_j est au moins 10^{-1}. [Rappel: $y = 0$.] En effet, après ça on peut dire que les distances entre les divers points $x_i/|x_i|$ sur la sphère sont au moins 10^{-2}, donc il n'y en a pas plus de C_2. [Même genre d'argument de comptage sur la sphère que pour C_1.]

On fait le dessin, et on se souvient aussi que x_j n'est pas dans B_{x_i}, alors que l'origine est dans les deux boules. Ca marche largement (et peut-être un peu moins largement dans le cas de cubes ou d'autres objets semblables). C'est même amusant, quand on voit ça, que le lemme devienne faux dans le groupe de Heisenberg. Fin de la première partie.

Maintenant on passe à la seconde assertion. On va garder la même collection I de points, et on va les répartir en au plus N familles.

On prend les x_k un par un et on les met dans des boites $A_1, \cdots A_N$; en fait dans la première boite qui contient pas encore de x_i tel que B_i rencontre B_k. C'est clairement le meilleur moyen simple d'essayer, et il faut juste démontrer qu'il n'y a besoin que de N boites, où N est une constante dimensionelle. Ou encore que chaque B_k rencontre au plus $N - 1$ boules B_i précédentes.

On fait comme ci-dessus: on se réduit aisément au cas de boules de tailles vraiment différentes; ensuite on suppose que $x_k = 0$, et on veut vérifier que les vecteurs x_i et x_j font un angle au moins 10^{-2} (dès qu'on en a deux différents). Cette fois, on a seulement que (avec des notations évidentes) $r_i \geq 10r_j \geq 100r_k$, que B_k rencontre B_i et B_j, que $x_k \notin B_j$ (donc x_j n'est pas trop près de l'origine) mais que $x_j \notin B_i$. On refait le dessin et on voit que ça marche encore (c'est comme si l'origine pouvait être juste un peu hors de B_i dans le dessin précédent). □

Remarque. Si, au lieu de savoir que Ω est borné, on sait que $M = \sup_x r_x$ est fini, la démonstration marche encore. Pour le choix des x_j, on procède encore E_k par E_k, mais éventuellement on prend une collection infinie de x_j à cette étape, obtenus en recouvrant les $E_k \cap B(0, M)$ l'un après l'autre (en complétant à chaque fois).

Exercice recommandé: Sard (version facile).

Exercice possible: la variante pour Vitali du lemme de presque-recouvrement ci-dessous, formulé avec Besicovitch.

On utilise souvent (y compris dans la suite) le lemme de presque-recouvrement suivant.

Théorème. *Soient μ une mesure de Radon (borélienne localement finie) sur $\mathbb{R}^n$, et $A \subset \mathbb{R}^n$ un ensemble borélien (ou mesurable pour μ). On se donne aussi une famille $\mathcal{C}$ de boules fermées B, $B \in \mathcal{C}$, telle que, pour tout $x \in A$, il existe des boules de $\mathcal{C}$, centrées en x, et de rayons arbitrairement petits. [On va appeler ça une famille de Vitali centrée par rapport à A.] Alors il existe*

$\mathcal{D} \subset \mathcal{C}$ *telle que*

$$(a) \qquad \text{les boules } B,\ B \in \mathcal{D}, \text{ sont disjointes,}$$

et

$$(b) \qquad \mu\Big(A \setminus \bigcup_{B \in \mathcal{D}} B\Big) = 0.$$

Noter que pour la démonstration, il est effectivement utile d'avoir des boules fermées. Mais si on aime beaucoup les boules ouvertes (comme moi), on peut souvent se débrouiller pour choisir des boules $B(x, r)$ telles que $\mu(\partial B(x, r)) = 0$, puisque ceci est vrai pour presque-tout $r > 0$ (à x fixé). [Voir ci-dessous.] Dans ce cas, (b) pour les fermetures implique (b) pour des boules ouvertes, et on sera content.

Si μ est une mesure borélienne pas nécessairement localement finie, mais A est un borélien tel que $\mu(A)$ est (localement) fini, rien n'empêche d'appliquer le résultat à la restrition à A de μ. On obtient ainsi le même résultat.

De même, on aurait pu se donner à l'avance un ouvert U contenant A, et exiger que les boules B, $B \in \mathcal{D}$, soient toutes contenues dans U. Ceci revient à remplacer $\mathcal{C}$ par la nouvelle famille de Vitali composée des boules $B \in \mathcal{C}$ qui sont contenues dans U.

Pour la démonstation, on va d'abord faire le cas un peu plus simple où $\mu(A) < +\infty$.

Si $\mu(A) = 0$, il suffit de prendre $\mathcal{D} = \emptyset$, donc on peut supposer que $0 < \mu(A) < +\infty$.

On construit des familles finies $\mathcal{D}_k$ par récurrence, et après on prendra $\mathcal{D} = \cup_{k \geq 1} \mathcal{D}_k$. L'idée sera de recouvrir tout ce qu'on peut en restant disjoint, puis de recommencer avec ce qui reste à recouvrir. En même temps, on définira des ensembles U_k et A_k.

On commence par $\mathcal{D}_0 = \emptyset$, $U_0 = \mathbb{R}^n$ et $A_0 = A$. En général, on prendra

$$(c) \qquad U_k = \mathbb{R}^n \setminus \bigcup_{l \leq k} \bigcup_{B \in \mathcal{D}_l} B,$$

qui est un ouvert parce que les boules sont fermées et que les $\mathcal{D}_k$ sont finis, et

$$(d) \qquad A_k = A \cap U_k = A \setminus \bigcup_{l \leq k} \bigcup_{B \in \mathcal{D}_l} B$$

(tout ce qu'on n'a pas encore réussi à recouvrir).

Prenons $k \geq 0$, supposons construits les $\mathcal{D}_l$, $l \leq k$ (et donc aussi U_k et A_k). Pour chaque $x \in A_k$, notre hypothèse dit qu'il existe des boules B_x de $\mathcal{C}$, centrées en x, et de rayons arbitrairement petits. On en choisit une telle que $B_x \subset U_k$ (c'est possible puisque $x \in A_k \subset U_k$ et U_k est ouvert), et dont le rayon est inférieur à 1. Ceci nous donne une sous-famille $\mathcal{C}'$ de $\mathcal{C}$.

On a choisi des rayons inférieurs à 1 pour pouvoir appliquer immédiatement le lemme de recouvrement de Besicovitch, version forte. Il existe une constante $N \geq 1$, qui ne dépend que de la dimension n, telle que dans la situation présente on puisse trouver des familles $\mathcal{C}_j$, $1 \leq j \leq N$, contenues dans $\mathcal{C}'$ avec les deux propriétés suivantes:

$$A_k \subset \bigcup_{1 \leq j \leq N} \bigcup_{B \in \mathcal{C}_{k,j}} B$$

et

$$(e) \qquad \text{Pour tout } j, \text{ les boules } B,\ B \in \mathcal{C}_j, \text{ sont disjointes.}$$

Noter que cette inclusion donne

$$\mu(A_k) \leq \sum_{1 \leq j \leq N} \mu\big(A_k \cap \bigcup_{B \in \mathcal{C}_j} B\big)$$

de sorte que par Chebyshev, ou peut choisir un indice j (celui qui donne la plus grosse contribution) tel que

$$\mu\big(A_k \cap \bigcup_{B \in \mathcal{C}_j} B\big) \geq \frac{1}{N} \mu(A_k).$$

On choisit pour $\mathcal{D}_{k+1}$ une partie finie de $\mathcal{C}_j$ suffisament grande pour que

$$(f) \qquad \mu\big(A_k \cap \bigcup_{B \in \mathcal{D}_{k+1}} B\big) \geq \frac{1}{N+1} \mu(A_k)$$

[se souvenir que comme les boules sont disjointes, il n'y en a qu'un nombre au plus dénombrable, puis utiliser la définition d'une mesure].

On note aussitôt que les nouvelles boules B, $B \in \mathcal{D}_{k+1}$ sont disjointes des précédentes (par (c) et puisqu'elles sont contenues dans U_k parce que chaque B_x l'est), et aussi entre elles (par (e)).

D'autre part,

$$\mu(A_{k+1}) = \mu\big(A_k \setminus \bigcup_{B \in \mathcal{D}_{k+1}} B\big) = \mu(A_k) - \mu\big(A_k \cap \bigcup_{B \in \mathcal{D}_{k+1}} B\big) \leq \frac{N}{N+1} \mu(A_k)$$

car (d) nous dit qu'on obtient A_{k+1} en otant de A_k la partie qui est dans les boules de $\mathcal{D}_{k+1}$, puis par (f). On itère et on trouve que $\mu(A_k) \leq \big(\frac{N}{N+1}\big)^k \mu(A)$, qui tend donc vers 0.

Et pour finir (b) a lieu car $\mu\Big(A \setminus \bigcup_{B \in \mathcal{D}} B\Big) \leq \mu(A_k)$ pour tout k, par (d).

Donc c'était le cas particulier où $\mu(A) < +\infty$. Dans le cas général, on va trouver dans $\mathbb{R}^n$ des ouverts disjoints U_m, et qui presque-recouvrent A. Il suffira ensuite d'appliquer le cas particulier déjà traité à chacun des $A \cap U_m$.

On commence par choisir, pour tout $m \geq 0$, un rayon r_m tel que $2^m < r_m \leq 2^{m+1}$, et $\mu(A \cap \partial B(0, r_m)) = 0$. C'est facile: les mauvais rayons $r \in]2^m, 2^{m+1}[$ pour lesquels $\mu(A \cap \partial B(0, r)) > 0$ sont en nombre au plus dénombrable, car les sphères $\partial B(0, r)$ sont disjointes, et la masse totale $\mu(A \cap B(0, 2^{m+1}))$ est finie (μ est une mesure de Radon).

On pose $U_0 = B(0, r_0)$ et, pour $m > 0$, $U_m = B(0, r_m) \setminus \overline{B}(0, r_{m-1})$. Ce sont des ouverts disjoints. Pour chaque m, la collection $\mathcal{C}_m$ des boules de $\mathcal{C}$ qui sont contenues dans U_m est encore une famille de Vitali pout $A \cap U_m$, et comme maintenant $\mu(A \cap U_m) < +\infty$, on peut trouver une sous-famille $\mathcal{D}_m$ de boules disjointes, telle que $\mu(A \cap U_m \setminus \bigcup_{B \in \mathcal{D}_m} B) = 0$.

On prend $\mathcal{D} = \bigcup_m \mathcal{D}_m$; ce sont encore des boules disjointes, parce que les U_m sont disjoints et les boules de $\mathcal{D}_m$ sont contenues dans U_m. Elles recouvrent μ-presque tout $A \cap \big(\bigcup_m U_m\big)$ par définition de $\mathcal{D}_m$. Et ce qui manque peut-être, à savoir $A \setminus \big(\cup_m U_m\big) = A \cap \big(\bigcup_m \partial B(0, r_m)\big)$, est de μ-mesure nulle par choix des r_m. $\square$

2. FONCTION MAXIMALE DE HARDY-LITTLEWOOD

On se donne une mesure de référence μ, borélienne et positive sur $\mathbb{R}^n$ (penser à la mesure de Lebesgue), et qu'on va aussi prendre localement finie (de Radon). On se donne aussi une seconde mesure (borélienne positive) ν. On pose

$$M_\mu(\nu(x)) = \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} d\nu \in [0,+\infty],$$

avec la convention que $0/0 = 0$ quand $\mu(B(x,r)) = 0$. C'est la fonction maximale (de Hardy-Littlewood, ou est-ce seulement dans le cas ou μ est la mesure de Lebesgue?) de ν. Pour la mesurabilité, voir la remarque plus bas, après la démonstration du lemme.

Le cas qui arrive le plus souvent est celui d'une mesure de densité, quand $d\nu(x) = |f(x)|d\mu(x)$, avec f localement intégrable pour μ, et alors cela donne

$$M_\mu f(x) = \sup_{r>0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} |f|\, d\mu.$$

Le plus souvent, $d\mu$ est la mesure de Lebesgue, et $M_\mu(\nu)$ est noté ν^*.

On se demande si ces fonctions sont intégrables ou dans des $L^p(d\mu)$. Ca sera utile parce que pas mal de problèmes avec des sup ou des limites dépendent de fonctions maximales; voir plus loin pour au moins un exemple.

Lemme. *Si ν est une mesure finie, $M_\mu(\nu) \in L^1_{faible}(d\mu)$, avec*

$$\mu\big(\{x \in \mathbb{R}^n\,; M_\mu(\nu(x)) > \lambda\}\big) \leq \frac{C_n \nu(\mathbb{R}^n)}{\lambda} \quad \textit{pour tout } \lambda > 0.$$

Notation: pour $p > 0$, on note $||g||_{L^p_{faible}(d\mu)} = \sup_{\lambda>0} \big\{\lambda^{-p}\, \mu\big(\{x \in \mathbb{R}^n\,; |g(x)| > \lambda\}\big)\big\}^{1/p}$, et ensuite $L^p_{faible}(d\mu)$ est la classe des fonctions mesurables g telles que $||g||_{L^p_{faible}(d\mu)} < +\infty$. C'est bien un espace vectoriel complet, mais attention, $||g||_{L^p_{faible}(d\mu)}$ n'est pas une norme, l'inégalité triangulaire n'est pas vérifiée avec la constante 1. Bien sûr, $L^p \subset L^p_{faible}$, par Tchebyshev.

Corollaire. *Si $f \in L^1(d\mu)$, alors $M_\mu f \in L^1_{faible}(d\mu)$, avec $||M_\mu f||_{L^1_{faible}(d\mu)} \leq C_n ||f||_{L^1(d\mu)}$.*

Démonstration. Le corollaire est immédiat dès qu'on a le lemme. Pour le lemme, on pose $O_\lambda = \{x \in \mathbb{R}^n\,; M_\mu(\nu(x)) > \lambda\}$, et pour tout $x \in O_\lambda$, on se donne une boule $B_x = B(x,r)$ centrée en x telle que $\nu(B_x) > \lambda\mu(B_x)$. On veut appliquer le lemme de recouvrement de Besicovitch, donc on va se contenter de recouvrir $\Omega = O_\lambda \cap B(0,R)$ (pour n'importe quel $R > 0$) par une collection au plus dénombrable de boules B_x, $x \in X \subset \Omega$, qui sont de recouvrement borné comme dans le lemme. Alors

$$\begin{aligned}\mu(\Omega) \leq \mu\Big(\bigcup_{x\in X} B_x\Big) \leq \sum_{x\in X}\mu(B_x) \leq \lambda^{-1}\sum_{x\in X}\nu(B_x) = \lambda^{-1}\sum_{x\in X}\int_{\mathbb{R}^n}\mathbf{1}_{B_x}d\nu \\ = \lambda^{-1}\int_{\mathbb{R}^n}\Big(\sum_{x\in X}\mathbf{1}_{B_x}\Big)d\nu \leq \lambda^{-1}\int_{\mathbb{R}^n}C_1 d\nu = C_1\lambda^{-1}\nu(\mathbb{R}^n)\end{aligned}$$

par définition des boules, puis le corollaire de Beppo-Lévi sur les séries, et enfin le fait que $\sum_{x\in X}\mathbf{1}_{B_x} \leq C_n$ partout. Il ne reste plus qu'à faire tendre R vers $+\infty$ pour récupérer le fait que $\mu(O_\lambda) \leq C_1\lambda^{-1}\nu(\mathbb{R}^n)$. □

Remarque sur la mesurabilité. Mauvaise conscience aidant, voici quelques mots sur la mesurabilité de $M_\mu(\nu)$.

D'abord, notons que pour tout $r > 0$, l'application f définie par $f(x) = \nu(B(x,r))$ est semi-continue inférieurement, ce qui signifie que pour tout λ, l'ensemble $A_\lambda = \big\{x \in \mathbb{R}^n\,;\, f(x) > \lambda\big\}$ est ouvert.

C'est facile: si $f(x) > \lambda$, et si $\{r_k\}$ est une suite strictement croissante qui tend vers r, on a que $f(x) = \nu(B(x,r)) = \lim_{k\to+\infty}\nu(B(x,r_k))$ (par union croissante et définition d'une mesure), donc il existe k tel que $\nu(B(x,r_k)) > \lambda$. Mais alors, pour $y \in B(x, r-r_k)$, $B(y,r)$ contient $B(x,r_k)$, donc $f(y) = \nu(B(y,r)) \geq \nu(B(x,r_k)) > \lambda$ et $y \in A_\lambda$.

Maintenant vérifions que pour tout $x \in \mathbb{R}^n$,

$$M_\mu(x) = \sup\big\{\nu(B(x,r))/\mu(B(x,r))\,;\, r\in\mathbb{Q}\big\} \tag{1}$$

L'inégalité $\geq$ est triviale, donc parlons de la réciproque. Fixons $x \in \mathbb{R}^n$ et $r > 0$, et choisissons une suite croissante de rationnels r_k qui tend vers r. Alors $\nu(B(x,r)) = \lim_{n\to+\infty}\nu(B(x,r_k))$, et pareil pour μ; vérifions que par conséquent $\nu(B(x,r))/\mu(B(x,r)) = \lim_{n\to+\infty}\nu(B(x,r_k))/\mu(B(x,r_k))$. Quand $\mu(B(x,r) \neq 0$, c'est clair. Sinon, $\mu(B(x,r_n)) = 0$ pour tout n, et alors ou bien $\nu(B(x,r)) = 0$ (et alors tout le monde vaut 0), ou bien $\nu(B(x,r)) > 0$, et alors $\nu(B(x,r_n)) > 0$ pour n grand, donc la limite est bien $+\infty$. La morale est que $\nu(B(x,r))/\mu(B(x,r))$ est inférieur au membre de droite de (1), et comme ceci vaut pour tout r, on prend le sup et on trouve (1).

Comme pour chaque r rationnel, $\mu(B(x,r))$ est une fonction semi-continue inférieurement (sci) de x, le rapport $\nu(B(x,r))/\mu(B(x,r))$ est une fonction borélienne de x, et donc M_μ est borélienne (comme borne supérieure dénombrable de fonctions boréliennes).

Et dans le cas où μ est la mesure de lebesgue, $\mu(B(x,r))$ est une fonction continue (de x et de r, d'ailleurs), donc de rapport $\nu(B(x,r))/\mu(B(x,r))$ est en fait sci pour chaque r, de sorte qu'en prenant le sup on trouve encore que la fonction maximale est sci.

Je n'ai pas eu le courage de vérifier ce qui se serait passé si on avait défini la fonction maximale avec des boules fermées. [Il faudrait essayer d'approximer par des suites décroissantes de rationnels, mais il faut peut-être faire plus attention au cas où $\mu(\overline{B}(x,r)) = 0$.]

Au cas où il y aurait une erreur ci-dessus, ou alors si les boules fermées posaient un problème, je me permets de renvoyer à Federer qui, si j'ai bien compris, montre que $M_\mu(\nu)$ est mesurable pour la mesure extérieure μ^* associée à μ, mais pas qu'elle est borélienne.

Ceci dit, noter aussi que même si l'on ne savait pas que $M_\mu(\nu)$ est mesurable, cela n'empêcherait pas de l'estimer comme ci-dessus, en montrant que la mesure extérieure de $O_\lambda = \big\{x \in \mathbb{R}^n\,;\, M_\mu\nu(x) > \lambda\big\}$ est inférieure à $C_1\lambda^{-1}\nu(\mathbb{R}^n)$. En effet, notre recouvrement de Besicovitch par des boules B_x donne un ouvert $\bigcup_{x\in X} B_x$ qui contient O_λ et dont la mesure est estimée ci-dessus.

Une variante de M_μ est la fontion maximale non centrée définie par

$$M'_\mu(\nu(x)) = \sup_B \frac{1}{\mu(B)} \int_B d\nu \in [0, +\infty],$$

où le sup est pris sur toutes les boules B qui contiennent x. Mattila prend les boules fermées, mais sauf erreur de ma part ou pourrait les prendre ouvertes aussi sans rien changer. La fonction $M'_\mu f = M'_\mu(f d\mu)$ est définie pareillement. Si la mesure μ est doublante, c.-à.-d. s'il existe $C \geq 1$ tel que

$$\mu(B(x,2r)) \leq C\mu(B(x,r)) \;\text{ pour tout } x \in \mathbb{R}^n \text{ et tout } r > 0,$$

on a le même genre de résultat que plus haut:

$$||M'_\mu(\nu)||_{L^1_{faible}(d\mu)} \leq C(n,\mu)\,\nu(\mathbb{R}^n),$$

et pareil pour les fonctions de L^1:

$$||M'_\mu f||_{L^1_{faible}(d\mu)} \leq C(n,\mu)||f||_{L^1(d\mu)}.$$

Démonstration sans se fatiguer: si B est une boule qui contient x, et si r est son diamètre, elle est contenue dans $B(x,2r)$, donc

$$\frac{1}{\mu(B)}\int_B d\nu \leq \frac{1}{\mu(B)}\int_{B(x,2r)} d\nu \leq C^2\frac{1}{\mu(B(x,2r))}\int_{B(x,2r)} d\nu \leq C^2 M_\mu(\nu(x)),$$

car $B(x,2r)\subset 3B\subset 4B$, donc $\mu(B(x,2r))\leq C^2\mu(B)$. On prend le sup et on trouve que $M'_\mu(\nu(x))\leq C^2 M_\mu(\nu(x))$ partout.

La vraie démonstration raisonnable est d'utiliser le lemme de 5-recouvrement (qui est plus facile à démontrer et marche dans d'autres contextes), et de faire comme plus haut en passant par les $5B$. En plus, ça donne une inégalité plus précise (restricted type inequality):

$$\mu\big(\{x\in\mathbb{R}^n\,;M'_\mu(\nu(x))>\lambda\}\big)\leq C_n\lambda^{-1}\nu\big(\{x\in\mathbb{R}^n\,;M'_\mu(\nu(x))>\lambda\}\big).$$

Voici en gros comment ça marche. On pose $O'_\lambda=\{x\in\mathbb{R}^n\,;M'_\mu(\nu(x))>\lambda\}$, et on recouvre O'_λ par des boules $B_j=B(x_j,r_j)$ telles que $\nu(B_j)>\lambda\mu(B_j)$. [Chaque $x\in O'_\lambda$ est contenu dans une telle boule par définition.] Pour ne pas avoir d'ennui avec la taille des boules, occupons-nous seulement pour l'instant de la variante de $M_\mu^{(R)}$ où l'on ne considère que des boules de rayon inférieur à R, où l'on s'est donné R (très grand) à l'avance.

Alors tous les rayons des B_j sont inférieurs à R, et le lemme de 5-recouvrement donne une famille B_j, $j\in J$, de boules disjointes, mais telle que les $5B_j$ recouvre O'_λ. Et

$$(*)\qquad \begin{aligned}\mu(O'_\lambda)&\leq\mu(\cup_{j\in J}5B_j)\leq\sum_{j\in J}\mu(5B_j)\leq C^3\sum_{j\in J}\mu(B_j)\\&\leq C^3\lambda^{-1}\sum_{j\in J}\nu(B_j)\leq C^3\lambda^{-1}\nu(\cup_j B_j)\leq C^3\lambda^{-1}\nu(O'_\lambda),\end{aligned}$$

où la dernière inégalité vient de ce que $M'_\mu(\nu(x))>\lambda$ pour tout $x\in B_j$, puisque $\nu(B_j)>\lambda\mu(B_j)$ et par définition de M'_μ.

Dans le cas général, on note que l'ensemble O'_λ associé à $M'_\mu(\nu)$ est l'union croissante dénombrable des ensembles associés à $M_\mu^{(R)}(\nu)$, et on passe à la limite dans $(*)$.

Théorème (Hardy-Littlewood). *Pour $1 < p \leq +\infty$, l'opérateur maximal $f \to M_\mu f$ est borné sur L^p. plus préciément, il existe $C = C(p,n)$ tel que $||M_\mu f||_{L^p(d\mu)} \leq C||f||_{L^p(d\mu)}$. Même énoncé pour la fonction maximale non centrée quand μ est doublante.*

Quand $p = +\infty$, c'est clair (et même $C = 1$). Le cas général s'en déduit, par interpolation avec la continuité $L^1 \to L^1_{faible}$; voir ci-dessous.

Exercice. Est-ce que la fontion maximale non centrée est également mesurable dans les circonstances ci-dessus (disons, avec des boules ouvertes)?

3. UN PEU D'INTERPOLATION REELLE

Voici le prototype de théorème d'interpolation réelle.

Theorème (Marcinkiewicz). *Soient* $(\Omega, \mathcal{A}, \mu)$ *et* $(\Omega', \mathcal{A}', \mu')$ *deux espaces mesurés,* $p_1 < p_2$ *et* $q_1 < q_2$ *des nombres strictement positifs. On autorise aussi* $+\infty$*, mais alors* L^∞_{faible} *est juste* L^∞*. On suppose aussi que* $q_1 \geq p_1$ *et* $q_2 \geq p_2$*. Soit* T *un opérateur sous-linéaire, défini sur* $L^{p_1}(d\mu) + L^{p_2}(d\mu)$*, et qui envoie de manière bornée* $L^{p_1}(d\mu)$ *dans* $L^{q_1}_{faible}(d\mu')$ *et* $L^{p_2}(d\mu)$ *dans* $L^{q_2}_{faible}(d\mu')$*. Alors, pour tout* $p \in]p_1, p_2[$*,* T *envoie de manière bornée* $L^p(d\mu)$ *dans* $L^q(d\mu)$*, où* q *est donné par la formule* $\dfrac{1}{q} = \dfrac{t}{q_1} + \dfrac{1-t}{q_2}$*, avec* $0 < t < 1$ *tel que* $\dfrac{1}{p} = \dfrac{t}{p_1} + \dfrac{1-t}{p_2}$*.*

Sous-linéaire signifie que $|T(f+g)(x)| \leq |Tf(x)| + |Tg(x)|$ pour tout choix de x, f, g.
"Envoie de manière bornée L^p dans L^q_{faible} signifie que $||Tf||_{L^q_{faible}} \leq C||f||_{L^p}$, où l'on note $||g||^q_{L^q_{faible}}$ le sup des $\lambda^q \mu'\big(\{x \in \Omega' \, ; \, |g(x)| > \lambda\}\big)$ (ce qu'on dominerait pour $g \in L^p$ par Tchebyshev); noter que c'est plus faible que la continuité de L^p dans L^q, puisque $||Tf||_{L^q_{faible}} \leq ||Tf||_{L^q}$.
Dans le cas de L^∞, il se trouve que la continuité de L^∞ dans BMO (voir plus loin) suffit pour interpoler. Chercher un théorème de Fefferman-Stein, si je me souviens bien, et il y a peut-être une hypothèse a priori à vérifier.

Pour la démonstration, et au moins dans un premier temps, on va se contenter du cas où $p_1 = q_1 = 1$ et $p_2 = q_2 = +\infty$, comme dans le cas du théorème maximal. Ca va simplifier pas mal, surtout à cause de la partie L^∞, mais dans le cas général, le shéma reste le même.

On prend $f \in L^p$, et on va la découper en deux morceaux de la manière la plus profitable. Après on appliquera la continuité à chacun des deux morceaux, et ça donnera le résultat. C'est le principe de l'interpolation réelle; il y a toute une théorie qu'on ne fera pas ici.

Avant de commencer, voici un moyen pratique de calculer les normes L^p, qu'on utilisera sur $g = Tf$ (mais gardons quand même l'exposant p parce qu'on a l'habitude).

Lemme. *Soit g mesurable, et pour tout $s > 0$, posons $G(s) = \mu'\big(\{|g(x)| > s\}\big)$. Alors $||f||_p^p = p\int_0^{+\infty} s^{p-1}G(s)ds$ pour $1 \leq p < +\infty$.*

On applique Fubini à la fonction mesurable positive $(x,s) \to ps^{p-1}\mathbf{1}_{0\leq s<|g(x)|}$, à intégrer contre $d\mu'(x)ds$ (qui est bien un produit de mesures sigma-finies dans les cas auxquels on pense vraiment). Si on intègre d'abord en x, on trouve $\int ps^{p-1}G(s)ds$. Si on commence par intégrer en s, on remarque que $\int_0^{|g(x)|} ps^{p-1}ds = |g(x)|^p$, de sorte que l'intégrale double vaut aussi $\int |g(x)|^p d\mu'(x)$. Si μ' n'est pas sigma-finie, on s'y ramène en remplaçant Ω' par $\{x \in \Omega'; |g(x)| > 0\}$. $\square$

N.B.: si vous n'aimez pas Fubini, ou ne croyez pas mes histoires de mesures sigma-finies, vous pouvez toujours découper Ω' en ensembles $E_k = \{x\,;\, 2^{k-1} < |g(x)| \leq 2^k\}$ et estimer brutalement les contributions. Ca fait juste perdre une constante. La même remarque vaut pour la suite des calculs.

Le plus simple pour notre décomposition est de couper f au niveau d'un $\lambda > 0$: on écrit $f = f_1 + f_2$, avec $f_2 = f\,\mathbf{1}_{\{|f(x)|\leq\lambda\}}$ (qui est bornée) et $f_1 = f\,\mathbf{1}_{\{|f(x)|>\lambda\}}$ (qui est dans L^1 si $f \in L^p$, par Hölder ou bêtement parce que $|f_1(x)| \leq \lambda^{1-p}|f(x)|^p$ partout). Notons que

$$||Tf_2||_\infty \leq M||f_2||_\infty \leq M\lambda. \tag{1}$$

Par sous-linéarité, $|Tf(x)| \leq M\lambda + |Tf_1(x)|$, donc $\{|Tf(x)| > 2M\lambda\}$ est contenu dans $\{|Tf_1(x)| > M\lambda\}$, ce qui donne, en posant $G(s) = \mu\big(\{|Tf(x)| > s\}\big)$ comme dans le lemme

$$\begin{aligned} G(2M\lambda) &\leq \mu'\big(\{|Tf_1(x)| > M\lambda\}\big) \leq (M\lambda)^{-1}||Tf_1(x)||_{L^1_{faible}} \\ &\leq \lambda^{-1}||f_1||_1 = \lambda^{-1}\int_{|f(x)|>\lambda} |f(x)|d\mu(x) \end{aligned}$$

par continuité de T de L^1 dans L^1_{faible}. Ou encore, en posant $s = 2M\lambda$

$$G(s) \leq 2Ms^{-1}\int_{|f(x)|>\frac{s}{2M}} |f(x)|d\mu(x)$$

On applique le lemme et on trouve

$$\begin{aligned}
||Tf||_p^p = p\int_0^{+\infty} s^{p-1}G(s)ds &\leq 2Mp\int_0^{+\infty} s^{p-2}\int_{|f(x)|>\frac{s}{2M}} |f(x)|d\mu(x) \\
&= 2Mp\int_x |f(x)|\int_{s<2M|f(x)|} s^{p-2}ds\ d\mu(x) \\
&= \frac{2Mp}{p-1}\int_x |f(x)||2Mf(x)|^{p-1}d\mu(x) = \frac{(2M)^p p}{p-1}\,||f||_p^p.
\end{aligned}$$

Bref, on a montré que T est borné sur L^p, avec une norme au plus $2M(p/p-1)^{1/p}$. C'est logique que l'estimation diverge près de $p=1$, puisque la continuité faible n'entraîne pas la continuité forte. □

Essayons encore dans le cas où $p_1 = q_1$ et $p_2 = q_2$, et où p_2 et q_2 ne sont pas infinis. Le calcul donne évidemment $q = p$ aussi. Alors

$$\begin{aligned}
G(\lambda) = \mu'\big(\{|Tf(x)| > \lambda\}\big) &\leq \mu'\big(\{|Tf_1(x)| > \lambda/2 \text{ ou } |Tf_2(x)| > \lambda/2\}\big) \\
&\leq \mu'\big(\{|Tf_1(x)| > \lambda/2\}\big) + \mu'\big(\{|Tf_2(x)| > \lambda/2\}\big) \\
&= \mu'\big(\{|Tf_1(x)|^{q_1} > (\lambda/2)^{q_1}\}\big) + \mu'\big(\{|Tf_2(x)|^{q_2} > (\lambda/2)^{q_2}\}\big) \\
&\leq C\lambda^{-q_1}||Tf_1||^{q_1}_{L^{q_1}_{faible}} + C\lambda^{-q_2}||Tf_2||^{q_2}_{L^{q_2}_{faible}} \\
&\leq C\lambda^{-q_1}||f_1||^{q_1}_{p_1} + C\lambda^{-q_2}||f_2||^{q_2}_{p_2} \\
&= C\lambda^{-q_1}\Big\{\int_{|f|>\lambda} |f|^{p_1}\Big\}^{q_1/p_1} + C\lambda^{-q_2}\Big\{\int_{|f|\leq\lambda} |f|^{p_2}\Big\}^{q_2/p_2}
\end{aligned}$$

Notre hypothèse que $p_i = q_i$ semble bien pratique maintenant, pour oter les puissances; ensuite, pour calculer $||Tf||_p^p$, on doit multiplier tout ça par $p\lambda^{p-1}$, puis intégrer en λ. La partie avec f_1 donne

$$\int p\lambda^{p-1}\lambda^{-p_1}\int_{|f(x)|>\lambda} |f(x)|^{p_1}d\mu(x)d\lambda$$

Comme $p > p_1$, l'intégrale en λ converge, et on trouve

$$p(p-p_1)^{-1}\int |f(x)|^{(p-p_1)+(p_1)}d\mu(x) = p(p-p_1)^{-1}||f||_p^p.$$

Le second morceau marche pareil, l'intégrale pour λ grand convergeant cette fois parce que $p < p_2$. On en déduit le résultat.

Pour le cas général, je n'ai toujours pas vérifié si on peut s'y ramener par des manipulations de compositions avant et après l'application de T (et en notant que l'argument ci-dessus résiste à la sous-linéarité faible où l'on suppose seulement que $|T(f_1 + f_2)| \leq CTf_1 + CTf_2$). En tout cas, il semble que ce soit fait, comme de juste, dans le livre de Zygmund (Trigonometric series).

4. DIFFERENTIATION DES MESURES

Vous avez peut-être déjà vu le théorème de Radon-Nikodym; on va en voir une version plus concrète avec des estimations (mais seulement dans $\mathbb{R}^n$ ou de bons espaces métriques).

Avant ça, un théorème qui a un rapport (mais on verra pourquoi plus tard) et qui est assez facile (maintenant qu'on a le théorème maximal). Pour changer, on se place dans $\mathbb{R}^n$, avec la mesure de Lebesgue.

Théorème de différentiation de Lebesgue. *Pour toute fonction f localement intégrable pour dx, on a*

$$\lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(x)-f(y)|dy = 0 \tag{1}$$

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

Corollaire souvent utile, même quand $d\mu = dx$: le cas de la fonction caractéristique de B borélien. Voir plus tard.

Notons aussi qu'ici on se donne un représentant (quelconque) de f avant d'écrire (1); bien sûr, si on modifie f sur un ensemble de mesure nulle, on modifie la liste des x tels que (1) a lieu, mais pas le théorème.

On a noté $|B(x,r)|$ la mesure de Lebesgue de $B(x,r)$. Une conséquence facile, à cause de l'inégalité triangulaire, est que

$$f(x) = \lim_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} f(y)dy \tag{2}$$

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

Démonstration. On peut se contenter du cas où $f \in L^1$, puisque pour regarder ce qui se passe dans $B(0,R)$, on peut remplacer f par $f\,\mathbf{1}_{B(0,2R)} \in L^1$.

Pour f continue, le théorème est trivial. Le cas général sera obtenu, de manière en fait très standard, par un argument de densité et de fonction maximale. Posons, pour $x \in \mathbb{R}^n$ et $r > 0$,

$$\omega_f(x,r) = \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(x)-f(y)|dy, \tag{3}$$

puis

$$\omega_f(x) = \limsup_{r\to 0} \omega_f(x,r). \tag{4}$$

On veut prouver que $\omega_f(x) = 0$ presque-partout. La chose est vraie sur une classe dense, car

$$\omega_f(x) = 0 \text{ partout quand } f \text{ est continue.} \tag{5}$$

Par ailleurs, l'inégalité triangulaire dit que

$$\omega_{f+g}(x,r) \leq \omega_f(x,r) + \omega_g(x,r) \tag{6}$$

pour tout choix de f, g, x, et r, ce qui donne aussitôt

$$\begin{aligned}\omega_{f+g}(x) &= \limsup_{r\to 0} \omega_{f+g}(x,r) \leq \limsup_{r\to 0} [\omega_f(x,r) + \omega_g(x,r)] \\ &\leq \limsup_{r\to 0} \omega_f(x,r) + \limsup_{r\to 0} \omega_g(x,r) \leq \omega_f(x) + \omega_g(x).\end{aligned} \tag{7}$$

Enfin, on va utiliser le fait, assez clair, que

$$\omega_f(x) \leq |f(x)| + M_{dx} f(x). \tag{8}$$

Soient $f \in L^1$ et $\varepsilon > 0$. Choisissons g continue telle que $||f-g||_{L^1} \leq \varepsilon$. Alors

$$\omega_f(x) \leq \omega_{f-g}(x) + \omega_g(x) = \omega_{f-g}(x) \leq |f-g|(x) + M_{dx}(f-g)(x), \tag{9}$$

par (5), donc (9) donne

$$\begin{aligned}|\{x\,;\,|\omega_f(x)| > \lambda\}| &\leq |\{x\,;\,|(f-g)(x)| > \lambda/2\}| + |\{x\,;\,|M_{dx}(f-g)(x)| > \lambda/2\}| \\ &\leq 2||f-g||_{L^1}/\lambda + 2||M_{dx}(f-g)||_{L^1_{faible}}/\lambda \leq 2\varepsilon\lambda + 2C\varepsilon\lambda\end{aligned}$$

par le théorème maximal. On fixe λ et on applique ça avec ε aussi petit qu'on veut, et on trouve $|\{x\,;\,|\omega_f(x)| > \lambda\}| = 0$. Finalement, $\omega_f(x) = 0$ presque-partout. □

Après en principe un rappel court sur ce qu'est le théorème de Radon-Nikodym (passé sous silence ici), on passe à la différentiation des mesures, pour laquelle je vais suivre Mattila assez servilement. On se donne deux mesures (boréliennes positives), la mesure de base μ et la mesure à étudier ν. On les suppose toutes les deux localement finies (de Radon). Pour ne pas avoir trop de problème de mesurabilité, on suppose que l'on travaille avec des boules ouvertes, ou qu'on a vérifié ce qu'il faut pour des boules fermées, ou que $\mu(\partial B(x,r)) = 0$ pour $x \in \mathbb{R}^n$ et $r > 0$, [de sorte qu'alors les démonstrations données plus haut marchent encore].

On va essayer, dans le cas où ν est absolument continue par rapport à μ, d'écrire ν comme mesure de densité par rapport à μ. Mais la différence avec la démonstration plus générale du théorème de Radon-Nikodym basée sur le théorème de représentation de Riesz (sur le dual de L^2) est que, grâce à la situation géométrique plus particulière, on pourra construire la densité comme limite de rapports $\nu(B(x,r))/\mu(B(x,r))$.

On commence par définir des densités ; pour alléger, on supposera que μ est fixée et on l'oubliera dans les notations. On peut penser à la mesure de Lebesgue. On pose

$$\begin{aligned} D\nu(x,r) &= \nu(B(x,r))/\mu(B(x,r)), \\ \overline{D}\nu(x) &= \limsup_{r\to 0} D\nu(x,r), \\ \underline{D}\nu(x) &= \liminf_{r\to 0} D\nu(x,r). \end{aligned}$$

On utilise encore la règle $0/0 = 0$ quand les deux mesures sont nulles; elle ne sont jamais infinies, puisque les mesures sont de Radon.

Deux mots de mesurabilité. Sauf erreur de ma part, nous sommes dans une situation où les limsup et liminf ne changent pas quand on se restreint à r rationnel (remplacer chaque rayon d'une suite qui tend vers 0 par un rationnel un peu inférieur pour lequel $\nu(B(x,r))$, donc aussi $D\nu(x,r)$, est très proche). Alors les sup, inf, limsup, et liminf sont mesurables par les arguments habituels.

Un commentaire aussi sur le mot de différentiation. Je ne suis pas certain de ce qui suit, mais en tout cas la seul lien que je peux voir avec la notion usuelle de différentiation est le suivant. Dans le cas où les deux mesures sont sur $\mathbb{R}$, on peut les écrire comme mesure de Stiljes à partir d'une fonction de répartition. La fonction associée à μ est une fonction croissante $F : \mathbb{R} \to \mathbb{R}$ telle que par exemple $F(x) = \mu([a, x[) + C$ pour $x > a$, et des formules semblables pour $x \leq a$. Dans ce cadre, differentier μ par rapport à la mesure de Lebesgue, c'est effectivement dériver F.

On aura besoin du lemme de presque-recouvrement obtenu en TD à partir du lemme de recouvrement de Besicovitch (Chapitre 1, près de la page 5). Le point central sera le lemme suivant.

Lemme 1. *Soient μ et ν comme plus haut, $A \subset \mathbb{R}^n$ quelconque, et $t \in]0, +\infty[$.*
1. Si $\underline{D}\nu(x) \leq t$ pour tout $x \in A$, alors $\nu(A) \leq t\mu(A)$.
2. Si $\overline{D}\nu(x) \geq t$ pour tout $x \in A$, alors $\nu(A) \geq t\mu(A)$.

Noter que c'est comme pour le théorème des accroissements finis: on part d'une information à l'échelle infinitésimale, et on obtient une information globale toute intégrée.

Si A n'est pas mesurable, la mesure est la mesure extérieure. On commence par 1. On se donne $\varepsilon > 0$. On trouve U ouvert contenant A tel que $\mu(U) \leq \mu(A)+\varepsilon$. Pour tout $x \in A$, on peut trouver des boules $B(x, r)$ centrées en x, aussi petites qu'on veut, et pour lesquelles $\underline{D}\nu(x, r) \leq t + \varepsilon$. Comme dans la définition de $\underline{D}\nu(x, r)$ on a pris des boules ouvertes, et que pour le lemme de recouvrement on veut des boules fermées, notons qu'on peut passer de l'un à l'autre. Soit $B(x, r)$ telle que $\underline{D}\nu(x, r) \leq t + \varepsilon$. Donc

$$\nu(B(x, r)) \leq (t + \varepsilon)\mu(B(x, r)).$$

Quand $r' < r$ tend vers r, $\nu(\overline{B}(x, r'))$ tend vers $\nu(B(x, r))$, et pareil pour $\mu(B(x, r))$. Alors pour $r' < r$ assez proche de r, $\nu(\overline{B}(x, r')) \leq (t+2\varepsilon)\mu(\overline{B}(x, r'))$ (distinguer 2 cas: ou bien $\mu(B(x, r)) > 0$, on peut diviser, et c'est facile, ou bien $\mu(B(x, r)) = 0$, et tout le monde est nul). Donc on peut choisir des boules fermées B_x, centrées en x, de rayons arbitrairements petits, et telles que $\nu(B_x) \leq (t + 2\varepsilon)\mu(B_x)$.

Par le lemme de presque-recouvrement, on peut trouver de telles boules disjointes et contenues dans U, avec $\nu(A \setminus \bigcup_x B_x) = 0$. Alors

$$\begin{aligned}\nu(A) \leq \nu(\bigcup_x B_x) = \sum_x \nu(B_x) &\leq (t + 2\varepsilon) \sum_x \mu(B_x) \\ = (t + 2\varepsilon)\mu(\bigcup_x B_x) &\leq (t + 2\varepsilon)\mu(U) \leq (t + 2\varepsilon)(\mu(A) + \varepsilon)\end{aligned}$$

On fait tendre ε vers 0 et on obtient le résultat.

Pour la partie 2., on fait pareil en échangeant le rôle de μ et ν; tout va bien, essentiellement parce que $0 < t < +\infty$ donc les hypothèses sont symétriques.

Lemme 2. *La densité $D\nu(x) = \lim_{r\to 0} D\nu(x,r)$ existe et est finie pour μ-presque tout $x \in \mathbb{R}^n$.*

Noter que par contre, la densité pourrait être infinie pour tout x dans le support de ν: penser à la mesure de longueur sur une droite.
Démonstration. On fixe une grosse boule B, et on pose, pour $0 < s < t < +\infty$, $A_{s,t} = \{x \in B\,;\, \underline{D}\nu(x) \leq s < t \leq \overline{D}\nu(x)\}$. Vérifions que $\mu(A_{s,t}) = 0$. On applique le lemme 1 et on trouve que

$$t\mu(A_{s,t}) \leq \nu(A_{s,t}) \leq s\mu(A_{s,t}) < +\infty$$

où la dernière inégalité vient de ce que μ est de Radon et on s'est restreint à B. Comme $t < s$ on en déduit bien que $\mu(A_{s,t}) = 0$.

De même, posons $A_t = \{x \in B\,;\, \overline{D}\nu(x) \geq t\}$. Alors $t\mu(A_t) \leq \nu(A_t) \leq \nu(B) < +\infty$, donc $\lim_{t\to+\infty} \mu(A_t) = 0$, et la μ-mesure de l'intersection des A_t est nulle. Bref, pour μ-presque tout $x \in B$, la densité supérieure est finie, et comme $\mu(A_{s,t}) = 0$ pour $0 < t < s < +\infty$, la densité existe μ-presque partout sur B. On prend B de plus en plus grand et on en déduit le lemme. □

Lemme 3. *Pour $B \subset \mathbb{R}^n$ borélien, on a $\nu(B) \geq \int_B D\nu(x)d\mu(x)$. Et il y a égalité si $\nu << \mu$.*

D'abord, $D\nu(x)$ existe μ-presque partout (Lemme 2), et est mesurable (car l'ensemble où les deux densités (inférieure et supérieure) diffèrent est mesurable), donc l'intégrale a un sens.

Absolument continue signifie que tout ensemble de ν-mesure nulle est de μ-mesure nulle.

Ça commence à ressembler à Radon-Nikodym, puisque dans le cas absolument continu, on dit que ν est donnée par la densité $D\nu(x)$.

Pour démontrer l'inégalité, on coupe B en tranches: on se donne $t > 1$ proche de 1, et on pose $B_k = \{x\,;\, D\nu(x) \text{ existe et } t^k \leq D\nu(x) < t^{k+1}\}$. Alors

$$\int_B D\nu(x)d\mu(x) = \sum_k \int_{B_k} D\nu(x)d\mu(x) \leq \sum_k t^{k+1}\mu(B_k) \leq \sum_k t\nu(B_k) \leq t\nu(B)$$

parce que l'on a oté un ensemble de mesure nulle où $D\nu$ n'est pas défini, et un ensemble où elle est nulle, puis à cause du lemme 1. On fait tendre t vers 1 et on obtient l'inégalité.

Maintenant on suppose $\nu << \mu$. Pour ν-pp x, la densité de μ par rapport à ν existe et est finie. Notons-la $D'(x)$. En un tel point x, $D\nu(x)$ est l'inverse de $D'(x)$, donc $D\nu(x) > 0$. En plus de ça, $D\nu(x) < +\infty$ pour μ-pp x, donc aussi pour ν-pp x. Alors, en reprenant les notations précédentes,

$$\nu(B) = \sum_k \nu(B_k) \leq \sum_k t^{k+1}\mu(B_k) \leq t \sum_k \int_{B_k} D\nu(x)d\mu(x) \leq t \int_B D\nu(x)d\mu(x).$$

(la première égalité vient de ce qu'on vient de dire, et après on applique l'autre moitié du lemme 1.) A nouveau on fait tendre t vers 1, et on obtient le lemme. $\square$

Lemme 4. *La mesure ν est absolument continue par rapport à μ si et seulement si $\underline{D}\nu(x) < +\infty$ ν-presque partout.*

Noter que c'est bien ν-p.p. cette fois. La partie directe est facile, puisqu'on a même vu que $D\nu(x)$ existe et est finie μ-p.p. Réciproquement, supposons que $\underline{D}\nu(x) < +\infty$ ν-p.p., donnons nous $A \subset \mathbb{R}^n$ négligeable pour μ, et montrons qu'il l'est aussi pour ν. La partie où la densité inférieure est infinie est de ν-mesure nulle par hypothèse, il ne reste plus qu'à traiter les ensembles $F = \{x \in A \cap B(0,R)\,;\, \underline{D}\nu(x) \leq N\}$. Mais le lemme 1 dit que $\nu(F) \leq N\mu(F) \leq N\mu(A) = 0$. $\square$

On s'occupe maintenant du cas singulier, où il y aura forcément moins à dire parce que justement μ et ν ne se parlent pas.

Rappelons que deux mesures μ et ν sont dites <u>mutuellement étrangères</u> (ou <u>singulières</u>) s'il existe $A \subset \mathbb{R}^n$ tel que $\nu(A) = \mu(\mathbb{R}^n \setminus A) = 0$ (mesures extérieures; mais du coup, pour nos mesures de Radon, et sauf erreur de ma part, on peut prendre A borélien). Notation: $\mu \perp \nu$. C'est symétrique.

Lemme 5. *Si la mesure ν est étrangère à μ, alors $D\nu(x) = 0$ μ-presque partout.*

Soit A comme ci-dessus. Pour $t > 0$, on applique le lemme 1 à $Z_t = \{x \in \mathbb{R}^n\,;\, \overline{D}\nu(x) \geq t\}$, et on trouve

$$t\mu(Z_t) = t\mu(A \cap Z_t) \leq \nu(A \cap Z_t) \leq \nu(A) = 0.$$

Par union croissante d'ensembles de mesure nulle, on trouve bien que que $\overline{D}\nu(x) = 0$ pour μ-p.p. x, et évidemment $D\nu(x) = 0$ pour ces points.

On est prêt pour la version précisée de Radon-Nikodym.

Théorème. *On se donne (comme depuis le début) deux mesures de Radon (boréliennes localement finies) sur* $\mathbb{R}^n$*. On pose* $A = \big\{x \in \mathbb{R}^n \,;\, \underline{D}\nu(x) < +\infty\big\}$*,* $\nu_1 = \mathbf{1}_A \nu$ *et* $\nu_2 = \mathbf{1}_{\mathbb{R}^n \setminus A}\nu$*. Alors* $\nu = \nu_1 + \nu_2$*,* $\nu_1 << \mu$*,* $\nu_2 \perp \mu$*, et* $d\nu_1(x) = D\nu_1(x)d\mu(x) = D\nu(x)d\mu(x)$*.*

Rappelons qu'à cause du fait qu'on travaille avec des boules ouvertes, $\underline{D}\nu$, $\overline{D}\nu$, et A sont boréliens. Si on avait pris des boules fermées, disons pour ne pas se fatiguer qu'on aurait supposé que μ ne charge pas les sphères, et qu'alors $\underline{D}\nu$, $\overline{D}\nu$, et A sont boréliens aussi. Ou alors, suivre le livre de Mattila. De plus, par le lemme 3, $D\nu(x)$ et $D\nu_1(x)$ sont définies μ-presque partout, donc $D\nu(x)d\mu(x)$ et $D\nu_1(x)d\mu(x)$ sont définies.

Il est clair que $\nu = \nu_1 + \nu_2$. De plus, $\nu_2 \perp \mu$ parce que ν_2 est portée par $\mathbb{R}^n \setminus A$ et $\mu(\mathbb{R}^n \setminus A) = 0$ à cause du lemme 2.

Par le lemme 5, $D\nu_2(x) = 0$ μ-presque partout. On en déduit (petit calcul avec la définition) que $D\nu(x) = D\nu_1(x)$ μ-presque partout. De sorte que $D\nu_1(x)d\mu(x) = D\nu(x)d\mu(x)$.

Pour voir que $\nu_1 << \mu$, on va appliquer le lemme 4. Par définition de A, $\underline{D}\nu(x) < +\infty$ partout sur A, donc ν_1-presque partout. C'est encore vrai pour $\underline{D}\nu_1(x)$, qui est plus petite. Le lemme 4 dit donc que $\nu_1 << \mu$.

Maintenant le lemme 3, appliqué à ν_1, dit que $d\nu = D\nu_1 d\mu$. On sait déjà que $D\nu_1 d\mu = D\nu d\mu$. On a donc fini. □

Remarques.

1. Mattila suppose μ et ν finies, ça semble un peu inquiétant (est-ce que quelque chose m'aurait échappé?), mais en fait non: notre résultat est local, et comme on a supposé nos mesures localement finies (et pas seulement sigma-finies), c'est en fait pareil.
2. Le théorème de Radon-Nikodym (décomposition unique de ν en une partie singulière et une partie absolument continue, donnée par intégration contre une densité) est vrai dans un cadre plus général (mesures abstraites sigma-finies).

L'unicité est assez facile (exercice?). Noter aussi que $\nu << \mu$ dans le théorème ssi $\nu_2 = 0$, essentiellement par définition. La base de la démonstration est le théorème de Riesz. (pour trouver une fonction à partir d'une forme linéaire continue sur L^2, obtenue come on s'en doute par intégration des fonctions).

Pour ce qu'on vient de faire, l'intérêt est aussi dans la manière constructive de trouver la décomposition. Ceci dit, un exercice intéressant, mais que je n'ai pas fait par paresse, consiste à essayer de voir jusqu'où nous mène le théorème de Radon-Nikodym abstrait dans le cas présent.

Voici en tout cas comment on commence. Par Radon-Nikodym, $\nu = \nu_1 + \nu_2$, avec $\nu_1 << \mu$ et $\nu_2 \perp \mu$, et même $d\nu_1 = f d\mu$. Il faudra vérifier (sans doute par les lemmes 1 et 5) que $D\nu_2(x) = 0$ μ-p.p., mais ensuite pour savoir que $f(x) = D\nu(x) < +\infty$ μ-p.p., on peut essayer le théorème de différentiation de Lebesgue.

Le lemme de recouvrement de Besicovitch est faux dans certains espace métriques qui ont l'air raisonnables, l'exemple le plus classique étant le groupe de Heisenberg, muni de sa métrique de Carnot-Carathéodory. Le théorème de différentiation des mesures y est quand même vrai, mais il faut être subtil et trouver une démonstration différente.

Exercice. Redémontrez sans fonction maximale que si μ est comme ci-dessus et $f : \mathbb{R}^n \to [-\infty, +\infty]$ est localement intégrable, alors

$$\lim_{r\to 0} \frac{1}{\mu(B(x,r))} \int_{B(x,r)} f(y) d\mu(y)$$

pour μ-presque tout x. [Commencer par le cas positif.]

Exercice. Si $A \subset \mathbb{R}^n$ est borélien et μ est comme ci-dessus, alors
$\lim_{r\to 0} \mu(A \cap B(x,r))/\mu(B(x,r)) = 1$ pour μ-p.p. $x \in A$;
$\lim_{r\to 0} \mu(A \cap B(x,r))/\mu(B(x,r)) = 0$ pour μ-p.p. $x \in \mathbb{R}^n \setminus A$.
[Appliquer les résultats ci-dessus au cas où $d\nu = 1_A \mu$, et utiliser l'unicité de la densité de Radon-Nikodym.]
C'est un résultat plus facile, mais ça sert souvent.

5. INTERPOLATION COMPLEXE

A nouveau, il s'agit surtout de dire comment ça marche, parce que c'est rigolo, et pas de donner un énoncé trop général.

Le centre de la démonstration est de dire que certains objets ont un prolongement analytique, et d'appliquer le principe du maximum. Mais comme le domaine où l'on veut appliquer le principe du maximum n'est pas borné, il nous faut un petit énoncé spécial, que je prends dans Rudin.

Un premier énoncé simple, d'abord.

Théorème. *Soient* $\Omega = \{z = x + iy\,;\, a < x < b\} \subset \mathbb{C}$ *une bande verticale ouverte,* $f : \Omega \to \mathbb{C}$ *une application analytique. On suppose que* f *est bornée, et qu'elle admet un prolongement continu à la bande fermée* $\{z = x+iy\,;\, a \leq x \leq b\}$*. On pose* $M(x) = \sup_{y\in\mathbb{R}} |f(x + iy)|$ *pour* $a \leq x \leq b$*. Alors*

$$M(x) \leq M(a)^{\frac{b-x}{b-a}} M(b)^{\frac{x-a}{b-a}} \quad \text{pour } a \leq x \leq b. \tag{1}$$

Autrement dit, $\ln(M(x))$ est convexe (l'équivalence est facile à vérifier). Le cas particulier le plus important est quand $M(a) \leq 1$ et $M(b) \leq 1$, qui donne $M(x) \leq 1$. On va commencer par s'y ramener en divisant par g, où

$$g(z) = M(a)^{\frac{b-z}{b-a}} M(b)^{\frac{z-a}{b-a}}.$$

Ici on suppose que $M(a) > 0$ et $M(b) > 0$; sinon, si par exemple $M(a) = 0$, noter que la démonstration s'applique en remplaçant $M(a)$ par n'importe quel $\varepsilon > 0$; il reste ensuite à faire tendre ε et noter que le membre de droite de (1) tend vers 0.

Noter que g est bien définie et holomorphe sur un voisinage de la bande (utiliser $M^w = \exp(w \ln(M))$ pour $M > 0$ et $w \in \mathbb{C}$), et que $|g(x+iy)| = M(x)$ pour $z = x + iy \in \Omega$. Autrement dit, son module sur la droite $x + i\mathbb{R}$ est le membre de droite de (1). Donc g ne s'annule pas, et même $1/g$ est bornée sur la bande, f/g vérifie les hypothèses du théorème avec $M(a) \leq 1$ et $M(b) \leq 1$, et les conclusions pour f se déduisent des conclusions pour f/g.

Donc il suffit de supposer que $M(a) \leq 1$ et $M(b) \leq 1$ et prouver que $M(x) \leq 1$. On veut se ramener au principe du maximum sur un domaine borné, donc on multiplie par

$$h_\varepsilon(z) = (1 + \varepsilon(z - a))^{-1}$$

avec ε petit qu'on fera tendre vers 0 [ou, si vous préférez, $h_\varepsilon(z) = (1 + \varepsilon(z - a + 1))^{-1}$]. Noter que $h_\varepsilon(z) \leq 1$ sur la bande fermée (car $|1 + \varepsilon(z - a)| \geq 1 + \varepsilon(x - a) \geq 1$), donc $h_\varepsilon f$ vérifie encore les hypothèses. Mais l'avantage est que le maximum de $h_\varepsilon f$ est à trouver dans un domaine borné. En effet soit M tel que $|f(z)| \leq M$ partout; noter que $h_\varepsilon(x + iy) \leq \varepsilon|y|^{-1}$, donc $|h_\varepsilon f(z)| \leq 1$ pour $|y| \geq \varepsilon^{-1}M$. On applique le principe du maximum sur le domaine borné $\Omega \cap \{|y| < \varepsilon^{-1}M\}$, et on trouve que $|h_\varepsilon f(z)| \leq 1$ dans ce domaine aussi, donc partout.

Ainsi, $|h_\varepsilon f(z)| \leq 1$ pour $z \in \Omega$ et tout $\varepsilon > 0$. On fait tendre ε vers 0 et on trouve que $|f(z)| \leq 1$. $\square$

Une certaine hypothèse (comme "f est bornée") est nécessaire, parce qu'autrement on a des contre-exemples. Par exemple, faisons un quart de tour pour simplifier, et considérons le domaine

$$\Omega_0 = \{x + iy\,;\, -\frac{\pi}{2} < y < \frac{\pi}{2}\}$$

et $f(z) = \exp(\exp(z))$. Alors $f(x \pm \frac{i\pi}{2}) = \exp(\pm ie^x)$, donc f est de module 1 sur les deux bords, mais sa restriction à l'axe réel $\{y = 0\}$ est e^{e^x}, qui n'est pas bornée du tout.

En fait, ceci est le pire contre-exemple possible, à cause du résultat suivant.

Théorème (Phragmen-Lindelöf). *Soit f une fonction holomorphe sur Ω_0. On suppose que f a une extension continue à $\overline{\Omega}_0$, telle que $|f(x \pm \frac{i\pi}{2})| \leq 1$ pour $x \in \mathbb{R}$. On suppose également qu'il existe $\alpha < 1$ et $A \geq 0$ tels que*

$$|f(x + iy)| \leq \exp(A \exp(\alpha|x|)) \quad \textit{pour } x + iy \in \Omega_0. \tag{2}$$

Alors $|f(z)| \leq 1$ pour $z \in \Omega_0$.

Noter que ceci permet d'améliorer le théorème précédent, en supposant un peu moins sur la fonction f. Par exemple, une estimation du type $|f(z)| \leq C \exp(A|z|)$ (au lieu de f bornée) est largement suffisante dans tous les cas, sans avoir à faire le changement de variable précis pour se ramener à Ω_0.

La démonstration est du même type, on doit juste changer de fonction écraseuse. Prenons β tel que $\alpha < \beta < 1$ et $\varepsilon > 0$ petit, et essayons

$$h_\varepsilon(z) = \exp\{-\varepsilon(e^{\beta z} + e^{-\beta z})\}. \tag{3}$$

Il faut voir si c'est assez petit à l'infini pour tuer f. Prenons $z \in \overline{\Omega}_0$, et notons que

$$\Re(e^{\beta z} + e^{-\beta z}) = e^{\beta x} \Re(e^{i\beta y}) + e^{-\beta x} \Re(e^{-i\beta y}) = (e^{\beta x} + e^{-\beta x}) \cos(\beta y).$$

On note que $\beta|y| \leq \beta\pi/2 < \pi/2$, donc $\delta = \cos(\beta\pi/2)$ est strictement positif et par conséquent $\Re(e^{\beta z} + e^{-\beta z}) \geq \delta(e^{\beta x} + e^{-\beta x})$. Et en reprenant (3), on trouve que

$$|h_\varepsilon(z)| \leq \exp\big\{ - \varepsilon\delta(e^{\beta x} + e^{-\beta x})\big\}.$$

D'une part, $|h_\varepsilon(z)| \leq 1$, donc $|h_\varepsilon f| \leq 1$ sur les bords de la bande. D'autre part, on va vérifier que $|h_\varepsilon f|$ tend vers 0 à l'infini, ce qui permettra d'appliquer l'argument précédent pour se ramener à appliquer le principe du maximum sur un ensemble borné, et dire que $|h_\varepsilon f| \leq 1$ partout. Et de conclure en faisant tendre ε vers 0, puisque $\lim_{\varepsilon \to 0} h_\varepsilon(x) = 1$ partout.

Vérification: compte tenu de (2), il s'agit de voir que $A \exp(\alpha|x|) - \varepsilon\delta(e^{\beta x} + e^{-\beta x})$ tend vers $-\infty$ quand $|x|$ tend vers $+\infty$. C'est clair parceque $\beta > \alpha$. $\square$

Donc voici un théorème d'interpolation complexe. A nouveau, c'est peut-être le principe de la démonstration qui est la chose la plus intéressante.

Notations pour le prochain théorème. Soient, pour $i = 1, 2$, $(\Omega_i, \mathcal{A}_i, \mu_i)$ des espaces mesurés, avec des mesures sigma-finies. On note $\mathcal{E}_i$ la classe des fonctions étagées sur Ω_i et intégrables pour μ_i. On se donne aussi des exposants $p_1, p_2, r_1, r_2 \in [1, +\infty]$. Et on se donne un opérateur linéaire T, défini sur $\mathcal{E}_1$ et à valeurs dans $L^{r_1}(\Omega_2) \cap L^{r_2}(\Omega_2)$.

Théorème. *On suppose que T a une extension continue de $L^{p_1}(\Omega_1)$ dans $L^{r_1}(\Omega_2)$, de norme au plus M_1, et une extension continue de $L^{p_2}(\Omega_1)$ dans $L^{r_2}(\Omega_2)$, de norme au plus M_2. Pour $0 \leq t \leq 1$, déterminons p et r par les formules*

$$\frac{1}{p} = \frac{1-t}{p_1} + \frac{t}{p_2} \quad \textit{et} \quad \frac{1}{r} = \frac{1-t}{r_1} + \frac{t}{r_2}$$

Alors T a aussi une extension continue de $L^p(\Omega_1)$ dans $L^r(\Omega_2)$, de norme au plus M, avec $M = M_1^{1-t} M_2^t$.

Quelques remarques pour commencer.

Les exposants p et r sont bien déterminés, même si certains de nos exposants sont $+\infty$.

L'énoncé autorise le croisement $p_1 < p_2$ et $r_1 > r_2$.

Noter que $\mathcal{E}_i$ est seulement dense dans $L^p(d\mu_i)$ quand $p < +\infty$, donc, par exemple, nous n'avons pas vraiment défini T entièrement sur L^{p_1} si $p_1 = +\infty$. Ca n'est pas grave. En fait, on va caractériser la phrase "T a une extension continue de $L^p(d\mu_1)$ dans $L^r(d\mu_2)$" par la formule (1) suivante, qui utilise la dualité. On note q l'exposant conjugué de r (donc $\frac{1}{q} + \frac{1}{r} = 1$). Alors montrons que T a une extension continue de $L^p(d\mu_1)$ dans $L^r(d\mu_2)$ si et seulement s'il existe $M \geq 0$ tel que

(1)
$$|\langle Tf, g\rangle| \leq M \text{ pour tout choix de } f \in \mathcal{E}_1, g \in \mathcal{E}_2 \text{ telles que } ||f||_p \leq 1 \text{ et } ||g||_q \leq 1.$$

Et en fait le meilleur M donne la meilleure norme d'extension.

Si T a une extension continue, c'est un opérateur borné, et (1) vient de ce que $||Tf||_r \leq M||f||_p$, et de Hölder.

Réciproquement, si on a (1), et si $f \in \mathcal{E}_1$ pour commencer, (1) dit que Tf définit une forme linéaire bornée sur la fermeture de $\mathcal{E}_2$ dans L^q, de norme au plus $M||f||_p$. On en déduit que $||Tf||_r \leq M||f||_p$ (on sait déjà, par hypothèse, que $Tf \in L^r$). Si $p < +\infty$, on en déduit que T a un unique prolongement (par uniforme continuité). Si $p = +\infty$, on a un unique prolongement à l'adhérence de $\mathcal{E}_1$, mais rien n'empêche ensuite de reprolonger à L^∞ par Hahn-Banach.

La démonstration du théorème consiste maintenant à utiliser le théorème de Phragmen-Lindelöf, ou une version plus faible, pour déduire l'inégalité (1) de ses analogues pour les indices de l'énoncé.

On doit donc essayer de définir des fonctions analytiques. On fixe $f \in \mathcal{E}_1$ et $g \in \mathcal{E}_2$, et on pose, pour z dans la bande verticale $\Omega = \{z = x + iy\,;\, 0 \leq x \leq 1\}$

$$(2) \qquad f_z(u) = \frac{f(u)}{|f(u)|}|f(u)|^{az+b} \quad \text{et} \quad g_z(u) = \frac{g(u)}{|g(u)|}|g(u)|^{cz+d},$$

avec $a, b, c, d \in \mathbb{R}$ sont à choisir bientôt. Plus précisément, on garde $f_z(u) = 0$ quand $f(u) = 0$, et autrement on prend $|f(u)|^{az+b} = \exp((az+b)\ln(|f(u)|)$; pareil pour g. Dans tous les cas, on a un résultat holomorphe en z (à u fixé).

En fait, comme f et g sont étagées intégrables, les choses sont assez simples. On peut écrire $f = \sum_j \lambda_j \mathbf{1}_{A_j}$, avec des λ_j non nuls et des A_j disjoints de μ_1-mesures finies, et pareillement $g = \sum_k \eta_k \mathbf{1}_{B_k}$. Noter qu'alors

$$(3) \qquad f_z = \sum_j \lambda_{j,z} \mathbf{1}_{A_j} \quad \text{et} \quad g_z = \sum_k \eta_{k,z} \mathbf{1}_{B_k},$$

avec des coefficients $\lambda_{j,z}$ et $\eta_{k,z}$ facilement calculables (avec des puissances), et qui sont tous des fonctions holomorphes de z dans la bande verticale Ω. Noter au passage que f_z et g_z sont étagées intégrables.

On veut appliquer le théorème de Phragmen-Lindelöf à la fonction H sur Ω définie par

$$H(z) = \int_{\Omega_2} Tf_z(u) g_z(u) d\mu_2(u). \tag{4}$$

Noter que f_z est étagée intégrable, donc Tf_z est bien défini, est même dans L^r, et s'intègre donc bien contre la fonction étagée intégrable g_z. Par linéarité,

$$H(z) = \sum_j \sum_k \lambda_{j,z}\, \eta_{k,z} a_{j,k} \tag{5}$$

où $a_{j,k} = \langle T\mathbf{1}_{A_j}, \mathbf{1}_{B_k} \rangle$ est juste un coefficient numérique.

Comme les fonctions $\lambda_{j,z}$ et $\eta_{k,z}$ sont juste des sommes de puissances, H a une croissance au plus exponentielle, ce qui permettra d'appliquer le théorème de Phragmen-Lindelöf.

On peut choisir f et g pour que

$$||f||_p = ||g||_q = 1, \tag{6}$$

ce qui simplifiera les calculs.

Notons que puisque $\frac{1}{r} = \frac{1-t}{r_1} + \frac{t}{r_2}$ et $\frac{1}{q} = 1 - \frac{1}{r}$, il vient $\frac{1}{q} = \frac{1-t}{q_1} + \frac{t}{q_2}$, où q_1 et q_2 sont les exposants conjugués de r_1 et r_2.

Posons maintenant $p(x) = \frac{p}{ax+b}$; alors

$$||f_z||_{p(x)}^{p(x)} = \int_{\{f(u)\neq 0\}} |f(u)|^{(ax+b)p(x)} = ||f||_p^p = 1. \tag{7}$$

De même, posons $q(z) = \frac{q}{cx+d}$; alors

$$||g_z||_{q(x)}^{q(x)} = \int_{\{g(u)\neq 0\}} |g(u)|^{(cx+d)q(x)} = ||g||_q^q = 1. \tag{8}$$

Maintenant on veut choisir a, b, c, d. On veut que

$$p(0) = p_1,\ p(1) = p_2,\ q(0) = q_1,\ \text{et } q(1) = q_2. \tag{9}$$

ou encore, en prenant les inverses, on veut que la fonction $(ax+b)/p$ vaille $1/p_1$ en 0 et $1/p_2$ en 1. Ceci détermine a et b, et par linéarité, $(ax+b)/p$ vaut $1/p$ en t. De sorte que $at+b=1$, donc $f_t = f$ (par (2)).

De même, les conditions sur les q déterminent c et d, et on trouve $(ct+d)/q = 1/q$, puis $ct+d=1$ et $g_t = g$. Alors $H(t) = \langle Tf, g\rangle$ (par (4)).

D'un autre côté, quand $z \in i\mathbb{R}$,

(10)

$$|H(z)| \leq |\langle Tf_z, g_z\rangle| \leq |||T|||_{p_1,r_1}||f_z||_{p_1}||g_z||_{q_1} = |||T|||_{p_1,r_1}||f_0||_{p_1}||g_0||_{q_1} \leq M_1$$

parce que $|f_z|$ ne dépend que de sa partie réelle (et pareil pour g_z), puis (7), et parce que $p(0) = p_1$ et $q(0) = q_1$.

Pour les mêmes raisons, $|H(z)| \leq M_2$ pour $z \in 1+i\mathbb{R}$. Finalement, on applique Phragmen-Lindelöf (et le lemme précédent) et on trouve $|\langle Tf, g\rangle| = |H(t)| \leq M_1^{1-t}M_2^t$, comme voulu pour (1), qui lui-même donne le théorème. $\square$

Pour ne pas alourdir trop, on n'a pas mis de poids dans l'énoncé. Mais on aurait pu en mettre. Soit ω_i une fonction mesurable positive sur Ω_i, et donnons-nous encore des puissances α_i et β_i. Supposons cette fois que pour $i = 1, 2$, T envoie continument $L^{p_i}(\Omega_1, \omega_1^{\alpha_i}d\mu_1)$ dans $L^{r_i}(\Omega_2, \omega_2^{\beta_i}d\mu_2)$, avec p_i et r_i comme dans l'énoncé. Supposons également ce qu'il faut pour que les fonctions f_z qu'on est amené à écrire dans l'argument soient dans les bons L^p (c'est un peu plus délicat à écrire à cause des exposants variables, mais dans des cas pratiques ça serait assez facile à arranger). Alors T envoie aussi $L^p(\Omega_1, \omega_1^{\alpha}d\mu_1)$ dans $L^r(\Omega_2, \omega_2^{\beta}d\mu_2)$, où sauf nouvelle erreur de ma part $\frac{1}{\alpha} = \frac{1-t}{\alpha_1} + \frac{t}{\alpha_2}$, et $\frac{1}{\beta} = \frac{1-t}{\beta_1} + \frac{t}{\beta_2}$ La démonstration et la vérification sont laissées en exercice. En principe, le point essentiel est de poser $f_z(u) = \frac{f(u)}{|f(u)|}|f(u)|^{az+b}\omega_1(u)^{a'z+b'}$ et pareillement pour g_z, et de s'arranger pour choisir les constantes a', b', c', d' pour pouvoir appliquer les continuités aux extrémités; si les exposants de l'énoncé sont exacts, on trouve ce qu'il faut pour la continuité cherchée en regardant au point t.

Moralités du chapitre: il est très amusant que le principe du maximum puisse servir dans des circonstances (continuité des opérateurs, même dans un cadre réel) où on ne l'aurait pas attendu; c'est souvent aux extrémités de l'intervalle de validité que les majorations sont les plus difficiles à obtenir; pourtant, c'est parfois bien pratique de ne pas avoir à faire la vérification, surtout quand les espaces extrêmes sont plus simples.

6. $W^{1,p}$ ET INÉGALITÉS DE POINCARÉ

On va essayer de faire le minimium vital pour arriver aux inégalités de Sobolev et Poincaré. [En fait, apparemment on s'arrêtera un peu avant Sobolev, mais vous le verrez peut-être en exercice.] On se place dans un ouvert U (en principe, assez simple et connexe) de $\mathbb{R}^n$.

On va commencer par définir les espaces $W^{k,p}(U)$ lorsque k est un entier positif (ou nul) et $p \in [1, +\infty]$.

Pour $k = 0$, prendre $W^{0,p}(U) = L^p(U)$, avec la même norme.

Pour $k = 1$ (le plus important pour nous), on dira que $f \in W^{1,p}(U)$ si $f \in L^p$ et si de plus ses dérivée premières (les $\frac{\partial f}{\partial x_i}$), prises au sens des distributions, sont dans $L^p(U)$. [Voir la traduction ci-dessous.] Et on prendra la norme $||f||_{W^{1,p}(U)} = ||f||_p + \sum_i ||\frac{\partial f}{\partial x_i}||_p$.

Pour k entier plus grand, on dit que $f \in W^{k,p}(U)$ si f, ainsi que toutes ses dérivées jusqu'à l'ordre k (toujours prises au sens des distributions) sont dans L^p, et on peut prendre pour norme la somme des normes L^p.

Traduction. Soit f localement intégrable sur U. Alors, elle définit une distribution sur U, en posant $\langle f, \varphi \rangle = \int_U f(x)\varphi(x)dx$ pour $\varphi \in C_c^\infty(U)$ (effet de la distribution sur la fonction test).

Ensuite, on définit les distributions $\frac{\partial f}{\partial x_i}$ par $\langle \frac{\partial f}{\partial x_i}, \varphi \rangle = -\langle f, \frac{\partial \varphi}{\partial x_i} \rangle$ pour $\varphi \in C_c^\infty(U)$.

Donc on dira que la dérivée partielle $\frac{\partial f}{\partial x_i}$ est dans L^p s'il existe $g_i \in L^p$ telle que

$$(1) \qquad -\langle f, \frac{\partial \varphi}{\partial x_i} \rangle = \langle g_i, \varphi \rangle \text{ pour toute fonction test } \varphi \in C_c^\infty(U).$$

Facile à vérifier (= exercice!): s'il existe $g \in L^p$ telle que (1) ait lieu, elle est unique. Ou, de manière à peu près équivalente, si la fonction $g \in L^1_{loc}$ définit une distribution nulle, elle est nulle presque-partout.

Parfois, le fait que $\frac{\partial f}{\partial x_i}$ est dans L^p se décide plus facilement par dualité. Commençons par discuter dans le cas où $1 < p \leq +\infty$. Soit q l'exposant conjugué de p.

Si $\frac{\partial f}{\partial x_i}$ est dans L^p, alors, en utilisant (1) on trouve que

$$(2) \qquad |\langle f, \frac{\partial \varphi}{\partial x_i} \rangle| \leq C||\varphi||_q \text{ pour tout } \varphi \in C_c^\infty(U),$$

avec $C = ||g_i||_p$. Par un petit passage à la limite, ceci implique d'ailleurs que

$$(3) \qquad |\langle f, \frac{\partial \varphi}{\partial x_i} \rangle| \leq C||\varphi||_q \text{ pour tout } \varphi \in C_c^1(U)$$

(avec la même constante C). Et réciproquement, si on a (2) ou (3) (et si $p > 1$), ceci implique que l'application $\varphi \to \langle f, \frac{\partial \varphi}{\partial x_i} \rangle$ s'étend en une forme linéaire continue sur L^q, donc (par Riesz), est donnée par une fonction de L^p. Bref, quand $p > 1$, (2) ou (3) implique que $\frac{\partial f}{\partial x_i}$ est dans L^p.

Quand $p = 1$, il y a un autre espace un peu plus grand que $W^{1,1}(U)$, l'espace BV des fonctions $f \in L^1$ (ou alors, $f \in L^1_{loc}$, cela pourrait dépendre des auteurs) telles que pour tout i, $\frac{\partial f}{\partial x_i}$ est une mesure (signée ou complexe) finie. Il se trouve que le dual de la fermeture, pour la norme sup, de l'ensemble des fonctions bornées à support compact, est justement l'espace vectoriel des mesures finies sur U, muni de la norme $||\nu||$ = variation totale de ν. Donc, quand $p = 1$ et $q = +\infty$, (2) et (3) caractérisent le fait que $\frac{\partial f}{\partial x_i}$ est une mesure (signée) finie. On en reparlera au chapitre sur BV.

Quand $p = 2$ et $U = \mathbb{R}^n$, les choses s'expriment très bien en termes de transformée de Fourier. L'espace $W^{k,2}$ est alors le plus souvent noté H^k, et il est assez facile de voir que $f \in H^k(\mathbb{R}^n)$ si et seulement si

$$(4) \qquad ||f||^2_{H^k} = \int_{\mathbb{R}^n} (1 + |\xi|^2)^k |\widehat{f}(\xi)|^2 d\xi < +\infty.$$

Noter au passage que quand $f \in L^2$, $\widehat{f}(\xi)$ est définie presque-partout, donc le membre de droite a un sens.

Du coup, il est facile de définir H^s pour tout $s \geq 0$, en utilisant (2) (en fait, même pour pour tout $s \in \mathbb{R}$). On voyage entre les H^s en appliquant diverses puissances réelles de l'opérateur positif $I - \Delta$, dont l'effet en transformée de Fourier est juste de multiplier $\widehat{f}(\xi)$ par $1 + |\xi|^2$.

Et on peut faire pareil chez les $W^{k,p}(\mathbb{R}^n)$ (au moins pour $1 < p < +\infty$, SVP vérifiez dans les autres cas avant de dire que c'est ma faute); il y a quelques petites vérifications à faire, pour savoir par exemple que $(I - \Delta)^{l/2}$ est un isomorphisme de $W^{k+l,p}$ dans $W^{k,p}$, ou pour vérifier que pour $k = 1$ nos deux définitions de $W^{1,p}$ coincident, ou encore pour définir $W^{k,p}$ pour k non entier.

Deux mots des espaces homogènes aussi. On notera $\dot{W}^{k,p}(U)$ l'espace des fonctions localement intégrables sur U, dont les dérivées d'ordre k sont dans L^p, et muni de la norme "homogène" obtenue en sommant les normes L^p des dérivées d'ordre k. [On ne met rien pour contrôler le caractère L^1_{loc}, mais celui-ci découle des estimations sur les dérivées, comme on le verra de manière implicite plus bas.] Sur $\mathbb{R}^n$ et pour les k non entiers, on utiliserait les puissances de $-\Delta$ au lieu de $I - \Delta$, en faisant un peu plus attention.

Remarque (produit et localisation). Soient $f \in W^{1,p}$ et g de classe C^1 sur U, disons bornée et avec une dérivée bornée. Alors $fg \in W^{1,p}$, avec $\frac{\partial(fg)}{\partial x_i} = \frac{\partial f}{\partial x_i}\, g + \frac{\partial g}{\partial x_i}\, f$.

En effet, le membre de droite est bien dans L^p, donc il s'agit seulement de montrer que pour $\varphi \in C^\infty_c$, $-\int fg\frac{\partial \varphi}{\partial x_i} = -\int f\frac{\partial(g\varphi)}{\partial x_i} + \int \frac{\partial g}{\partial x_i}\, f\,\varphi$. Toutes les intégrales portent sur un compact (le support de φ), et il s'agit seulement d'intégrer l'identité $\frac{\partial(g\varphi)}{\partial x_i} = g\frac{\partial \varphi}{\partial x_i} + \frac{\partial g}{\partial x_i}\varphi$ contre $f \in L^1_{loc}$.

L'intérêt de cette remarque est aussi qu'elle permet de définir l'espace $W^{1,p}_{loc}(U)$ comme étant l'espace des fonctions telles que $fg \in W^{1,p}(U)$ pour $g \in C^1_c(U)$. La dérivée $\frac{\partial f}{\partial x_i}$ est alors la fonction de $L^p_{loc}(U)$ telle que si $K \subset U$ est compact, et $g \in C^1_c(U)$ est égale à 1 sur un voisinage de K, alors $\frac{\partial f}{\partial x_i} = \frac{\partial(fg)}{\partial x_i}$ sur K.

Exemple. Plaçons-nous sur $U =]a,b[\subset \mathbb{R}$. Vérifions d'abord que pour obtenir $f \in W^{1,p}(U)$, il suffit de prendre $g \in L^p(U)$ et $\lambda \in \mathbb{R}$, et de poser

$$f(x) = \lambda + \int_a^x g(t)dt \quad \text{pour } x \in U. \tag{5}$$

Et alors $f' = g$ (au sens des distributions).

En effet, il suffit de voir que $\int_a^b f(x)\varphi'(x)dx = -\int_a^b g(t)\varphi(t)dt$ pour pour $\varphi \in C^\infty_c(U)$. On note d'abord que λ ne contribue ni au membre de gauche (trivialement), ni au membre de droite (parce que $\int \varphi'(x)dx = 0$ puisque φ est à support compact). On applique ensuite le théorème de Fubini pour calculer $\displaystyle\int_U\int_U g(t)\,\varphi'(x)\mathbf{1}_{x>t}\,dxdt$. Quand on intègre d'abord en x, on trouve le membre de droite. Quand on intègre d'abord en t, on trouve le membre de gauche.

Exercice: généraliser ce qui précède à une mesure finie (disons, signée). Il faut quand même faire un peu attention à la contribution des bornes quand on écrit Fubini (mais pas dans la définition de f).

Réciproquement, si $f \in W^{1,p}$, alors il existe λ telle qu'on ait la formule plus haut. En effet, on pose $g = f'$ et $F(x) = \int_a^x g(t)dt$, et on note que

$F \in W^{1,p}$, avec $F' = g = f'$. Donc la dérivée de $F - f$ au sens des distributions est nulle, i.e., $\int_U (f - F)\varphi' = 0$ pour toute fonction test φ. Il reste à vérifier que $G = f - F$ est (égale presque-partout à une) constante, dès qu'elle est localement intégrable et que sa dérivées est nulle. Il y a divers moyens de le vérifier, mais le plus amusant (compte tenu de ce qu'on a fait) sera d'utiliser le théorème de densité de Lebesgue.

Soient x et y deux points de densité de Lebesgue de $G = f - F$, avec $x \neq y$, et vérifions que $G(x) = G(y)$. On en déduira le résultat, en fixant x et en notant que presque-tout point $y \in U$ est un point de densité de Lebesgue, donc tel que $G(y) = G(x)$. On se donne une fonction bosse ψ de classe C^∞ à support dans $[0, 1]$ et d'intégrale 1, on pose $\psi_k(t) = 2^k \psi(2^k t)$ (toujours d'intégrale 1), et ensuite φ_k à support compact, dont la dérivée est $\varphi_k'(t) = \psi_k(t - x) - \psi_k(t - y)$. Le support de φ_k est contenu dans U pour k assez grand, donc on sait que $\int_U G(t)[\psi_k(t - x) - \psi_k(t - y)]dt = 0$. Mais

$$\Big|G(x) - \int_U G(t)\psi_k(t-x)dt\Big| = \Big|\int_{x-t}^{x+t} [G(x) - G(t)]\psi_k(t-x)dt\Big|$$
$$\leq \int_{x-t}^{x+t} \big|G(x) - G(t)\big||\psi_k(t-x)|dt \leq ||g||_\infty \frac{1}{t}\int_{x-t}^{x+t} \big|G(x) - G(t)\big|dt,$$

qui tend vers 0 par la version un peu forte du théorème de densité de Lebesgue (en en prenant la définition correspondante de "point de Lebesgue"). Voir (1) du chapitre 4 sur la différentiation, page 14.

Pareillement, $\int_U G(t)\psi_k(t-x)dt$ tend vers $G(y)$; on trouve $G(x) = G(y)$, et ceci termine notre description par (5) des fonctions de $W^{1,p}$ sur un intervalle borné. Bien sûr, on a noté que pour la discussion ci-dessus, on aurait pu se contenter de regarder le cas où $p = 1$.

Signalons encore que puisqu'on a (1), on peut prouver que f est différentiable presque-partout, avec la dérivée $g = f'$. En fait, si f est un point de Lebesgue pour g, on a

$$\lim_{r \to 0} \frac{f(x+r) - f(x)}{r} = \lim_{r \to 0} \frac{1}{r}\int_x^{x+r} g(t)dt = g(x),$$

et de même $\lim_{r \to 0} \dfrac{f(x-r) - f(x)}{-r} = g(x)$. Mais par contre il existe des fonctions non constantes dont la dérivée existe et est nulle presque-partout (fonction de Lebesgue, dont le graphe est aussi surnomé escalier du diable), qui est obtenue en intégrant une mesure singulière.

Exercice: vérifier qu'en effet la primitive d'une mesure singulière (par rapport à Lebesgue) a bien une dérivée nulle Lebesgue-p.p..

Dirigeons-nous doucement vers Sobolev et Poincaré. On aura aussi besoin d'un lemme d'approximation des fonctions de $W^{1,p}$ par des fonctions régulières.

Lemme 1. *On se donne $U \subset \mathbb{R}^n$, $1 \leq p < +\infty$, et $f \in W^{1,p}_{loc}(U)$. Pour tout compact $K \subset U$, il existe une suite $\{f_k\}$ dans $C_c^\infty(U)$ telle que*

$$\lim_{k\to+\infty} f_k = f \text{ dans } L^p(K) \text{ et, pour } 1 \leq i \leq n, \lim_{k\to+\infty} \frac{\partial f_k}{\partial x_i} = \frac{\partial f}{\partial x_i} \text{ dans } L^p(K).$$

Sans trop de détails: on essaie $f_k = f * \psi_k$, où $\{\psi_k\}$ est une approximation de l'identité, avec support$(\psi_k) \subset B(0, 2^{-k})$. Pour k assez grand pour que $K + B(0, 2^{-k}) \subset\subset U$, f_k est bien définie sur K. Le fait que $\{f_k\}$ converge vers f dans $L^p(K)$ est très classique. Pour la seconde partie, vérifions d'abord que $\frac{\partial f_k}{\partial x_i} = \frac{\partial f}{\partial x_i} * \psi_k$ dans un voisinage de K (et pour k assez grand). On vérifie ceci sur une fonction test φ (de support contenu dans ce voisinage de K, et pour peu que celui-ci reste à distance $> 2^{-k}$ du bord). Le membre de droite donne

$$\begin{aligned}
&\int_U \Big\{ \int_{B(0,2^{-k})} \frac{\partial f}{\partial x_i}(x-u)\psi_k(u)du \Big\} \varphi(x)dx \\
&\qquad = \int_{B(0,2^{-k})} \psi_k(u) \Big\{ \int_U \frac{\partial f}{\partial x_i}(x-u)\varphi(x)dx \Big\} du \\
&\qquad = -\int_{B(0,2^{-k})} \psi_k(u) \int_U f(y) \frac{\partial \varphi}{\partial x_i}(y+u)dydu \\
&\qquad = -\int_U \int_{B(0,2^{-k})} f(y)\psi_k(u) \frac{\partial \varphi}{\partial x_i}(y+u)dudy
\end{aligned}$$

(on a utilisé la définition de $\frac{\partial f}{\partial x_i}$ après avoir vérifié que pour k assez grand, le support de $\varphi(x+u)$ est encore contenu dans U). Le membre de gauche donne

$$\begin{aligned}
-\int_U f_k \frac{\partial \varphi}{\partial x_i} &= -\int_U \int_{B(0,2^{-k})} f(x-u)\psi_k(u) \frac{\partial \varphi}{\partial x_i}(x) \\
&= -\int_U \int_{B(0,2^{-k})} f(y)\psi_k(u) \frac{\partial \varphi}{\partial x_i}(y+u)\, dxdu,
\end{aligned}$$

et on constate que c'est pareil.

Cette vérification faite, on vérifie que $\frac{\partial f_k}{\partial x_i} = \frac{\partial f}{\partial x_i} * \psi_k$ converge bien vers $\frac{\partial f}{\partial x_i}$ dans $L^p(K)$ comme plus haut. Enfin, nos fonctions ne sont pas à support compact, donc on les multiplie par une fonction C^∞ qui vaut 1 dans un voisinage de K mais est à support compact dans U. □

Lemme 2. *On se donne $p \geq 1$, $f \in L^p(V)$, et une suite $\{f_k\}$ dans $W^{1,p}(V)$. On suppose que $f = \lim_{k\to\infty} f_k$ dans $L^p(V)$ et que les $\frac{\partial f_k}{\partial x_i}$ convergent dans $L^p(V)$ vers des fonctions g_i. Alors $f \in W^{1,p}(V)$ et $\frac{\partial f}{\partial x_i} = g_i$.*

Soit $\varphi \in C_c^\infty(V)$. Alors

$$
\begin{aligned}
(3) \quad \Big\langle \frac{\partial f}{\partial x_i}, \varphi \Big\rangle &= -\Big\langle f, \frac{\partial \varphi}{\partial x_i} \Big\rangle = -\lim_{k\to+\infty} \int_V f_k(x)\, \frac{\partial \varphi}{\partial x_i}(x) dx \\
&= \lim_{k\to+\infty} \int_V \frac{\partial f_k}{\partial x_i}(x)\, \varphi(x)\, dx = \int_V g_i(x)\, \varphi(x)\, dx = \langle g_i, \varphi \rangle.
\end{aligned}
$$

Il n'y a pas de problème de convergence, puisque $\frac{\partial \varphi}{\partial x_i}$ et φ sont bornées et à supports compacts dans V.

En fait, le lemme marche encore en supposant seulement que f et les g_i sont dans L^p, que $f = \lim_{k\to\infty} f_k$ faiblement, et que $g_i = \lim_{k\to\infty} \frac{\partial f_k}{\partial x_i}$ faiblement. □

On est à peu près prêts pour essayer de majorer $|f(x) - f(y)|$ en fonction du gradient de f près de x et y. On verra après comment en déduire des inégalités fonctionnelles sur les $W^{1,p}$.

Pour l'instant, supposons que x et y sont donnés et que f est de classe C^1 près de x et y (on aura besoin ce ça dans un petit double cône $C_{x,y}$ autour de x et y).

Quelques notations. On se donne $\alpha \in]0,1]$ pour faire général, mais $\alpha = 1$ devrait suffire. On note

(4)

$$D(x,y) = D_\alpha(x,y) = \Big\{ z \in \mathbb{R}^n : |z-x| = |z-y| \text{ et } \Big| z - \frac{x+y}{2} \Big| \leq \alpha |x-y| \Big\},$$

un bout d'hyperdisque à mi-chemin, et

$$(5) \qquad C(x,y) = C_\alpha(x,y) = \operatorname{conv}\big(\{x\} \cup \{y\} \cup D(x,y)\big)$$

(l'enveloppe convexe), qui est une sorte de double cône (une carotte de bureau de tabac) avec ses pointes en x et y.

On paramètre aussi des chemins γ_ℓ, $\ell \in D(x,y)$: pour $\ell \in D(x,y)$, γ_ℓ est le chemin de x à y obtenu en joignant x à ℓ puis à y par les segments $[x,\ell]$ et $[\ell,y]$. On paramètre γ_ℓ par sa projection sur l'intervalle $[x,y]$, et on note aussi $\gamma_\ell : [x,y] \to C(x,y)$ le paramétrage.

Pour la suite des calculs, on suppose que $C(x,y) \subset U$ et que la fonction f est de classe C^1 sur $C(x,y)$. Les accroissements finis, appliqués à $f \circ \gamma_\ell$, donnent

$$|f(x) - f(y)| \leq \int_{[x,y]} |\nabla f(\gamma_\ell(t))||\gamma'_\ell(t)|dt = \int_{\Gamma_\ell} |\nabla f| d\mathcal{H}^1 \tag{6}$$

où la dernière égalité ne sert qu'à essayer de rendre la chose plus géométrique (on ne l'utilisera pas) et $\Gamma_\ell = [x,\ell] \cup [\ell,y]$. On moyennise cela par rapport à $\ell \in D(x,y)$, et on trouve

$$|f(x) - f(y)| \leq \frac{1}{\mathcal{H}^{n-1}(D(x,y))} \int_{D(x,y)} \int_{[x,y]} |\nabla f(\gamma_\ell(u))|\,|\gamma'_\ell(u)|\,du\,d\ell \tag{7}$$

où l'on a noté indifféremment $d\ell$ et $d\mathcal{H}^{n-1}(\ell)$ la mesure de surface. On fait le changement de variable $(\ell,u) \to \gamma_\ell(u)$ et on trouve

$$|f(x) - f(y)| \leq \int_{C(x,y)} |\nabla f(z)|\theta(z)\,dz, \tag{8}$$

où dz est la mesure de Lebesgue et $\theta(z)$ vient du déterminant Jacobien. Le calcul donne

$$\theta(z) \leq C(n,\alpha)\,w_{x,y}(z), \tag{9}$$

avec

$$w_{x,y}(z) = \mathbf{1}_{C(x,y)}(z)\big\{|z-x|^{1-n} + |z-y|^{1-n}\big\}. \tag{10}$$

Autrement dit, la mesure image de $\mathcal{H}^{n-1}(D(x,y))^{-1} du\,d\ell$ par $(\ell,u) \to \gamma_\ell(u)$ est inférieure à $C(n,\alpha)\,w_{x,y}(z)$.

Remarque. Les calculs suivants donnent des majorations (parce que c'est ce qui nous intéresse ici), mais on aurait pu calculer de manière plus précise, en écrivant vraiment la formule des accroissements finis. On aurait obtenu une formule du type $f(y) - f(x) = \int_{C(x,y)} K_{x,y}(z)\nabla f(z)\,dz$, où $K_{x,y}(z)$ est une matrice qui dépend de z (et avec certaines propriétés d'invariance par rapport à $x-y$). Noter que si on fait tendre y vers l'infini (ici, en partant dans une direction donnée), on obtient une formule qui permet de calculer f à partir de ∇f, par convolution avec un noyau matriciel (et en supposant que f est C^1 à support compact, au moins pour commencer). A cause de la manière dont on s'y est pris (avec la forme de notre cône), la formule n'est pas invariante par rotations, mais en intégrant sur toutes les droites partant de x, on aurait une formule invariante par rotation. A savoir, $f(x) = -c_n \int_{\mathbb{R}^n} \nabla f \cdot \frac{y-x}{|y-x|^n} dy$, et où c_n est la mesure de la sphère unité. [Ca ne peut être que ça, par homogénéité, invariance par dilatations, et en calculant la constante sur des fonctions particulières, par exemple tendant vers la fonction caractéristique d'une sphère.]

Et ce serait la même formule qu'on peut obtenir en partant de ∇f, et en lui appliquant l'opérateur $\Delta^{-1}\nabla$ (à définir plus facilement en transformée de Fourier), qui est bien un opérateur de convolution. Bref, les calculs faits plus hauts sont moins jolis et homogènes, mais ils sont robustes et marchent tout seul dans des domaines.

Il va falloir commencer à distinguer des cas. Si on veut une vraie estimation ponctuelle de $|f(x) - f(y)|$, valable pour $f \in W^{1,p}$ quelconque, il faudra s'assurer que $|\nabla f(z)|$ est intégrable contre $w_{x,y}(z)$, et ça ne marchera bien tel quel que si $p > n$. On va commencer par ce cas-là. Au fait, le cas $n = 1$ est trop facile; voir le cas traité plus haut.

Notons $q = \frac{p}{p-1}$ l'exposant conjugué de p. Si $p > n$, alors $q < \frac{n}{n-1}$, le poids $|z|^{q(1-n)}$ est localement integrable, et

(11)
$$||w_{x,y}||_q \leq 2\Big\{\int_{|z-x|\leq|x-y|} |z-x|^{q(1-n)}\Big\}^{1/q} \leq C\,|x-y|^{(1-n)+\frac{n}{q}} = C\,|x-y|^{\frac{p-n}{p}}$$

(ce qui sera utile pour la normalisation). Finalement Hölder et (6) donnent

(12)
$$|f(x) - f(y)| \leq C\,||\nabla f||_{L^p(C(x,y))}||w_{x,y}||_q \leq C\,|x-y|^{\frac{p-n}{p}}\Big\{\int_{C(x,y)} |\nabla f|^p\Big\}^{1/p}.$$

Proposition 1. *Soient U un ouvert de $\mathbb{R}^n$, $p \in]n, +\infty]$, et $f \in W^{1,p}_{loc}(U)$. Soient $x, y \in U$, tels que de double cône $C(x,y)$ soit contenu dans U. Alors*

$$|f(x) - f(y)| \leq C\,|x-y|^{\frac{p-n}{p}}\Big\{\int_{C(x,y)} |\nabla f|^p\Big\}^{1/p}. \tag{13}$$

Ici C ne dépend que de n, p, et α (l'ouverture). En fait, il faudrait dire que f est égale presque-partout à une fonction continue, qui vérifie (13).

Pour démontrer ceci, on se donne f, et on l'approxime par des fonctions régulières f_k comme au lemme 1. A cause de la façon dont on a écrit le lemme, où la suite $\{f_k\}$ dépend du compact, on va s'y reprendre à deux fois.

D'abord, on se donne une boule fermée $B = \overline{B}(x_0, r)$ contenue dans U, et on construit les f_k comme au lemme 1. Quitte à extraire une sous-suite, on peut supposer que la suite $\{f_k\}$, qui converge déja dans $L^1(B)$ vers f, converge aussi presque-partout. Noter que pour $x, y \in B_1 = B(x_0, r/10)$, le double cône $C(x, y)$ est contenu dans $B \subset U$, donc on peut appliquer (12) aux f_k et il vient

$$(14) \qquad |f_k(x) - f_k(y)| \leq C\,|x-y|^{\frac{p-n}{p}} \Big\{ \int_B |\nabla f_k|^p \Big\}^{1/p} \leq A|x-y|^{\frac{p-n}{p}}$$

pour un nombre $A < +\infty$ (car les ∇f_k convergent dans $L^p(B)$, donc leur norme est bornée). Si $f(x)$ est la limite des $f_k(x)$ (ce qui arrive presque-partout), et pareil pour $f(y)$, on en déduit que $|f(x) - f(y)| \leq A|x-y|^{\frac{p-n}{p}}$. Ceci vaut pour x, y dans une partie de mesure pleine de B_1. On en déduit la continuité de f, après changement sur une partie de mesure nulle.

Maintenant on peut supposer f continue, on fixe x et y comme dans l'énoncé, et on applique le lemme 1 avec le compact $C(x, y)$. La démonstration du lemme donne aussi la convergence uniforme des f_k vers f, et (13) se déduit de (12), et du fait que les ∇f_k convergent vers ∇f dans $L^p(C(x, y))$. $\square$

Passons très brièvement au cas où $p = n$. Prenons la fonction $f(x) = \ln(\ln(\frac{1}{|x|}))$, au voisinage de 0. Alors $|\nabla f(x)| = \frac{1}{|x|\ln(\frac{1}{|x|})}$, qui est bien dans L^n près de 0 (car $n \geq 2$). Donc $f \in W^{1,n}_{loc}$, et pourtant elle n'est ni continue ni bornée. Mais f est quand même exponentiellement intégrable quand $f \in W^{1,n}$; voir le chapitre sur BMO, la seconde remarque juste après le théorème de John et Nirenberg.

Quand $p < n$, f n'est pas continue, mais par contre elle est localement dans des L^q meilleurs que L^p. Le plus pratique pour écrire ceci semble être en termes d'inégalités de Poincaré, qui sont une manière de dire comment $f \in L^q$ localement. Si on voulait des propriétés globales sur tout $\mathbb{R}^n$, on serait facilement canulé par des questions d'homogénéité. [Sauf peut-être justement pour $p = n$.]

On va utiliser les notations suivantes pour simplifier: $|B|$ est la mesure de Lebesgue de B et, si $f \in L^1(B)$, $m_B f = \frac{1}{|B|}\int_B f(x)dx$ est la moyenne de f sur B.

Proposition 2. *Soient $p \in [1,n]$ et $r \in [1, \frac{np}{n-p}[$. Il existe une constante $C(n,p,r)$ telle que pour toute boule ouverte B et tout $f \in W^{1,p}(B)$,*
(15)
$$\left\{\frac{1}{|B|}\int_B |f(x) - m_B f|^r\right\}^{1/r} \leq C(n,p,r)\,|B|^{1/n}\left\{\frac{1}{|B|}\int_B |\nabla f(x)|^p\right\}^{1/p}.$$

Quelques commentaires avant de commencer la démonstration.

Quand $p = n$, tout $r < +\infty$ est autorisé. Et même, f est localement exponentiellement intégrable, comme on le verra au prochain chapitre sur BMO. Quand $p > n$ tout $r < +\infty$ est autorisé aussi, mais la Proposition 1 donne directement une borne sur $|f(x) - m_B f|$ qui est meilleure que (15). Donc on on s'est carrément mis dans le cas où $p \leq n$, et on n'a rien perdu.

J'ai mis les puissances de manière à simplifier la vérification d'homogénéité. Il est logique que chaque membre soit homogène de degré 1 en f. Pour ce qui est de l'homogénéité en fonction du rayon R de B, penser que quand $|\nabla f(x)|$ est de l'ordre de 1, les deux membres sont en principe de l'ordre de R ou $|B|^{1/n}$.

Noter que par Hölder, l'inégalité pour r entraîne celle pour $r' < r$. C'est pourquoi on pourra se contenter du cas où $r > p$ (noter que $\frac{np}{n-p} > p$).

On va commencer par le cas où f est de classe C^1, il faudra faire un petit passage à la limite à la fin.

Comme l'homogénéité est correcte, on pourrait se ramener facilement au cas où $B = B(0,1)$, mais on va essayer de faire directement le cas général où $B = B(X,R)$.

Posons $B_1 = B(X, R/10)$, et notons que quand $x \in B$ et $y \in B_1$, alors le bi-cône $C(x,y)$ est contenu dans B. Alors (8) et (9) disent que

$$|f(x) - f(y)| \leq C\int_{C(x,y)} |\nabla f(z)|\, w_{x,y}(z)dz \tag{16}$$

On fait la moyenne par rapport à $y \in B_1$ et on trouve

$$|f(x) - m_{B_1} f| \leq CR^{-n}\int_{y\in B_1}\int_{z\in C(x,y)} |\nabla f(z)|\,\omega_{x,y}(z)\,dzdy, \tag{17}$$

On se souvient que $w_{x,y}(z) = \mathbf{1}_{C(x,y)}(z)\Big\{|z-x|^{1-n}+|z-y|^{1-n}\Big\}$ (par (10)), et quand on intègre ceci par rapport à $y \in B_1$, on trouve moins que $CR^n(R^{1-n}+|z-x|^{1-n})$. Donc

$$(18)\qquad |f(x) - m_{B_1} f| \leq C \int_{z\in B} |\nabla f(z)|\,(R^{1-n} + |z-x|^{1-n})dz.$$

On pose $h(u) = (R^{1-n} + |u|^{1-n})\mathbf{1}_{B(0,2R)}(u)$, et on garde (18) sous la forme synthétique

$$(19)\qquad |f(x) - m_{B_1} f| \leq C\,[(\mathbf{1}_B\,|\nabla f|) * h](x).$$

Puis on se souvient que la convolution envoie $L^p \times L^q$ dans L^r, où $\frac{1}{p}+\frac{1}{q} = \frac{1}{r}+1$. Ici, on a déjà p (puisque $\nabla f \in L^p(B)$) et r (avec $p < r < \frac{np}{n-p}$). Si on prenait $r = \frac{np}{n-p}$, on trouverait $\frac{1}{q} = \frac{n-p}{np} + 1 - \frac{1}{p} = \frac{n-1}{n}$ et $q = \frac{n}{n-1}$. Ici on prend r plus petit, donc q plus petit. Bref, $q < \frac{n}{n-1}$, ce qui tombe bien, parce que cela implique que h est localement dans L^q. Noter encore que $r = p$ donne $q = 1$, donc on est dans des valeurs acceptables de q.

Retournons à (19). On trouve

$$(20)\qquad ||f(x) - m_{B_1} f||_r \leq C\,||h||_q\,||\nabla f||_{L^p(B)}$$

et en plus

$$(21)\qquad ||h||_q \leq C\Big\{R^{(1-n)q}R^n\Big\}^{1/q} = R^{1-n+\frac{n}{q}}.$$

C'est ce qu'il fallait pour (15), modulo deux choses. D'abord, la vérification de l'homogénéité en R. La puissance de R est $1-n+\frac{n}{q} = 1-n+n(1+\frac{1}{r}-\frac{1}{p}) = 1+\frac{n}{r}-\frac{n}{p}$. C'est exactement ce qu'il faut puisque $|B|$ est de l'ordre de R^n.

La seconde chose est qu'on a remplacé $m_B f$ par $m_{B_1} f$. Mais

(22)

$$\begin{aligned}|m_B f - m_{B_1} f| &\leq \frac{1}{|B|}\int_B |f(x) - m_{B_1} f|\,dx \leq \left\{\frac{1}{|B|}\int_B |f(x) - m_B f|^r\right\}^{1/r}\\ &\leq C(n,p,r)\,|B|^{1/n}\left\{\frac{1}{|B|}\int_B |\nabla f(x)|^p\right\}^{1/p}.\end{aligned}$$

par inégalité triangulaire, puis Hölder, puis ce qu'on vient de montrer; le remplacement de $m_B f$ par $m_{B_1} f$ ne coûte donc pas plus que le second membre.

Notons encore qu'en fait, dans les calculs ci-dessus, on peut remplacer $B_1 = B(X, R/10)$ par n'importe quelle boule $B(X, R_1)$, avec $R/11 \leq R_1 \leq R/9$. Soit en notant que la démonstration marche encore, soit en faisant comme pour (22).

On a presque fini. On a vérifié (15), mais seulement pour les fonctions f de classe C^1. Dans le cas général, on se donne une boule $B' = B(X, R') \subset B$, avec R' juste un peu plus petit que R, et on commence par utiliser le lemme 1 pour trouver une suite de fonctions f_k de classe C^1, qui convergent vers f dans $L^p(B')$, et telles que les ∇f_k convergent vers ∇f dans $L^p(B')$. Quitte à extraire une sous-suite on peut même s'arranger pour que $f_k(x)$ tende vers $f(x)$ pour presque tout $x \in B'$. On sait que

$$\left\{ \int_{B'} |f_k(x) - m_{B_1'} f_k|^r \right\}^{1/r} \leq C \left\{ \int_{B'} |\nabla f_k|^p \right\}^{1/p}, \tag{23}$$

où en principe on devrait prendre $B_1' = B(X, R'/10)$, mais où, à cause de la remarque précédente sur $R/11 \leq R_1 \leq R/9$, on peut carrément garder $B_1' = B_1$ (ce qui simplifie un peu le passage à la limite ci-dessous).

On fait tendre k vers $+\infty$, et on trouve que

$$\left\{ \int_{B'} |f(x) - m_{B_1'} f|^r \right\}^{1/r} \leq C \left\{ \int_{B'} |\nabla f|^p \right\}^{1/p} \leq C \left\{ \int_{B} |\nabla f|^p \right\}^{1/p}, \tag{24}$$

ou l'on a utilisé Fatou pour l'intégrale de gauche. On fait tendre R' vers R et on trouve l'analogue de (15) avec $m_{B_1} f$. Et ensuite on remplace $m_{B_1} f$ par $m_B f$ comme avant, avec (22). La proposition s'en déduit. $\square$

Exercice. Montrer que, sous les hypothèses de la proposition 2, on a aussi

$$\left\{ \frac{1}{|B|^2} \int_B \int_B |f(x) - f(y)|^r \right\}^{1/r} \leq C(n, p, r)\, |B|^{1/n} \left\{ \frac{1}{|B|} \int_B |\nabla f(x)|^p \right\}^{1/p}. \tag{25}$$

Exercice (Au cas où, mais c'est peut-être un peu délicat sans questions intermédiaires; il faut passer d'une boule à une autre par une chaine des boules en utilisant la connexité). Soient $\Omega \subset \mathbb{R}^n$ un domaine connexe, $p \geq 1$, et $\{f_k\}_{k \geq 0}$ une suite dans $W^{1,p}(\Omega)$. On suppose que

$$\{\nabla f_k\} \text{ converge dans } L^p(\Omega) \text{ vers une limite } h \in L^p(\Omega), \tag{26}$$

et qu'on peut trouver une boule $B \subset \Omega$ telle que

$$\lim_{k \to +\infty} m_B f_k \text{ existe.} \tag{27}$$

Montrer que $\{f_k\}$ converge dans $L^1_{loc}(\Omega)$ vers une limite f, que $f \in W^{1,p}(\Omega)$, et que $\nabla f = h$.

6b. RESTRICTIONS, TRACES, ET POSSIBLEMENT EXTENSIONS

Chapitre ajouté en 2016, donc attention aux fautes de frappe et erreurs graves.

On commence par ce que je vais appeler restrictions, faute de mieux peut-être, qui consiste à prendre une fonction $f \in W^{1,p}$, et demander si en moyenne sa restriction à un plan est aussi dans $W^{1,p}$, avec une norme moyenne bornée. Ceci va être facile, mais pas forcément inutile. Ensuite on parlera d'un problème a priori plus délicat, qui est de donner un sens à la restriction, mais je dirai trace, de f, sur un plan donné à l'avance (et donc qui pourrait être exceptionnel, je veux dire mauvais, pour le probeème précédent.

Les notations pour ce qui suit sont les suivantes. On se donne un ouvert Ω_1 de $\mathbb{R}^m$ et un ouvert Ω_2 de $\mathbb{R}^n$, et une fonction $f \in W^{1,p}(\Omega_1 \times \Omega_2)$, et on va voir que pour presque tout $x \in \Omega_1$, la fonction $f(x, \cdot)$ est bien dans $W^{1,p}(\Omega_1)$, avec les dérivées partielles escomptées. En fait, toute la difficulté est dans le cas où $p = 1$, auquel on va se restreindre dans un premier temps.

Voici l'énoncé. Notons encore $x = (x_1, \dots, x_m)$ le point courant de $\Omega_1 \subset \mathbb{R}^m$ et $y = (y_1, \dots, y_m)$ le point courant de $\Omega_2 \subset \mathbb{R}^n$.

Proposition 1. *On se donne $\Omega = \Omega_1 \times \Omega_2$ comme ci-dessus, et une fonction $f \in L^1(\Omega)$, et on suppose que la dérivée partielle $\frac{\partial f}{\partial y_1}$ (au sens des distribution) est dans $L^1(\Omega)$. Alors, pour presque tout $x \in \Omega_1$, la fonction $f(x, \cdot)$ a une dérivée partielle par rapport à y_1 qui est dans $L^1(\Omega_2)$, et qui est égale à $\frac{\partial f}{\partial y_1}(x, \cdot)$.*

Quelques remarques avant de commencer. Dans cet énoncé, le plus simple est de prendre un représentant de f et un de $g = \frac{\partial f}{\partial y_1}$, et de vérifier que la conclusion vaut presque partout, au sens ou pour presque tout $x \in \Omega_1$, $g(x, \cdot) \in L^1(\Omega_1)$ et représente la dérivée partielle de $f(x, \cdot)$.

Le fait que presque tout $x \in \Omega_1$, $g \in L^1(\Omega_1)$ est juste une application de Fubini, donc le problème est bien dans le calcul de dérivée partielle. Ainsi, si on avait supposé mieux (disons que notre domaine est borné), à savoir que $\frac{\partial f}{\partial y_1} \in L^p(\Omega)$, on aurait trouvé aussitôt que pour presque tout $x \in \Omega_1$, la dérivée partielle $g(x, \cdot)$ est dans $L^p(\Omega_1)$, avec l'estimation $\int_{\Omega_1} ||g(x, \cdot)||_{L^p(\Omega_1)}^p dx = ||g||_{L^p(\Omega)})^p$. Donc on se contente de $p = 1$.

Le problème est local: il faut juste être capable d'identifier les dérivées partielles des restrictions. Donc l'énoncé avec $W^{1,p}_{loc}(\Omega)$ est vrai aussi, et se déduit du précédent.

Il y a une réciproque facile (mais moins utile sans doute); voir plus bas.

Passons à la démonstration. Continuons de noter $g = \frac{\partial f}{\partial y_1}$. Notons aussi f^x la fonction $y \to f(x,y)$ et $g^y : y \to g(x,y)$. L'hypothèse est d'abord que $|f| + |g| \in L^1(\Omega)$, donc

$$\int_\Omega |f| + |g| dxdy < +\infty \tag{1}$$

et il existe $Z \subset \Omega_1$ de mesure nulle, tel que pour $x \in \Omega_1 \setminus Z$,

$$\int_{\Omega_2} |f^x| + |g^x| dy < +\infty. \tag{2}$$

Ensuite, $g = \frac{\partial f}{\partial y_1}$ donc pour toute fonction test ϕ sur Ω,

$$\int_\Omega f \frac{\partial \phi}{\partial y_1} = - \int_\Omega g\phi. \tag{3}$$

On s'intéresse à des fonctions décomposées $\phi(x,y) = \varphi(x)\psi(y)$, où $\varphi \in C_c^\infty(\Omega_1)$ et $\psi \in C_c^\infty(\Omega_2)$. Alors puisque $\frac{\partial \phi}{\partial y_1}(x,y) = \varphi(x)\frac{\partial \psi}{\partial y_1}(y)$, (2) s'écrit, en utilisant (1) et le fait que ϕ et sa dérivée sont bornées pour justifier Fubini
(4)

$$\int_{x\in\Omega_1} \Big\{\varphi(x) \int_{y\in\Omega_2} f(x,y)\frac{\partial \psi}{\partial y_1}(y)dy\Big\}dx = - \int_{x\in\Omega_1} \Big\{\varphi(x) \int_{y\in\Omega_2} g(x,y)\psi(y)dy\Big\}dx.$$

Fixons ψ et posons

$$A(x) = \int_{y\in\Omega_2} \Big[f(x,y)\frac{\partial \psi}{\partial y_1}(y) + g(x,y)\psi(y)\Big]dy. \tag{5}$$

C'est une fonction intégrable (toujours par (2)), et son intégrale contre n'importe quelle fonction test φ est nulle(ou autrement dit elle définit une distribution nulle). On a dit plus haut (et peut-être laisseé en exercice) que dans ce cas $A(x) = 0$ presque partout. Bref, pour tout $\psi \in C_c^\infty(\Omega_2)$, il existe un ensemble Z_ψ de mesure nulle tel que

$$\int_{y\in\Omega_2} f(x,y)\frac{\partial \psi}{\partial y_1}(y)dy = - \int_{y\in\Omega_2} g(x,y)\psi(y)dy \tag{6}$$

Un petit ennui pas grave est que l'on voudrait un ensemble de mesure nulle hors duquel (6) ait lieu pour tout $\psi \in C_c^\infty(\Omega_2)$. Donc on se donne une famille dénombrable $\{\psi_i\}$ de fonctions de $\mathcal{D} = C_c^\infty(\Omega_2)$, et qui est dense dans $\mathcal{D}$, muni de la norme (qu'on notera $||\cdot||_a$) de la convergence uniforme, pour ψ et sa première dérivée. Pour trouver ceci, il s'agit simplement, pour chaque compact $K \subset \Omega_2$, de trouver une famille dénombrable de fonction de $\mathcal{D}$, telle que pour tout $\psi \in \mathcal{D}$ à support dans K et tout $\varepsilon > 0$, il existe η dans la famille telle que $||\psi - \eta||_a \leq \varepsilon$. C'est pas difficile, mais un peu pénible; par exemple des fonctions polynômes à conficients rationnels, et correctement aplaties pour être nulles loin de K, doivent faire l'affaire. Revenons à $\{\psi_i\}$, et choisissons pour Z' l'union du Z de (2), et des Z_{ψ_i} pour cette famille. Pour $x \in \Omega_1 \setminus Z'$, et tout i, on a (6) pour tout ψ_i. Mais pour cet x, f^x et g^x sont intégrables (car $x \notin Z$), donc par convergence uniforme, (6) est encore vrai pour tout $\varphi \in \mathcal{D}$. Bref, g^x est la dérivée partielle de f^x, comme souhaité. □

On a promis une réciproque facile (rien de plus que la définition et Fubini, comme on va le voir). La voici.

Proposition 2. *Soient $\Omega = \Omega_1 \times \Omega_2$ comme ci-dessus, et deux fonctions $f, g \in L^1(\Omega)$. Notons $f^x(y) = f(x,y)$ et $g^x(y) = g(x,y)$. Ainsi, par Fubini, f^x et g^x sont dans $L^1(\Omega_2)$ pour presque tout $x \in \Omega_1$. Supposons que $g^x = \frac{\partial f^x}{dy_1}$ pour presque tout $x \subset \Omega_1$. Alors $g = \frac{\partial f}{dy_1}$ (le tout au sens des distributions).*

Notons qu'à nouveau, il ne s'agit pas ici de voir dans quelle espace L^p se trouve g en moyenne (ça, c'est Fubini), mais juste que g est la dérivée de f. On se donne donc une fonction test $\phi \in C_c^\infty(\Omega)$, et on calcule:

$$\begin{aligned}
\langle \frac{\partial f}{dy_1}, \phi \rangle &= -\int_\Omega f \frac{\partial \phi}{dy_1} = -\int_{\Omega_1} \Big\{ \int_{\Omega_2} f(x,y) \frac{\partial \phi}{dy_1}(x,y) dy \Big\} dx \\
&= -\int_{\Omega_1} \Big\{ \int_{\Omega_2} f^x \frac{\partial \phi^x}{dy_1}(y) dy \Big\} dx = \int_{\Omega_1} \Big\{ \int_{\Omega_2} g^x \phi^x dy \Big\} dx \\
&= \int_{\Omega_1} \Big\{ \int_{\Omega_2} g(x,y) \phi(x,y) dy \Big\} dx = \langle g, \phi \rangle
\end{aligned}$$

où les calculs d'intégrales sont justifiés par Fubini (et le fait que f, $g \in L^1$), et on a utilisé le fait que g^x est la dérivée partielle de f^x au milieu du calcul. □

Revenons à la proposition 1. Le plus souvent on l'applique avec $n = 1$, donc par exemple à un produit $\Omega_1 \times I$, où I est un intervalle. On trouve que pour presque tout x, f^x est une fonction de $W^{1,1}(I)$, ce qu'on appelle une fonction absolument continue sur l'intervalle: f^x est l'intégrale indéfinie d'une fonction g^x intégrable sur I, comme on l'a vu au début du chapitre sur $W^{1,p}$. On résume ceci en disant que f est absolument continue sur les lignes (mais on devrait dire absolument continue sur presque toutes les lignes).

Comme on a vu aussi que quand $f^x \in W^{1,1}(I)$, la fonction f^x a une dérivée en presque tout point, on en déduit (par Fubini) que pour presque tout point $(x, y) \in \Omega_1 \times I$, f a une dérivée partielle (cette fois, au sens usuel) $\frac{\partial f}{\partial y}$ au point x, y.

En appliquant ceci plusieurs fois dans des directions différentes, on voit que pour $\Omega \subset \mathbb{R}^n$ et $f \in W^{1,1}(\Omega)$, la fonction f a des dérivées partielles, dans les directions des axes, en presque tout point. Du coup, en changeant de coordonnées, aussi dans toute direction d'une famille dénombrable donnée à l'avance.

C'est assez rigolo quand on pense que f n'est pas forcément continue. Dans l'autre sens, on verra plus tard (au chapitre de différentiabilité presque partout) que si $f \in W^{1,p}$ avec $p > n$, elle est différentiable presque partout (donc dans toutes les directions à la fois, et avec un peu d'uniformité sur la direction).

Revenons maintenant à la thématique "restriction à presque tout plan", opposée à "**trace** sur un plan donné". Le sujet des traces est bien plus général que ce qu'on va décrire ici, où l'on va juste donner un exemple en entier et un autre en filigrane. Prenons le domaine $\Omega = B \times I$, où $B = B(0, 1) \subset \mathbb{R}^n$ et $I = [0, 1]$. On se donne $F \in W^{1,1}(\Omega)$, on se demande quel sens donner à la fonction trace (ou dans ce cas, valeur au bord) de F, sur le bord $B \simeq B \times \{0\}$.

Mais on aurait pu aussi se poser la question de la valeur à donner à F sur $B \times \{1/2\}$, qui est en plein milieu du domaine. Sachant qu'on ne peut se contenter de la définition de F presque partout, puisque $B \times \{1/2\}$ est de mesure nulle.

Utilisons l'absolue continuité sur les lignes. Pour presque tout $x \in B$, la fonction $y \to F(x, y)$ est dans $W^{1,1}(I)$, donc pour ces x il existe

$$f(x) = \lim_{y \to 0} F(x, y). \tag{7}$$

C'est une définition raisonnable de la trace (définie p.p. sur B). Mais quelques observations donnent un peu de bonne conscience supplémentaire.

D'abord, on pourrait choisir de prendre une autre direction que la direction verticale, et de prendre les limites suivant cette nouvelle direction oblique. Il se trouve, et ce n'est pas trop difficile à vérifier, que la trace obtenue serait la même. Ce serait encore mieux si F avait des limites non tangentielles (quand on tend vers le point $(x, 0)$ en restant dans un cône du type $\{(u, y); |u - x| \leq \alpha y\}$), mais c'est probablement faux (j'avoue, je n'ai pas vérifié). Sauf si on a des informations supplémentaires sur F (par exemple, savoir que F est harmonique).

Le procédé ci-dessous marche encore si la frontière, au lieu d'être le morceau de plan lisse B, était le graphe d'une fonction Lipschitzienne $A : B \to I$. D'ailleurs, un changement de variable bi-Lipschitzien préserve $W^{1,2}$, donc à la place de limites radiales ci-dessus, on pourrait avoir convergence le long d'autres réseaux (ou devrais-je dire fibrés) de courbes vaguement parallèles qui terminent sur B.

Enfin, on a un certain contrôle sur la trace f ainsi définie.

Lemme. *Avec les notations ci-dessus, et en notant aussi m la moyenne de F sur $B \times I$,*

$$\int_B |f(x) - m| dx \leq C \int_{B \times I} |\nabla F(x, y)| dx dy. \tag{8}$$

Ainsi, si par exemple on se restreint à l'hyperplan H de $W^{1,1}(B \times I)$ composé des fonctions de moyenne nulle, nous avons défini par (7) un opérateur de trace linéaire, à valeurs dans $L^1(B)$, et qui est borné. On peut faire plus précis, au sens où il y a des fonctions de L^1 qui ne sont la trace d'aucun $F \in W^{1,1}(B \times I)$.

D'ailleurs, une autre manière de définir la trace, juste apparemment un peu moins constructive, et qui finalement est à peu près équivalente, aurait été de dire que pour les fonctions lisses, la trace est bien définie, puis de noter que cette trace bien définié est un opérateur linéaire borné, en metant des normes convenables sur l'espace de départ (ici, $W^{1,1}$, avec bien sûr sa norme (ou la norme homogène où on n'ajoute même pas la norme L^1)), et l'espace d'arrivée (ici, on décide de prendre une norme L^1, modulo une constante additive, et probablement qu'on aurait pu prendre une norme un peu plus forts). Et en étendant l'opérateur de trace par continuité.

La démonstration du lemme est assez facile. On utilise la continuité absolue de F sur le segment vertical $L_x = \{x\} \times I$ (la proposition 1 ci-dessus), qui dit que

$$\sup_{y \in I} |f(x) - F(x, y)| \leq \sup_{y \in I, y' \in I} |F(x, y') - F(x, y)| \leq \int_{y \in I} \Big| \frac{\partial F}{\partial y}(x, y) \Big| dy. \tag{9}$$

Ensuite,

$$(10) \quad \int_B |f(x)-m|dx \leq \int_{B\times I} \{|f(x)-F(x,y)|+|F(x,y)-m|\}dxdy = A+A'$$

avec

(11)

$$A \leq \int_{x\in B} \sup_{y\in I} |f(x) - F(x,y)|dx \leq \int_{x\in B}\int_{y\in I} \Big|\frac{\partial F}{\partial y}(x,y)\Big|dy = \int_{B\times I}\Big|\frac{\partial F}{\partial y}\Big|$$

à cause de (9), et

$$(12) \qquad A' \leq C\int_{B\times I} |\nabla F|$$

par Poincaré sur le produit $B\times I$. Noter que Poincaré sur $B\times I$ se démontre comme Poincaré sur une boule. Si vous n'aimez pas cet argument, vous pouvez aussi commencer par travailler sur $B\times\{y\}$, surlaquelle la proposition 1 dit que F est aussi dans $W^{1,1}$ pour presque tout y, de sorte que si $m(y) = \fint_B F(x,y)dx$, Poincaré sur une boule dit que

$$(13) \qquad \int_B |F(x,y) - m(y)|dx \leq C\int_B |\nabla F(x,y)|dx.$$

Puis

(14)

$$A' = \int_{B\times I} |F(x,y)-m| \leq \int_{B\times I} |F(x,y)-m(y)| + \int_{B\times I} |m(y)-m| = A_1'+A_2'$$

avec

$$(15) \qquad \begin{aligned} A_1' &= \int_{y\in I}\Big\{\int_B |F(x,y) - m(y)|dx\Big\}dy \\ &\leq C\int_{y\in I}\Big\{\int_B |\nabla F(x,y)|dx\Big\}dy = C\int_{B\times I}|\nabla F| \end{aligned}$$

par (13), et, puisque m est aussi la moyenne des $m(y)$

$$(16) \qquad A_2' \leq \sup_{y,y'\in I} |m(y) - m(y')|.$$

Or pour $0 < y' < y$,
(17)
$$\begin{aligned}|m(y) - m(y')| = \Big|\int_{x\in B} F(x,y) - F(x,y')dx\Big| \leq \int_{x\in B} |F(x,y) - F(x,y')|dx \\ \leq \int_{x\in B}\int_{y}^{y'} \Big|\frac{\partial F}{\partial y}(x,t)\Big| dtdx \leq \int_{B\times I}\Big|\frac{\partial F}{\partial y}(x,t)\Big|\end{aligned}$$

où l'on a encore utilisé l'absolue continuité de F sur presque toute ligne verticale. On additionne tout et on retrouve (12), puis le lemme. □

Notons quand même qu'en appliquant le lemme, on a perdu une dérivée. On aurait pu faire mieux, mais pas au point de garder une dérivée de f sur B. Avec 2 dérivées, on aurait facilement trouver une trace sur un disque de dimension encore un de moins, mais à nouveau ceci n'est pas optimal. En fait, la règle de pouce (pardon pour l'anglicisme) est qu'on pert en gros une demie dérivée par dimension d'espace.

J'aurais peut-être du dire plus tôt que le théorème de Sobolev est aussi un théorème de trace, puisqu'il permet de définir la valeur de f en un point, sachant seulement que $f \in W^{1,p}_{loc}$ pour un $p > n$.

Il y a des cas où on a une caractérisation exacte des traces. Par exemple, si on part de $f \in W^{1,2}(B\times I)$, on trouve une trace dans l'espace $H^{1/2}(B)$ des fonctions avec une demie dérivée dans $L^2(B)$, dont une définition équivalente est que la quantité suivante est finie:

$$||f||^2_{1/2} = \int_{B\times B} \frac{|f(x) - f(y)|^2}{|x-y|} dxdy. \tag{18}$$

Par description exacte, j'entends que dans ce cas il y a aussi un opérateur linéaire d'extension, qui à $f \in H^{1/2}(B)$ associe une fonction $F \in W^{1,2}(B\times I)$, et dont la trace est f.

Si j'ai un peu de temps et de courage, j'essaierai de donner une indication de comment on fait tout ceci. Pour la trace, c'est un argument comme la démonstration de Poincaré, où l'on estime $|f(x) - f(y)|$ en intégrant ∇f sur des chemins contenus dans $C(x,y)$ et qui passent par $B \times I$ (et le plus simple est sans doute de commencer à démontrer les inégalités sur les fonctions lisses, pour ne pas avoir à essayer d'utiliser de l'absolue continuité sur des segments inclinés). Ensuite il reste pour estimer $||f||_{1/2}$ à utiliser Hölder et Fubini de la bonne manière.

Pour l'existence d'une bonne extension, les méthodes avec cubes de Whitney sont une bonne idée, mais penser à choisir des constantes qui sont plutôt des moyennes de valeurs de f.

Mais je voulais surtout insister sur le fait qu'on peut définir des traces raisonnables, à condition d'accepter de perdre de la régularité (ou plutôt, de partir avec assez de régularité).

7. BMO ET LE THÉORÈME DE JOHN ET NIRENBERG

La chose qui m'intéresse pour ce chapitre est le théorème de John et Nirenberg, et sa démonstration à l'aide de cubes dyadiques et d'un argument de temps d'arrêt. Il n'est pas exclu que l'on revoie BMO plus tard, mais ça n'est pas trop la question pour l'instant.

On va se placer sur $BMO(\mathbb{R}^n, dx)$, mais la notion a un sens sur les espaces de type homogènes, et je crois que le théorème marche pareil (sauf si j'oublie une petite hypothèse); en particulier, il y a souvent une notion de cubes dyadiques.

Definition. *Soit f une fonction localement intégrable (sur $\mathbb{R}^n$ muni de dx). On dit que f est dans BMO (à oscillation moyenne bornée) si*

$$||f||_{BMO} = \sup_B \frac{1}{|B|} \int_B |f(x) - m_B f| dx < +\infty,$$

où le sup est pris sur les boules ouvertes (disons mais ça ne change rien). Et comme toujours $m_B f = \frac{1}{|B|} \int_B f(x) dx$.

Quelques remarques:

• Si $f \in L^\infty$, alors $f \in BMO$, avec $||f||_{BMO} \leq 2||f||_\infty$ (je mets 2 pour être certain de ne pas me tromper, mais je parierais que 1 marche aussi). Le comportement de BMO vis-à-vis de l'homogénéité pour l'effet des translations et dilatations est le même que celui de L^∞. D'ailleurs, BMO apparaît souvent comme le substitut un peu plus gros de L^∞ quand celui-ci est juste un peu trop petit.

• Si f est constante, $||f||_{BMO} = 0$. De fait, BMO (et il faudrait plutôt dire son quotient par les constantes) est un espace de Banach de fonctions définies modulo une l'addition d'une constante. Mais quand même, quand je dirai “soit $f \in BMO$,” je penserai qu'un représentant a été choisi.

• Il est assez facile (mais un peu laborieux) de vérifier que $\ln(|x|) \in BMO$. C'est en fait la fonction typique de BMO pour son comportement (taille, type de singularités). Noter que $|f| \in BMO$ si $f \in BMO$ (inégalité triangulaire), mais la réciproque est fausse (par exemple $\text{sign}(x) \ln(|x|) \notin BMO(\mathbb{R})$).

• Un autre exemple (en fait un peu moins convaincant): si $f \in W^{1,n}(\mathbb{R}^n)$, alors $f \in BMO$. En effet, l'inégalité de Poincaré (la proposition 2 du paragraphe précédent), déjà avec $r = 1$, donne $\frac{1}{|B|} \int_B |f(x) - m_B f| dx \leq$

$C(n,n,1)\left\{\int_B |\nabla f|^n\right\}^{1/n} \leq C||f||_{W^{1,n}}$. Bref, les puissances de $|B|$ au second membre disparaissent miraculeusement. Mais on pert pas mal dans cette estimation, et $f \in W^{1,n}(\mathbb{R}^n)$ est loin d'être l'exemple typique de $f \in BMO$.

Par ailleurs, je devrais plutôt utiliser l'espace homogène $W^{1,n}$ avec un point dessus, composé des fonctions $f \in L^1_{loc}$ telles que $\nabla f \in L^p$. En signalant deux ou trois petites choses que j'aurais pu dire plus tôt. D'abord, le lemme d'approximation reste vrai pour les fonctions de l'espace $W^{1,p}(\Omega)$ homogène, sauf qu'au lieu de mettre que $\{f_k\}$ converge dans $L^p(K)$ vers f (où $K \subset \Omega$ est le compact où on approxime), on dit seulement que c'est vrai das $L^1(K)$ (puisque $f \in L^1_{loc}(\Omega)$). Ensuite, l'inégalité de Poincaré marche encore dans l'espace homogène (on n'utilise pas la convergence dans L^p de la suite approximante, juste la convergence presque partout, et Fatou). Du coup, les fonctions de l'espace $W^{1,p}(\Omega)$ homogène sont aussi dans $L^p_{loc}(\Omega)$, et même un espace L^r_{loc} meilleur, mais pas dans $L^p(\Omega$ si Ω a une forme bizarre. Et aussi, d'ailleurs, mon hypothèse que $f \in L^1_{loc}$ est juste plus pratique pour définir la distribution associée et les dérivées, mais pas utile; le fait que $f \in L^1_{loc}$ serait une conclusion. Enfin, à cause de ce qui vient d'être dit, les fonctions de l'espace homogène $W^{1,n}(\mathbb{R}^n)$ sont dans $BMO(\mathbb{R}^n)$, avec la même démonstration que ci dessus. C'est un énoncé plus satisfaisant, puisque ce sont deux espaces homogènes et de fonctions définies à une constante additive près.

• J'ai pris la définition logique avec des boules, à cause de son invariance. Mais on aurait pu remplacer les boules par des cubes à côtés parallèles aux axes dans la définition, et

$$||f||_* = \sup_{Q cube} \frac{1}{|Q|}\int_Q |f(x) - m_B f| dx,$$

où implicitement "cube" signifie "cube à côtés parallèles aux axes," est une norme équivalente à $||f||_{BMO}$. On a aussi la proposition suivante, parfois bien utile, et qui permettra de faire le lien en se fatigant encore moins.

Proposition. *Soit $f \in L^1_{loc}(\mathbb{R}^n)$. Alors $f \in BMO$ si et seulement s'il existe une constante $C(f)$ telle que, pour toute boule B, il existe C_B tel que $\frac{1}{|B|}\int_B |f(x) - C_B| dx \leq C(f)$. Et $||f||_{BMO}$ est équivalent à l'inf des $C(f)$ qui marchent.*

Dans un sens c'est facile: on prend $C_B = m_B f$, qui marche. Dans l'autre sens, si B et C_B sont donnés, on trouve que $|m_B f - C_B| \leq \frac{1}{|B|}\int_B |f - C_B| \leq$

$C(f)$, et ensuite $\frac{1}{|B|}\int_B |f(x)-m_Bf|dx \le |m_Bf - C_B| + \int_B |f(x)-C_B|dx \le 2C(f)$. $\square$

On a un énoncé analogue en remplaçant les boules B par des cubes Q et $||f||_{BMO}$ par $||f||_*$. On passe de l'énoncé avec C_B avec un énoncé semblable avec des cubes en notant que si Q est un cube et B la plus petite boule qui le contient, $\frac{1}{|Q|}\int_Q |f-C_B| \le \frac{1}{|Q|}\int_B |f-C_B| \le 2^n \frac{1}{|B|}\int_B |f-C_B| \le 2^n C(f)$. On fait pareil dans l'autre sens. Finalement, toutes les normes mentionnées ci-dessus sont équivalentes. Il y en a bien d'autres, comme le dira le théorème ci-dessous.

Lemme. *Si $f \in BMO$ et si I, J sont deux cubes tels que $I \subset J$ et $|J| \le 2|I|$, alors $|m_If - m_Jf| \le 2||f||_*$. Un peu plus généralement, si $f \in BMO$ et si I, J sont deux cubes tels que $I \subset J$ et $|J| \le 2^m|I|$, alors $|m_If - m_Jf| \le 2m\,||f||_*$.*

En effet $\int_I |f-m_Jf| \le \int_J |f-m_Jf| \le |J|||f||_*$ et $|m_If - m_Jf| = \frac{1}{|I|}\left|\int_I (f-m_Jf)\right| \le \frac{1}{|I|}\int_I |f-m_Jf| \le \frac{1}{|I|}\int_J |f-m_Jf| \le \frac{|J|}{|I|}||f||_* \le 2||f||_*$, d'où la première partie. Pour la seconde, on passe par un cube intermédiaire K tel que $I \subset K \subset J$, et $|K| = 2^k|I|$, avec k maximal. Alors $|m_If - m_Kf| \le 2k||f||_*$ en itérant la première partie, et $|m_Kf - m_Jf| \le 2||f||_*$ par la première partie, ce qui donne le résultat en additionnant. Noter que $k \le \log_2(|J|/|I|)$. $\square$

Théorème (John-Nirenberg). *Il existe une constante $C_n > 0$ telle que, si $f \in BMO(\mathbb{R}^n)$ (possiblement à valeurs complexes), et si Q est un cube, alors*

$$\frac{1}{|Q|}\int_Q \exp\left(\frac{1}{10n||f||_*}\,|f(x)-m_Qf|\right) dx \le C_n. \tag{1}$$

Surprise, on croyait seulement f localement intégrable, et voilà qu'elle est exponentiellement intégrable sur Q! C'est parce que l'on demande quelque-chose pour chaque cube, et que ces diverses informations se recoupent puisque la moyenne de f sur Q ne dépend pas n'importe comment de Q (voir le lemme).

Evidemment la constante $10n$ n'est pas optimale, et la norme $||f||_{BMO}$ donnerait une autre constante. La constante pour f à valeurs réelles est peut-être un peu plus petite.

Noter que $\ln(|x|)$ correspond au type de comportement donné par le lemme et le théorème. En particulier, on ne peut pas remplacer $10n$ par une constante trop petite. Noter qu'on a un énoncé semblable en remplaçant les cubes Q par des boules B, et en augmentant un peu la constante $10n$. Cet enoncé se déduit du précédent en appliquant le théorème au plus petit cube qui contient B.

Démonstration. Quelques réductions pour commencer. Notre énoncé est invariant par multiplication de f par une constante. Donc on peut supposer que $||f||_* \leq 1$ et prouver l'énoncé avec $\frac{1}{10n}$. Et même, en distinguant parties réelle et imaginaire, on peut supposer f réelle et démontrer (1) avec $\frac{1}{5n}$.

Par invariance par translations et dilatations, on peut supposer que Q est le cube unité Q_0. Comme le résultat ne change pas en ajoutant une constante, on peut supposer que $m_Q f = 0$. Enfin, puisque $e^{|x|} \leq e^x + e^{-x}$, et quitte à multiplier C_n par 2, on va pouvoir se contenter de démontrer que (après toutes ces réductions)

$$\int_{Q_0} \exp\left(\frac{f(x)}{5n}\right) dx \leq C_n. \tag{2}$$

On aura besoin de cubes dyadiques. Un cube dyadique de côté 2^{-k}, où $k \in \mathbb{Z}$ est donné, est un cube de la forme $Q = [0, 2^{-k}]^n + \ell 2^{-k}$, avec $\ell \in \mathbb{Z}^n$. On essaiera de noter $\mathcal{D}_k$ la classe des cubes dyadiques de côté 2^{-k}, et $\mathcal{D}$ l'union des $\mathcal{D}_k$

Noter que les cubes dyadiques de côté 2^{-k} recouvrent $\mathbb{R}^n$ et sont d'intérieurs disjoints, et aussi que si $j \geq k$, $Q \in \mathcal{D}_k$, et $R \in \mathcal{D}_j$, alors ou bien $R \subset Q$ ou alors R ne rencontre pas l'intérieur de Q. C'est très pratique, en particulier parce que si $\mathcal{P}$ est une propriété concernant des cubes, l'ensemble des cubes de $\mathcal{D}$ ayant la propriété $\mathcal{P}$ est automatiquement composé de cubes presque disjoints (j'entends par là d'intérieurs disjoints).

Ici on se contente de la classe $\mathcal{D}(Q_0)$ de tous les cubes dyadiques (de côtés divers) contenus dans Q_0. Pour chacun on note m_Q au lieu de $m_Q f$.

Pour $Q \in \mathcal{D}(Q_0)$, on note Q^* le parent de Q, c'est-à-dire l'unique cube de côté double qui le contient. On dira aussi le père de Q, c'est sexite mais rend plus clair le fait que chaque cube n'a qu'un parent. Noter que

$$|m_Q - m_{Q^*}| \leq 2n, \tag{3}$$

à cause du lemme et puisque $|Q^*| = 2^n |Q|$.

On veut contrôler les cubes où la moyenne est grosse, et montrer qu'ils sont en nombre exponentiellement décroissant. Et on va baser l'estimation sur un argument de temps d'arrêt (d'où l'intérêt d'utiliser des cubes dyadiques).

Posons $T = 5n$, et notons $\mathcal{F}_k$, pour $k \geq 1$, l'ensemble des cubes $Q \in \mathcal{D}$ tels que $m_Q > kT$, et qui sont maximaux avec cette propriété. Posons aussi $\mathcal{F}_0 = \{Q_0\}$. Notons $\Omega_k = \cup_{Q \in \mathcal{F}_k} Q$ pour $k \geq 1$. A cause de la maximalité, l'union est presque disjointe (les intérieurs des cubes sont disjoints). Noter aussi que les ensembles Ω_k sont emboîtés: $\Omega_{k+1} \subset \Omega_k$. On va montrer que pour $k \geq 1$,

$$|\Omega_k| \leq \frac{1}{3} |\Omega_{k-1}| \tag{4}$$

Vérifions, avant de commencer, que

$$kT < m_Q \leq kT + 2n \quad \text{pour } k \geq 1 \text{ et } Q \in \mathcal{F}_k. \tag{5}$$

La première inégalité vient de la définition de $\mathcal{F}_k$. Pour la seconde, notons que $m_{Q_0} = 0$, donc le père Q^* de Q est bien dans $\mathcal{D}(Q_0)$, et $Q^* \notin \mathcal{F}_k$ par maximalité. Donc $m_{Q^*} \leq kT$, et (5) se déduit de (3).

Maintenant fixons $k \geq 1$. Vérifions que pour $Q \in \mathcal{F}_k$, il existe un unique $R \in \mathcal{F}_{k-1}$ qui contient Q. Soit $Q \in \mathcal{F}_k$; bien sûr $m_Q > kT > (k-1)T$, et en essayant successivement tous les ancêtres de Q, on finit par trouver un cube R tel que $m_R > (k-1)T$, et qui soit maximal avec cette propriété. Ainsi $R \in \mathcal{F}_{k-1}$ et R contient Q. C'est le seul cube de $\mathcal{F}_{k-1}$ qui contient Q, puisque les cubes de $\mathcal{F}_k$ sont essentiellement disjoints.

Soit $R \in \mathcal{F}_{k-1}$. A cause de ce qu'on vient de dire, et parce que les autres cubes de $\mathcal{F}_{k-1}$ sont essentiellement disjoints de R, les cubes $Q \in \mathcal{F}_k$ sont soit contenus dans R, soit essentiellement disjoints de R. Donc $R \cap \Omega_k$ est, à un ensemble de mesure nulle près, l'union essentiellement disjointe des cubes $Q \in \mathcal{F}_k$ qui sont contenus dans R. Notons $\mathcal{F}_k(R)$ cet ensemble de cubes. Alors

$$\int_{R \cap \Omega_k} (f - m_R) = \sum_{Q \in \mathcal{F}_k(R)} \int_Q (f - m_R) = \sum_{Q \in \mathcal{F}_k(R)} (m_Q - m_R)\, |Q|. \tag{6}$$

Or $m_Q \geq kT$ puisque $Q \in \mathcal{F}_k$, et $m_R \leq (k-1)T + 2n$ par (5) (ou trivialement si $k = 1$). Donc

$$\int_{R \cap \Omega_k} (f - m_R) \geq \sum_{Q \in \mathcal{F}_k(R)} (T - 2n)\, |Q| = (T - 2n)\, |R \cap \Omega_k| = 3n\, |R \cap \Omega_k|. \tag{7}$$

Pour le reste de R,

$$(8) \qquad \Big|\int_{R\setminus\Omega_k}(f-m_R)\Big| \leq \int_{R\setminus\Omega_k}|f-m_R| \leq |R|||f||_* \leq |R|$$

par définition de $||f||_*$ et parceque $||f||_* \leq 1$. Mais $\int_R(f-m_R)=0$ par définition de m_R, donc (7) et (8) donnent

$$(9) \qquad 3n\,|R\cap\Omega_k| \leq \int_{R\cap\Omega_k}(f-m_R) \leq \Big|\int_{R\setminus\Omega_k}(f-m_R)\Big| \leq |R|$$

donc $|R\cap\Omega_k| \leq |R|/3$.

Maintenant Ω_k est l'union essentiellement disjointe des $\Omega_k\cap R$, $R\in\mathcal{F}_{k-1}$, parce que tout $Q\in\mathcal{D}_k$ est contenu dans un $R\in\mathcal{F}_{k-1}$, donc on peut sommer sur R et on trouve (4). En conclusion, $|\Omega_k|\leq 3^{-k}$ pour $k\geq 1$.

Pour presque-tout $x\in Q_0$, $0=\lim\limits_{r\to 0}\dfrac{1}{|B(x,r)|}\displaystyle\int_{B(x,r)}|f(y)-f(x)|dy$. Alors, si Q est un cube dyadique de côté 2^{-j} qui contient x, et si l'on utilise $r=n2^{-j+2}$ (pour être sûr que $Q\subset B(x,r)$), alors $\dfrac{1}{|Q|}\displaystyle\int_Q|f(y)-f(x)|dy \leq \dfrac{C}{|B(x,r)|}\displaystyle\int_{B(x,r)}|f(y)-f(x)|dy$, qui tend vers 0 quand j tend vers $+\infty$. Et alors m_Q tend vers $f(x)$.

Ainsi, si l'on pose $E_k=\{x\in Q_0\,;\,f(x)>kT\}$, alors pour presque-tout $x\in E_k$, tous les petits cubes dyadiques qui contiennent x sont tels que $m_Q>kT$, donc $x\in\Omega_k$. Autrement dit, on a montré que

$$(10) \qquad |\{x\in Q_0\,;\,f(x)>kT\}| = |E_k| \leq |\Omega_k| \leq 3^{-k}.$$

Une estimation de ce type est en fait équivalente à l'exponentielle intégrabilité qu'on souhaite. Bref, on peut maintenant conclure:

$$(11) \qquad \begin{aligned}\int_{Q_0}\exp\left(\frac{f(x)}{5n}\right)dx &\leq \sum_{k\geq 1}\int_{\Omega_{k-1}\setminus\Omega_k}\exp\left(\frac{f(x)}{5n}\right)dx\\ &\leq \sum_{k\geq 1}|\Omega_{k-1}\setminus\Omega_k|\,\exp(kT/5n) \leq \sum_{k\geq 1}3^{-k}e^k \leq C\end{aligned}$$

parce que la mesure de l'intersection des Ω_k est nulle, puis que $f(x)\leq kT$ presque-partout sur $Q_0\setminus\Omega_k$. C'est ce qu'on voulait, et ceci donne le théorème de John et Nirenberg. $\square$

Deux commentaires. D'abord, bien que cela ne se voie qu'à peine, on a ici un exemple particulièrement simple d'argument de temps d'arrêt répété: on utilise les premiers cubes qui vérifient une propriété X, on s'en sert pour estimer quelque chose (typiquement, le nombre de cubes), et on recommence.

Le théorème s'applique aux fonctions de $W^{1,n}(\mathbb{R}^n)$, puisque $W^{1,n}(\mathbb{R}^n) \subset BMO$, et donne l'intégrabilité de leur exponentielle si leur norme est assez petite.

Corollaire. *On pose, pour* $1 \leq p < +\infty$ *et* $f \in L^1_{loc}(\mathbb{R}^n)$,

$$||f||_{BMO,p} = \sup_{(x,r)\in\mathbb{R}^n\times]0,+\infty[} \left\{ \frac{1}{|B(x,r)|} \int_{B(x,r} |f(x) - m_{B(x,r)}f|^p \right\}^{1/p}.$$

Pour chaque p, $||f||_{BMO,p}$ *est une norme sur* BMO, *équivalente aux normes* $||f||_{BMO}$ *et* $||f||_*$. *Même résultat en remplaçant les boules par des cubes.*

Démonstration laissée en exercice. □

8. MESURES DE CARLESON

a. Définition

La notion est bien utile pour les opérateurs de Calderón-Zygmund, mais pas uniquement. On va juste voir ici le théorème simple mais fondamental de Carleson (sur la fonction maximale d'accès non tangentiel), et l'exemple des fonctions de BMO. Rien de très difficile, mais j'aime bien dire que dans certain cas, prononcer les mots "mesure de Carleson" c'est déjà presque résoudre le problème.

On va se placer sur $\mathbb{R}^{n+1}_+ = \big\{(x,t)\,;\, x \in \mathbb{R}^n \text{ et } t > 0\big\}$. Il est souvent agréable de penser à $\mathbb{R}^{n+1}_+$ comme étant l'espace des boules $B(x,t)$, au sens où un point de $\mathbb{R}^{n+1}_+$ donne un lieu (x) et une échelle (t). L'idée est par exemple que certaines propriétés des fonctions (par exemple, de régularité) peuvent être mieux comprises en étudiant des fonctions associées sur $\mathbb{R}^{n+1}_+$. Je crois que le premier exemple de ce type est l'extension harmonique de f au demi-espace, mais on en verra d'autres.

La mesure naturelle sur cet espace est plutôt la mesure invariante par translations et dilatations $\frac{dxdt}{t}$.

Définition 1. *Une mesure de Carleson sur* $\mathbb{R}^{n+1}_+$ *est une mesure positive* μ *sur* $\mathbb{R}^{n+1}_+$ *telle qu'il existe* $C \geq 0$ *telle que*

(cm) $$\mu(B(x,r)\times]0,r]) \leq Cr^n \quad \textit{pour } x \in \mathbb{R}^n \textit{ et } r > 0.$$

La norme de Carleson de μ, qu'on notera volontiers $||\mu||_{CM}$, est alors la plus petite constante C telle que (cm) ait lieu.

Noter que l'invariance est n-dimensionelle. C'est fait exprès.

On aurait pu remplacer (cm) par

$$(cm') \quad \mu(Q\times]0,r]) \leq Cr^n \text{ pour tout cube } Q \subset \mathbb{R}^n, \text{ où } r \text{ est le côté de } Q$$

et où "cube" signifie en fait "cube à faces parallèles aux axes". Et on aurait aussi pu remplacer r par $2r$, ou par $r/2$, sans changer la notion. [Faire un peu de géométrie.]

Contrexemples. D'abord, $\frac{dxdt}{t}$ a la bonne homogénéité, mais ne marche pas: le membre de gauche de (1) est toujours $+\infty$.

La mesure $dxdt$ ne marche pas non plus, à cause des gros cubes. L'homogénéité n'est pas la bonne.

Exemples.

Pour $\lambda > 0$ et $A \geq 1$, $d\mu = \mathbf{1}_{\lambda \leq t \leq A\lambda} \frac{dxdt}{t}$ donne une mesure de Carleson, avec une norme $\leq C \log(A)$.

La mesure de Lebesgue (de dimension n) sur un hyperplan (parallèle à $\mathbb{R}^n$, ou perpendiculaire, ou même transverse.

On fixe $t > 0$, et on prend la somme des mesures de Dirac sur les couples (k, t), où $k \in t\mathbb{Z}^n$.

On en verra de plus jolies plus tard, sous la forme $F(x,t) \frac{dxdt}{t}$.

b. Le théorème de Carleson

En fait, il en a beaucoup. Il s'agit ici du petit théorème simple mais utile sur les mesures de Carleson et la fonction maximale non-tangentielle.

Passons à la définition des cônes d'accès non tangentiel. On se donne $A > 0$ (probablement, $A = 1$ suffirait) et, pour $x \in \mathbb{R}^n$, on note

$$(1) \qquad \Gamma_A(x) = \big\{(y,t) \in \mathbb{R}^{n+1}_+ \,;\, |y-x| \leq At\big\}$$

le cône d'approche au-dessus de x. Ensuite, pour f définie sur $\mathbb{R}^{n+1}_+$, on lui associe une fonction maximale définie sur $\mathbb{R}^n$, par

$$(2) \qquad \mathcal{M}_A f(x) = \sup\big\{|f(y,t)|\,;\, (y,t) \in \Gamma_A(x)\big\}.$$

Et voici le théorème qui va avec.

Théorème 1 (Carleson). *Soit μ une mesure de Carleson sur $\mathbb{R}^{n+1}_+$. Alors, pour tout f borélienne définie sur $\mathbb{R}^{n+1}_+$ et tout $\lambda > 0$,*

(3)
$$\mu\big(\big\{(y,t) \in \mathbb{R}^{n+1}_+ \,;\, |f(y,t)| > \lambda\big\}\big) \leq C||\mu||_{CM} \left|\big\{x \in \mathbb{R}^n \,;\, |\mathcal{M}_A f(x)| > \lambda\big\}\right|.$$

La constante C ne dépend que de n et A.

On verra bientôt à quoi ça sert, mais le plus simple est de le démontrer d'abord. Pour $(x,r) \in \mathbb{R}^{n+1}_+$, on définit la tente sous (x,r) par

$$T(x,r) = \big\{(y,t) \in \mathbb{R}^{n+1}_+ \,;\, 0 < t < r \text{ et } |y-x| \leq A(r-t)\big\} \tag{4}$$

et sa base $b(x,r) = \big\{y \in \mathbb{R}^n \,;\, |y-x| \leq Ar\big\}$.

Ensuite, pour $O \subset \mathbb{R}^{n+1}_+$, on note $T(O) = \bigcup_{(x,r)\in O} T(x,r)$ et $b(O) = \bigcup_{(x,r)\in O} b(x,r)$. On s'intéresse à $O = \big\{(x,r) \in \mathbb{R}^{n+1}_+ \,;\, |f(x,r)| > \lambda\big\}$. Vérifions déjà que

$$b(O) = \big\{y \in \mathbb{R}^n \,;\, |\mathcal{M}_A f(y)| > \lambda\big\}. \tag{5}$$

En effet, pour $y \in \mathbb{R}^n$, $|\mathcal{M}_A f(y)| > \lambda$ si et seulement si il existe $(x,r) \in \Gamma_A(y)$ tel que $|f(x,r)| > \lambda$. La condition sur (x,r) est que $|y-x| \leq Ar$, qui s'écrit aussi $y \in b(x,r)$. Donc $|\mathcal{M}_A f(y)| > \lambda$ si et seulement si $y \in b(x,r)$ pour un $(x,r) \in O$. Donc on a (5).

Ainsi, (3) s'écrit plus simplement $\mu(O) \leq C|b(O)|$, et c'est ceci qu'on va vérifier. On peut supposer que $|b(O)| < +\infty$, sinon il n'y a rien à démontrer. Notons qu'alors l'altitude de O est bornée: si $(x,r) \in O$, $b(x,r) \subset b(0)$, donc $cr^n = |B(x,Ar)| = |b(x,r)| \leq |b(O)|$.

Pour $z = (x,r) \in O$, on pose $B(z) = B(z,2r)$. Ceci donne un recouvrement de O par des boules, on vient de voir que les rayons sont bornés, donc le lemme de type Vitali s'applique et donne une famille $I \subset O$ telle que les boules $B(z)$, $z \in I$, sont disjointes, et les $5B(z)$ recouvrent O. Notons $z = (x_z, r_z)$ et $D(z) = B(z) \cap b(z) = B(x_z, \rho_z) \cap \mathbb{R}^n$, où $\rho_z = r_z \mathrm{Min}(\sqrt{3}, A)$ et avec un léger abus de langage concernant l'inclusion de R^n dans $\mathbb{R}^{n+1}$. Alors

(6)
$$\mu(O) \leq \sum_{z\in I} \mu(5B(z)) \leq C \sum_{z\in I} r_z^n \leq C \sum_{z\in I} \rho_z^n \leq C \sum_{z\in I} |D(z)| \leq C|\cup_{z\in I} b(z)| \subset |b(O)|$$

par définition d'une mesure de Carleson, puis parce que les $D(z) \subset B(z)$ sont disjoints, et $D(z) \subset b(z) \subset b(O)$. D'où (3) et le théorème. $\square$

Corollaire 2. *Sous les hypothèses du théorème, et pour tout* $p > 0$,

$$\int_{\mathbb{R}^{n+1}_+} |f(x,r)|^p d\mu(x,r) \leq C||\mu||_{CM} \int_{R^n} |\mathcal{M}_A f(y)|^p dy, \tag{7}$$

avec $C = C(n, A)$.

C'est facile, puisque $\int g^p d\mu = \int_0^{+\infty} p\lambda^{p-1}\mu(\{g > \lambda\} d\lambda$ pour g mesurable positive, et pareil pour le second membre. □

Corollaire 3. *Soient* ψ *une fonction radiale bornée décroissante positive intégrable sur* $\mathbb{R}^n$, *et* φ *une fonction mesurable telle que* $|\varphi(x)| \leq \psi(x)$ *pour* $x \in \mathbb{R}^n$. *Posons* $\varphi_t(x) = t^{-n}\varphi(x/t)$ *pour* $x \in \mathbb{R}^n$ *et* $t > 0$. *Soient encore* μ *une mesure de Carleson sur* $\mathbb{R}^{n+1}_+$ *et* $p \in [1, +\infty[$. *Pour* $f \in L^p(\mathbb{R}^n)$, *posons* $F(x,t) = \varphi_t * f(x)$ *pour* $(x,t) \in \mathbb{R}^{n+1}_+$. *Alors*

$$\int_{\mathbb{R}^{n+1}_+} |F(x,t)|^p d\mu(x,t) \leq C||\mu||_{CM} \int_{R^n} f^*(y)^p dy \tag{8}$$

pour une constante $C = C(n, \psi)$, *et où* f^* *est la fonction maximale de Hardy-Littlewood de* f.

Donc quand $p > 1$, le membre de droite est $\leq C||f||_p^p$. Ceci sera l'application typique: on part de f sur $\mathbb{R}^n$; on l'étend en une fonction F définie sur le demi-espace (le plus souvent, φ est d'intégrale 1, donc on peut s'attendre à ce que $F(x,r)$ tende vers $f(x)$ quand r tend vers 0, par exemple), et le corollaire donne un contrôle sur F.

En fait la même démonstration marche encore si F est donnée par une formule plus compliquée $F(x,t) = \int \varphi_t(x,y)f(y)dy$, où l'on suppose juste que $|\varphi_t(x,y)| \leq \psi_t(x-y) = t^{-n}\psi((x-y)/t)$.

Avec nos hypothèses, φ est dans l'espace L^q pour tout $q \geq 1$, donc il n'y a pas de problème pour définir $F(x,t)$ partout.

A cause du corollaire 2, il suffit de montrer que

$$\int_{R^n} |\mathcal{M}_A F(y)|^p dy \leq C \int_{R^n} f^*(y)^p dy, \tag{9}$$

donc on veut montrer que $\mathcal{M}_A F(y)$ est dominée par f^*. La chose serait plus simple si φ était bornée à support compact (pas besoin de sommes comme ci-dessous), mais faisons quand même le cas général tout de suite, en décomposant φ en morceaux.

Notons $A_0 = B(0,1)$ et, pour $k \geq 1$, $A_k = B(0,2^k) \setminus B(0,2^{k-1})$. Donc $\mathbb{R}^n$ est l'union disjointe des A_k, et $\varphi = \sum_k \varphi_k$, où $\varphi_k = \varphi \mathbf{1}_{A_k}$.

Notons $a_0 = \psi(0)$, et pour $k \geq 1$, prenons pour a_k la valeur constante de ψ sur $\partial B(0,2^{k-1})$. Donc $||\varphi_k|||_\infty \leq a_k$, et on sait par hypothèse que $a_0 < +\infty$ et que

$$(10) \qquad \sum_{k\geq 1} a_k 2^{nk} \leq C \sum_{k\geq 1} \int_{A_{k-1}} \psi(x)dx = \int_{\mathbb{R}^n} \psi(x)dx < +\infty$$

puisque $\psi(x) \geq a_k$ sur A_{k-1}.

Revenons à F et sa fonction maximale. Donnons-nous donc $y \in \mathbb{R}^n$, et soit ensuite $(x,t) \in \Gamma_A(y)$. Ceci veut dire que $|x-y| \leq At$. Ensuite on calcule $F(x,t) = \varphi_t * f(x) = \sum_k \varphi_{k,t} * f(x)$, où bien sûr $\varphi_{k,t}$ est la dilatée de φ_k avec normalisation L^1, comme plus haut. On estime chaque morceau et on sommera après:

(11)

$$\begin{aligned} |\varphi_{k,t} * f(x)| &\leq \int |\varphi_{k,t}(x-z)|\,|f(z)|dz = t^{-n} \int |\varphi_k((x-z)/t)|\,|f(z)|dz \\ &\leq t^{-n} a_k \int_{|z-x|\leq 2^k t} |f(z)|dz \leq t^{-n} a_k \int_{|z-y|\leq 2^k t + At} |f(z)|dz \\ &\leq C t^{-n} a_k [2^k t + At]^n f^*(y) \leq C a_k (2^k + A)^n f^*(y) \end{aligned}$$

puisque $|x-y| \leq At$ et par définition de $f^*(y)$.

On somme tout ceci sur k, on utilise (10), et on trouve que $|\varphi_t * f(x)| \leq C f^*(y)$ quand $(x,t) \in \Gamma_A(y)$ puis, en prenant bêtement le sup, que $\mathcal{M}_A F(y) \leq C f^*(y)$, et on en déduit aussitôt (9) et (8). $\square$

On appliquera ceci plus tard, quand on parlera d'opérateurs de Calderón-Zygmund, mais pour l'instant on s'arrête là et on donne un exemple standard de mesure de Carleson.

c. BMO et les mesures de Carleson

Théorème 4. *Soit $\psi \in C_c^\infty(\mathbb{R}^n)$ telle que $\int_{\mathbb{R}^n} \psi(x)dx = 0$. Posons $\psi_t(x) = t^{-n}\psi(x/t)$ pour $x \in \mathbb{R}^n$ et $t > 0$. Finalement soit $f \in BMO(\mathbb{R}^n)$. Alors*

$$(12) \qquad d\mu(x,t) = |\psi_t * f(x)|^2 \frac{dxdt}{t} \quad \textit{est une mesure de Carleson,}$$

avec une constante $\leq C(n,\psi)\,||f||^2_{BMO}$.

Noter avant de commencer que $\psi_t * f(x)$ ne change pas quand on ajoute une constante à f. C'est plutôt rassurant.

On n'a pas besoin de toutes ces dérivées de ψ. Pour la première partie de la démonstration, (à savoir (14) pour les fonctions de L^2), on a juste besoin de savoir, par exemple, que $\psi \in L^1$ (c'est plus pratique pour définir la convolution, et en plus ça nous dit que $\widehat{\psi}$ est continue bornée), et qu'il existe une constante $C_\psi \geq 0$ telle que pour tout ξ dans la sphère unité de $\mathbb{R}^n$,

$$\int_{t>0} |\widehat{\psi}(t\xi)|^2 \,\frac{dt}{t} \leq C_\psi. \tag{13}$$

Ces conditions sont assez faciles à déduire de nos hypothèses, puisqu'alors $\widehat{\psi}$ est régulière près de 0, donc $|\widehat{\psi}(u)| \leq C|u|$ puisque $\widehat{\psi}(0) = \int \psi = 0$, et décroît à l'infini, donc $|\widehat{\psi}(u)| \leq C|u|^{-1}$. Alors $\displaystyle\int_{t>0} |\widehat{\psi}(t\xi)|^2 \,\frac{dt}{t} \leq C\int_0^1 t^2\,\frac{dt}{t} + C\int_1^{+\infty} t^{-2}\,\frac{dt}{t} \leq C$ et (13) s'ensuit.

Avant de démontrer le théorème, montrons sa version plus simple sur L^2: il existe $C \geq 0$ telle que

$$\int_{\mathbb{R}^{n+1}_+} |\psi_t * g(x)|^2 \,\frac{dxdt}{t} \leq C||g||_2^2 \tag{14}$$

pour tout $g \in L^2(\mathbb{R}^n)$.

Ceci se démontre en appliquant Plancherel à t fixé: on dit que

$$\int_{R^n} |\psi_t * g(x)|^2\, dx = \int_{R^n} |\widehat{\psi_t * g}(\xi)|^2 d\xi = \int_{R^n} |\widehat{\psi}(t\xi)\widehat{g}(\xi)|^2 d\xi$$

puis on intègre le résultat contre dt/t et on trouve, après application de Fubini, que

$$\int_{\mathbb{R}^{n+1}_+} |\psi_t * g(x)|^2 \,\frac{dxdt}{t} = \int_{R^n} |\widehat{g}(\xi)|^2 \Big\{ \int_{t>0} |\widehat{\psi}(t\xi)^2 \frac{dt}{t} \Big\}\, d\xi.$$

L'intégrale intérieure ne change pas quand on remplace ξ par $\lambda\xi$, $\lambda > 0$ (poser $u = \lambda t$), et ensuite (13) dit qu'elle est inférieure à C. On trouve donc bien (14) en ré-appliquant Plancherel.

Revenons à $f \in BMO$ et $d\mu(x,t) = |\psi_t * f(x)|^2 \frac{dxdt}{t}$. On va utiliser le fait que ψ est à support compact (et on garde les hypothèses utiles pour (14)). On pourrait facilement trouver des hypothèses plus faibles.

Soit donc A tel que $\psi(x) = 0$ hors de $B(0,A)$. On doit prouver qu'il existe une constante C telle que

$$\int_{x\in B(y,r)} \int_{0<t<r} |\psi_t * f(x)|^2 \frac{dxdt}{t} \leq Cr^n |f||^2_{BMO} \tag{15}$$

pour tout $y \in \mathbb{R}^n$ et tout $r > 0$. Notons $B = B(y,(A+1)r)$. Comme on peut ajouter une constante à f sans rien changer, on peut supposer que $m_B f = 0$. Posons encoe $f_1 = f\mathbf{1}_B$ (qui est donc d'intégrale nulle).

Noter que pour tout $t < r$ et tout $x \in B(y,r)$, l'intégrale qui définit $\psi_t * f(x) = \int \psi_t(x-z)f(z)dz$ ne porte que sur des valeurs de $f(z)$ où $|z-x| \leq At$, donc $|z-y| \leq (A+1)r$. De sorte que $\psi_t * f(x) = \psi_t * f_1(x)$ et que

(16)
$$\begin{aligned}\int_{x\in B(y,r)} \int_{0<t<r} |\psi_t * f(x)|^2 \frac{dxdt}{t} &= \int_{x\in B(y,r)} \int_{0<t<r} |\psi_t * f_1(x)|^2 \frac{dxdt}{t}\\ &\leq \int_{x\in\mathbb{R}^n} \int_{t>0} |\psi_t * f_1(x)|^2 \frac{dxdt}{t}\\ &\leq C||f_1||_2^2 = \int_B |f(z)|^2 dz = \int_B |f(z) - m_B f|^2 dz\\ &\leq C||f||^2_{BMO}\, |B| = Cr^n ||f||^2_{BMO}\end{aligned}$$

par (14), puis la définition de f_1, puis le fait que $m_B f = 0$, et enfin par John et Nirenberg. On a prouvé (15) et le théorème suit. $\square$

A nouveau, ceci risque de servir pour les opérateurs d'intégrale singulière.

9. BASE DE HAAR

Dans ce qui suit je vais essayer de décrire une base de $L^2(\mathbb{R}^n)$, appelée base de Haar. L'idée est de procéder de manière telle que vous puissiez généraliser facilement pour construire d'autres bases d'ondelettes (je vais appliquer le même formalisme d'analyse multirésolution), et aussi des bases de Haar adaptées à d'autres situations (espaces de type homogène, théorème $T(b)$).

On va commencer par des notations. On munit $\mathbb{R}^n$ de ses décompositions en cubes dyadique. On note D_k l'ensemble des cubes dyadiques de taille 2^{-k}, et $D = \cup_{k\in\mathbb{Z}}\mathcal{D}_k$ l'ensemble de tous les cubes dyadiques.

Pour $k \in \mathbb{Z}$, on note V_k l'ensemble des fonctions de $L^2 = L^2(\mathbb{R}^n, dx)$ qui sont constantes sur chaque cube de $\mathcal{D}_k$. Plus précisément, pour chaque $Q \in \mathcal{D}_k$, f est égale presque-partout à une constante. En particulier, on ne cherche même pas à définir f sur les frontières de cubes, et on fera tout comme si ces cubes étaient disjoints. Donc $f \in V_k$ ssi

$$f = \sum_{Q\in\mathcal{D}_k} f_Q \mathbf{1}_Q, \tag{1}$$

avec $\sum |f_Q|^2 < +\infty$.

Les espaces V_k forment une suite croissante de sous-espaces vectoriels fermés de L^2 (la base de notre analyse multirésolution). C'est pour obtenir la croissance qu'on a pris les tailles 2^{-k} plus haut; la croissance vient de ce que chaque cube de $\mathcal{D}_k$ a une décomposition presque disjointe en cubes de D_{k+1}.

On peut voir V_k comme la classe des fonctions de L^2 qui sont mesurables pour la tribu engendrée par les cubes de $\mathcal{D}_k$; les projections E_k ci-dessous correspondent alors à l'espérance conditionnelle.

On note E_k la projection orthogonale sur V_k (on la note comme une espérance). Il est facile de voir que

$$E_k f(x) = \sum_{Q\in\mathcal{D}_k} \Big\{ \frac{1}{|Q|} \int_Q f(y)dy \Big\} \mathbf{1}_Q \tag{2}$$

par exemple parce que les fonctions $|Q|^{-1/2}\,\mathbf{1}_Q$, $Q \in \mathcal{D}_k$, forment une base orthonormée de V_k.

Notre suite a la propriété agréable supplémentaire que $f(x) \in V_k$ ssi $f(2^k x) \in V_0$. Ca ne sera pas si important dans la suite, mais le fait qu'on n'a une certaine invariance des estimations aidera.

Pour compléter les propriétés d'une analyse multirésolution, on doit encore signaler que pour $f \in L^2$,

$$\lim_{k\to-\infty} E_k f = 0 \quad \text{et} \quad \lim_{k\to+\infty} E_k f = f \tag{3}$$

où la limite a lieu dans L^2. Autrement dit, E_k tend fortement vers 0 quand $k \to -\infty$ et vers Id quand $k \to +\infty$. Comme les opérateurs E_k sont uniformément bornés (ils sont même de norme 1), il suffit par un argument de densité de vérifier (3) pour f dans une classe dense, par exemple les fonctions continues à support compact, pour lesquelles (3) est facile à vérifier. Pour $k \to -\infty$, par exemple, on observe que la somme porte sur au plus 2^n termes, que les coefficients $\left\{\frac{1}{|Q|}\int_Q f(y)dy\right\}$ sont en $O(1/|Q|) = O(2^k)$, et donc que $||E_k f||_2^2$ est en $O(1/|Q|)$.

Maintenant on note W_k le complémentaire orthogonal de V_k dans V_{k+1}. La projection orthogonale Δ_k sur W_k est donc donnée par

$$\Delta_k f = E_{k+1} f - E_k f. \tag{4}$$

[Penser qu'on peut compléter une base orthonormée de V_k en une BO de V_{k+1}; les vecteurs ajoutés forment une BO de W_k, et on peut se contenter de vérifier (4) sur la base.]

Noter que les W_k sont deux à deux orthogonaux car si $k < l$, $W_k \subset V_k$, qui est orthogonal à W_l. Donc pour $k < l$,

$$E_l f - E_k f = \sum_{k\le j\le l-1} D_j f \quad \text{et} \quad ||E_l f - E_k f||^2 = \sum_{k\le j\le l-1} ||D_j f||^2. \tag{5}$$

On fait tendre k vers $-\infty$ et l vers $+\infty$, et on trouve que

$$f = \sum_{j\in\mathbb{Z}} \Delta_j f \tag{6}$$

(avec convergence forte des sommes partielles de la série), et

$$||f||^2 = \sum_{j\in\mathbb{Z}} ||D_j f||^2. \tag{7}$$

On peut aller facilement un cran plus loin, et découper chaque $\Delta_k f$ en morceaux portés par les cubes de $\mathcal{D}_k$, et écrivant

$$\Delta_k f = \sum_{Q\in\mathcal{D}_k} \mathbf{1}_Q \, [\Delta_k f] \tag{8}$$

où les morceaux sont orthogonaux parce qu'ils sont portés par des ensembles disjoints.

On peut d'ailleurs calculer chaque morceau. Notons $\mathcal{F}(Q)$ la collections des enfants de Q (les cubes de $\mathcal{D}_{k+1}$ qui sont contenus dans Q); alors (4), puis (2) pour $k+1$ et k donnent

$$\Delta_k f = \sum_{Q \in \mathcal{D}_k} \sum_{R \in \mathcal{F}(Q)} a_R \mathbf{1}_R \tag{9}$$

où la valeur constante de $\Delta_k f$ sur R est donnée par

$$a_R = \frac{1}{|R|} \int_R f - \frac{1}{|Q|} \int_Q f = m_R f - m_Q f. \tag{10}$$

Mais venons-en à la base de Haar. Notons, pour $Q \in \mathcal{D}_k$, $W_k(Q)$ le sous-espace de W_k des fonctions qui sont à support dans Q. Ce sont donc les fonctions qui sont de la forme $g = \sum_{R \in \mathcal{F}(Q)} a_R \mathbf{1}_R$ (parce que $g \in V_{k+1}$), mais avec la contrainte supplémentaire que $\int_Q g = 0$ (ou de manière équivalente $\sum_R a_R = 0$), dûe au fait que $g \in V_k^{\perp}$.

Et pour tout k et tout $Q \in \mathcal{D}_k$, on choisit une base raisonnable de $W_k(Q)$.

Quand $n = 1$, c'est le plus facile, $W_k(Q)$ est seulement de dimension 1, donc on a assez peu le choix. On nomme R_- et R_+ les deux enfants de Q (des intervalles de longueur moitié), disons avec R_- à gauche de R_+, et on pose

$$h_Q = |Q|^{-1/2} [\mathbf{1}_{R_+} - \mathbf{1}_{R_-}] \tag{11}$$

Quand $n > 1$, il y a 2^n cubes dans $\mathcal{F}(Q)$, donc $W_k(Q)$ est de dimension 2^{n-1}. On choisit des fonctions $h_{Q,\varepsilon}$, où $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n) \in \{0,1\}^n \backslash \{(0, \dots, 0)\}$, que l'on définit de la manière suivante.

D'abord, on écrit $Q = I_1 \times I_2 \ldots \times I_n$. Pour chaque intervalle I, on découpe I en deux intervalles I_- et I_+ de même longueur $|I|/2$, et disons avec I_- à gauche de I_+. Ensuite on pose $\varphi_I^0 = |I|^{-1/2} \mathbf{1}_I$ et $\varphi_I^1 = |I|^{-1/2} [\mathbf{1}_{I_+} - \mathbf{1}_{I_-}]$, et avec ces notations bizarres, on pose

$$h_{Q,\varepsilon}(x_1, \dots, x_n) = \varphi_{I_1}^{\varepsilon_1}(x_1) \ldots \varphi_{I_n}^{\varepsilon_n}(x_n). \tag{12}$$

Il est facile de voir que toutes ces fonctions sont d'intégrale nulle (appliquer Fubini et noter que puisque $\varepsilon \neq (0, \dots, 0)$, l'une au moins des $\varphi_{I_j}^{\varepsilon_j}(x_j)$ est d'intégrale nulle. De plus, ces fonctions sont deux à deux orthogonales: si $\varepsilon \neq \varepsilon'$, on choisit j tel que $\varepsilon_j \neq \varepsilon'_j$, et on constate que $h_{Q,\varepsilon} h_{Q,\varepsilon'}$ est le produit de $\varphi^1_{I_j}(x_j)$ par une fonction des autres variables dont on se moque, puisque $\int \varphi^1_{I_j}(x_j) dx_j = 0$.

Comme on a le bon nombre de fonctions, on a obtenu que les $h_{Q,\varepsilon}$, $\varepsilon \in \{0,1\}^n \setminus \{(0,\dots,0)\}$, forment bien une base orthonormée de $W_k(Q)$. Et ainsi, on a aussi la décomposition orthogonale

$$\mathbf{1}_Q\,[\Delta_k f] = \sum_{\varepsilon} \Big\{ \int_Q f(y) h_{Q,\varepsilon}(y) dy \Big\} h_{Q,\varepsilon}, \tag{13}$$

où la dernière expression vient de ce que $\langle \mathbf{1}_Q\,[\Delta_k f], h_{Q,\varepsilon} \rangle = \int_Q f(y)[\Delta_k f(y)] h_{Q,\varepsilon}(y) dy \Big\}$ $\int_Q f(y) h_{Q,\varepsilon}(y) dy$ parce que $h_{Q,\varepsilon}$ est à support dans Q, puis est orthogonal à $f - \Delta_k f$ puisque $h_{Q,\varepsilon} \in W_k$.

En regroupant tout,

$$f = \sum_{k \in \mathbb{Z}} \sum_{Q \in \mathcal{D}_k} \sum_{\varepsilon} \Big\{ \int_Q f(y) h_{Q,\varepsilon}(y) dy \Big\} h_{Q,\varepsilon}, \tag{14}$$

et la somme est orthogonale. Autrement dit, la collection des $h_{Q,\varepsilon}$, $Q \in \mathcal{D}$ et $\varepsilon \in \{0,1\}^n \setminus \{(0,\dots,0)\}$, est une base orthonormée de $L^2(\mathbb{R}^n)$.

Et aussi, pour $f \in L^2$,

$$||f||_2^2 = \sum_{k \in \mathbb{Z}} \sum_{Q \in \mathcal{D}_k} \sum_{\varepsilon} \Big\{ \int_Q f(y) h_{Q,\varepsilon}(y) dy \Big\}^2. \tag{15}$$

Quelques commentaires. D'abord, le système de Haar est assez ancien; la formalisation en termes d'espaces V_k (analyse multirésolutions) est plus récente. Par contre, j'imagine que Haar a peut-être déja envisagé le coté "martingale".

Ici on a tout pris à valeurs réelles pour simplifier, mais bien sûr les $h_{Q,\varepsilon}$ forment aussi une base de $L^2_{\mathbb{C}}$.

On a une bonne invariance par dilatations dyadiques et translations entières dans V_0. Tous les $h_{Q,\varepsilon}$ sont obtenus à partir d'un seul (celui du cube unité) par translations entières puis dilatations dyadiques.

Si au lieu de prendre les espaces V_k ci-dessus, on avait pris d'autres espaces (par exemple, en dimension 1, les fonctions continues qui sont affine sur chaque intervalle dyadique, ou d'autres exemples plus subtils, on aurait encore pu construire W_k, puis (dans les bons cas) une base invariante par translation de chaque W_k, puis finalement une base d'ondelettes.

Nous serons éventuellement intéressé par deux autres généralisations de la base de Haar. Dans la première, on remplace $\mathbb{R}^n$ par un espace métrique mesuré doublant (espace de type homogène), muni d'un analogue idoine des cubes dyadiques, et on trouve une base de Haar quand même. Le fait que les cubes $Q \in \mathcal{D}$ ont un nombre variable d'enfants n'est pas gênant, on prend une base de chaque $W_k(Q)$ de manière indépendante.

L'autre généralisation est de fournir une base adaptée à une fonction bornée b à valeurs complexe, telle que $\left|\frac{1}{|Q|}\int_Q b(y)dy\right| \geq \delta > 0$ pour tout Q. on en parlera peut-être au moment de $T(b)$.

Pour finir ce court chapitre, donnons une caractérisation de "$f \in BMO$" en termes de condition de Carleson pour les coefficients $f_{Q,\varepsilon} = \int_Q f(y)h_{Q,\varepsilon}(y)dy$ de f dans la base de Haar. C'est évidemment à rapprocher du théorème du chapitre précédent sur les mesures de Carleson.

Proposition. *Pour $f \in BMO(\mathbb{R}^n)$, les coefficients $f_{Q,\varepsilon} = \int_Q f(y)h_{Q,\varepsilon}(y)dy$ vérifient la condition suivante ("Carleson packing condition"): il existe $C \geq 0$ ne dépendant que de n, telle que*

$$\sum_{Q\in\mathcal{D}(R)} \sum_{\varepsilon} |f_{Q,\varepsilon}|^2 \leq C|R|\,||f||_{BMO}^2 \tag{16}$$

pour tout cube $R \in \mathcal{D}$, où l'on note $\mathcal{D}(R)$ l'ensemble des cubes dyadiques contenus dans R.

Noter que l'on aurait aussi pu sommer sur les cubes contenus dans $B(x,r)$, et obtenir $\leq Cr^n\,||f||_{BMO}$.

La démonstration simple vaudrait pour toute base d'ondelettes. On pose $f_1 = \mathbf{1}_R[f - m_R f]$, et on note que pour $Q \subset R$, les $h_{Q,\varepsilon}$ sont à support dans $Q \subset R$ et d'intégrale nulle, donc $f_{Q,\varepsilon}$ est aussi le coefficient correspondant pour f_1. [Avec une base d'ondelettes à support compact, on aurait pris $f_1 = \mathbf{1}_B[f - m_B f]$, où B est une boule centrée sur R et de rayon $C\,\mathrm{diam}(R)$, et on aurait poursuivi pareil.] Donc

$$\sum_{Q\in\mathcal{D}(R)} \sum_{\varepsilon} |f_{Q,\varepsilon}|^2 = ||f_1||_2^2 = \int_R |f - m_R f|^2 \leq C|R|\,||f||_{BMO}^2$$

par (15) appliqué à f_1, et par John et Nirenberg. $\square$

Exercice. Vérifier que réciproquement, si $f \in L^1_{loc}$ et le membre de gauche de (16) est $\leq C|R|$, alors $f \in BMO$.

Beaucoup d'autres espaces fonctionnels ont des caractérisations en termes de la taille des $f_{Q,\varepsilon}$ (ou de coefficients en ondelettes).

10. OPÉRATEURS D'INTÉGRALE SINGULIÈRE

Je me lance un peu; j'espère ne pas avoir à revenir trop en arrière. Je vais aussi essayer d'aller un peu vite à l'essentiel, en oubliant des tas de résultats utiles et remarquables.

10.a. Définition des opérateurs d'intégrale singulière

On va donner une définition assez générale (c'est plus pratique pour les manipulations), mais pas trop. On se placera sur $\mathbb{R}^n$, même s'il y aurait moyen de se placer sur un espace de type homogène (espace métrique avec une mesure doublante). La terminologie qui suit est essentiellement celle de Coifman et Meyer (il y en a eu d'autres, mais en gros celle-ci a gagné).

Definition 1. *Un opérateur d'intégrale singulière (SIO) sur $\mathbb{R}^n$ est un opérateur borné $T : C_c^\infty(\mathbb{R}^n) \to C_c^\infty(\mathbb{R}^n)'$ pour lequel il existe un noyau standard K (voir la définition plus bas) tel que*

$$\langle Tf, g \rangle = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} K(x,y) f(y) g(x) dx dy \tag{1}$$

pour tout choix de $f \in C_c^\infty(\mathbb{R}^n)$ et $g \in C_c^\infty(\mathbb{R}^n)$ tels que les supports $\mathrm{supp}(f)$ *et* $\mathrm{supp}(g)$ *soient disjoints.*

Quelques remarques avant de définir les noyaux standards.

Au lieu de dire que T envoie $C_c^\infty(\mathbb{R}^n)$ dans son dual, on aurait pu définir directement T comme une forme bilinéaire sur $C_c^\infty(\mathbb{R}^n) \times C_c^\infty(\mathbb{R}^n)$. En tout cas, ce qu'on connaît de T, c'est bien les nombres $\langle Tf, g \rangle$ (l'effet de Tf sur g), où $f, g \in C_c^\infty(\mathbb{R}^n)$.

On n'a pas besoin de C_c^∞ en général. Le plus souvent on peut se contenter de C_c^1, ou au pire de C_c^N. Du coup la définition de la continuité est plus facile à écrire. Ce qu'on demande, c'est que pour tout choix de boule B (destinée à contenir des supports), l'application $(f, g) \to \langle Tf, g \rangle$, définie de $C_c^N(B) \times C_c^N(B)$ dans $\mathbb{R}$ (ou $\mathbb{C}$, ou parfois même dans un espace de Banach), est continue. Et ici $C_c^N(B)$ est l'ensemble des fonctions de classe C^N à support compact contenu dans B, muni par exemple de la norme $\sum_D ||Df||_\infty$, où l'on somme sur toutes les dérivées partielles D d'ordre $\leq N$. Noter que plus N est grand, plus T est facile à définir.

Dans un espace de type homogène, on serait obligé de se contenter des fonctions-test qui sont simplement à support compact et Höldériennes d'exposant α pour un certain $\alpha > 0$.

Comme on le verra dès la définition des noyaux standards, l'intégrale double dans (1) converge dès que f et g sont à supports disjoints. Par contre, l'intégrale double risquerait fort de diverger sans cette contrainte de supports.

Definition 2. *Un noyau standard sur $\mathbb{R}^n$ est une application. continue K : $\mathbb{R}^n \times \mathbb{R}^n \setminus \Delta$, où $\Delta = \big\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^n \,;\, x = y\big\}$ est la diagonale, telle qu'il existe des constantes $C \geq 0$ et $\delta > 0$ telles que*

$$|K(x,y)| \leq C|x-y|^{-n} \text{ pour } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \setminus \Delta, \tag{2}$$

$$|K(x,y) - K(x,y')| \leq C\frac{|y-y'|^\delta}{|x-y|^{n+\delta}} \text{ pour } x, y, y' \in \mathbb{R}^n \text{ tels que } |y'-y| \leq |x-y|/2, \tag{3}$$

et

$$|K(x,y) - K(x',y)| \leq C\frac{|x-x'|^\delta}{|x-y|^{n+\delta}} \text{ pour } x, x', y \in \mathbb{R}^n \text{ tels que } |x'-x| \leq |x-y|/2. \tag{4}$$

Noter le comportement comme si K était homogène de degré $-n$; (3) et (4) donne un peu de régularité supplémentaire (en gardant la même homogénéité), et sont impliquées par la condition plus brutale

$$|\nabla_x K(x,y| + |\nabla_y K(x,y| \leq C|x-y|^{-n-1} \text{ pour } (x,y) \in \mathbb{R}^n \times \mathbb{R}^n \setminus \Delta, \tag{5}$$

avec n'importe quelle valeur de $\delta \leq 1$.

Exemple typique: $K(x,y) = \frac{1}{x-y}$ quand $n = 1$, qui donne un multiple de la transformée de Hilbert, ou $K(x,y) = \frac{x_j - y_j}{|x-n|^{n+1}}$, qui correspond à une constante près à la j-ième transformée de Riesz.

Noter aussi que, même avec ces exemples simples, l'intégrale double dans (1) ne converge que parce qu'on a pris f et g à supports disjoints.

Un exemple d'opérateur d'intégrale singulière T est l'opérateur de multiplication par une fonction $m \in L^1_{loc}$, et alors on peut prendre $K = 0$. Autrement dit $\langle Tf, g\rangle = \int mfg$ définit bien un opérateur borné $T : C_c^\infty(\mathbb{R}^n) \to C_c^\infty(\mathbb{R}^n)'$, et $\langle Tf, g\rangle = 0$ quand f et g ont des supports disjoints. Donc K ne détermine pas T.

L'exemple ci-dessus nous plaira encore, surtout si m est bornée pour que T soit borné sur L^2, mais on n'aimera pas du tout l'exemple suivant: prendre

$n = 1$ et $Tf = f'$ (ou une dérivée d'ordre supérieur). Ici aussi, T est in SIO associé à $K = 0$, mais en plus l'homogénéité de T (dans sa partie non donnée par le noyau, et par rapport aux dilatations (comme ci-dessous)) n'est pas celle suggérée par (2)-(4).

Par contre T détermine K. Prendre des fonctions bosses f et g qui convergent vers des masses de dirac en y et $x \neq y$, et noter que $\langle Tf, g\rangle$ converge alors vers $K(x, y)$.

En fait, si $f \in C_c(\mathbb{R}^n)$, et si x n'est pas dans le support de f, on peut poser

$$Tf(x) = \int_{\mathbb{R}^n} K(x, y) f(y) dy. \tag{6}$$

L'intégrale converge parce que $\mathrm{dist}(x, \mathrm{supp}(f)) > 0$; ensuite (1) dit que Tf coincide hors de $\mathrm{supp}(f)$ avec la fonction $Tf(x)$ qu'on vient de définir, au sens où $\langle Tf, g\rangle = \int Tf(x) g(x) dx$ pour g à support hors de $\mathrm{supp}(f)$.

Notons encore que la classe des SIO est stable par conjugaison par translations, rotations, et dilatations, au sens où, par exemple, si l'on associe à l'application $x \to ax$ (avec $a > 0$) l'oprérateur A défini par $Af(x) = f(x/a)$, alors $A^{-1}TA$ est aussi un opérateur d'intégrale singulière, associé au noyau standard $\widetilde{K}$, où $\widetilde{K} = a^n K(ax, ay)$ qui vérifie (2)-(4) avec les mêmes constantes que K. [Les autres invariances sont plus faciles.]

Parlons un peu des <u>opérateurs tronqués</u> T_ε. L'avantage est que T_ε ne dépend que du noyau K. On définit simplement T_ε par

$$T_\varepsilon f(x) = \int_{|y-x|>\varepsilon} K(x, y) f(y), \tag{7}$$

disons pour f bornée à support compact, ou même dans L^q pour un $q < +\infty$ (pour que l'intégrale converge grâce à (2)).

On ne dit pas que les T_ε convergent vers T (penser à l'opérateur de multiplication par m), mais ils donnent quand même certaines informations sur T.

Par souci de contrôle, on définit aussi <u>l'opérateur maximal</u> T_* par

$$T_* f(x) = \sup_{\varepsilon>0} |T_\varepsilon f(x)| \tag{8}$$

pour les mêmes f.

Quand K est antisymétrique, c.-à.-d. quand $K(y,x) = -K(x,y)$ pour $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n \setminus \Delta$, il y a un opérateur T naturellement associé à K, que l'on définit par

$$\langle Tf, g\rangle = \lim_{\varepsilon \to 0} \langle T_\varepsilon f, g\rangle \tag{9}$$

pour $f, g \in C^1_c(\mathbb{R}^n)$ (on n'aura pas besoin de plus de dérivées). On appellera cet opérateur l'opérateur de valeur principale associé à K. Vérifions juste que la limite dans (9) existe. On utilise l'antisymétrie de K pour écrire que

(10)

$$\begin{aligned}\langle T_\varepsilon f, g\rangle &= \int\int_{|y-x|>\varepsilon} K(x,y) f(y) g(x)\, dydx \\ &= \frac{1}{2}\int\int_{|y-x|>\varepsilon} K(x,y) f(y) g(x) dydx - \frac{1}{2}\int\int_{|y-x|>\varepsilon} K(y,x) f(y) g(x) dydx \\ &= \frac{1}{2}\int\int_{|y-x|>\varepsilon} K(x,y)\big[f(y)g(x) - f(x)g(y)\big]\, dydx\end{aligned}$$

(parce que le domaine d'intégration est symétrique). Maintenant, il ne reste plus qu'à noter que pour f et g de classe C^1 à supports compacts, $|f(y)g(x) - f(x)g(y)| \leq |f(y)g(x) - f(x)g(x)| + |f(x)g(x) - f(x)g(y)| C|x-y|$, de sorte que l'intégrale double $\int\int K(x,y)[f(y)g(x) - f(x)g(y)]\, dydx$ converge, et que la limite dans (9) existe. On peut même définir T directement par

$$\langle Tf, g\rangle = \frac{1}{2}\int\int K(x,y)\big[f(y)g(x) - f(x)g(y)\big]\, dydx \tag{11}$$

pour $f, g \in C^1_c$. On laisse en exercice le fait que ceci définit bien un opérateur borné, disons de $C^1_c(\mathbb{R}^n)$ dans son dual et que son noyau est K. Et aussi, pour ceux qui connaîtraient la distribution $vp(\frac{1}{x})$ (valeur principale de $1/x$) sur $\mathbb{R}$, que $K(x,y) = 1/(x-y)$ donne bien l'opérateur T de convolution avec la distribution $vp(\frac{1}{x})$. Pareil pour les noyaux $K_j(x,y) = \frac{x_j - y_j}{|x-y|^{n+1}}$ qui donnent (à une constante multiplicative près) des transformées de Riesz.

Un exemple typique (mais sur lequel nous ne nous étendrons pas dans ce cours): les opérateurs pseudodifférentiels classiques d'ordre 0, qui sont donnés par une formule du type

$$Tf(x) = \int_{\mathbb{R}^n} \sigma(x,\xi) \widehat{f}(\xi) e^{ix\cdot\xi} d\xi \tag{12}$$

[on a multiplié par le symbole avant d'appliquer la transformée de Fourier inverse; si σ était un polynôme en ξ, on obtiendrait un opérateur différentiel] avec un symbole σ tel que $|\partial_\xi^\alpha \partial_x^\beta \sigma(x,\xi)| \leq C_{\alpha,\beta}(1+|\xi|)^{-\alpha}$. Les opérateurs en question sont également bornés sur $L^2(\mathbb{R}^n)$, et on pourrait également prendre des conditions plus faibles sur les symboles. Voir le petit Lecture Notes de Journé, ou sans doute des tas de références plus récentes en EDP.

On parlera sans doute plus tard d'un exemple célèbre, le noyau de Cauchy associé à un graphe Lipschitzien.

10.b. Opérateurs de Calderón-Zygmund; continuité L^p.

Définition 3. *Un opérateur de Calderón-Zygmund (CZO) est un SIO qui a une extension bornée sur $L^2(\mathbb{R}^n)$.*

Autrement dit, c'est un SIO tel qu'il existe une constante $C \geq 0$ telle que

$$|\langle Tf, g\rangle| \leq C||f||_2||g||_2 \quad \text{pour } f, g \in C_c^\infty(\mathbb{R}^n). \tag{13}$$

On va voir qu'automatiquement T est aussi borné sur L^p pour $1 < p < +\infty$. C'est le même intervalle que pour la fonction maximale et ça n'est pas étonnant. Pardon, j'ai un peu oublié les attributions des théorèmes célèbres ci-dessous.

Théorème 4. *Si T est un CZO, T a une extension continue $T : L^\infty(\mathbb{R}^n) \to BMO(\mathbb{R}^n)$.*

On se donne $f \in L^\infty$. On va en fait devoir définir Tf, à partir de l'extension déjà définie sur L^2 (et unique par densité des fonctions test), et Tf sera défini à une constante additive près (une vraie fonction de BMO). On fait tout ça parce que l'intégrale $\int K(x,y)f(y)dy$ est en général divergente.

Voyons comment faire. Soit $B = B(x,r)$ une boule. On découpe f en $f_1 + f_2$, où $f_1 = f\mathbf{1}_B$, et on va juste regarder ce qui se passe sur la boule moitié $B(x, r/2)$.

On a déjà une définition de Tf_1, qui est une fonction de L^2, et on va maintenant définir ce qui doit être $Tf_2(z) - Tf_2(x)$ pour $z \in B(x, r/2)$. On pose simplement

$$a_{x,r}(z) = \int_{\mathbb{R}^n} [K(z,y) - K(x,y)]\, f_2(y)dy. \tag{14}$$

L'intégrale porte sur $y \notin B$, donc $|z-y| \geq r/2$ (et bien sûr $|x-y| \geq r/2$), donc pas de problème de divergence locale, et comme à l'infini on note que

$$|K(z,y) - K(x,y)| \leq C|z-x|^{\delta}|x-y|^{-n-\delta} \leq Cr^{\delta}|x-y|^{-n-\delta}, \tag{15}$$

on trouve en intégrant hors de B que l'intégrale converge, et que

$$|a_{x,r}(z)| \leq C||f||_{\infty} \int_{|x-y|>r} r^{\delta}|x-y|^{-n-\delta}dy \leq C||f||_{\infty}. \tag{16}$$

Rassurons-nous maintenant, et donnons une définition de Tf. On va exiger que $\int_{B(0,1)} Tf = 0$. On choisit $B = B(0,R)$ (avec R grand), ce qui donne f_1, f_2, et aussi $Tf_1 \in L^2$. Ensuite, on pose

$$Tf(z) = Tf_1(z) + a_{0,R}(z) + \lambda \text{ pour } z \in B(0, R/2) \tag{17}$$

où la constante λ est juste choisie pour que $\int_{B(0,1)} Tf(z) = 0$. On vérifie sans difficulté que notre définition de $Tf(z)$ ne dépend pas du choix de R grand, parce que quand passe de $B(0,R)$ à $B(0,S)$, disons avec $S > R$, le terme $Tf_1(z)$ augmente de $\int_{B(0,S)\setminus B(0,R)} K(z,y)f(y)dy$, et $a_{0,R}(z)$ diminue de $\int_{B(0,S)\setminus B(0,R)} [K(z,y) - K(0,y)]f(y)dy = \int_{B(0,S)\setminus B(0,R)} K(z,y)f(y)dy + C$, et C disparait dans la normalisation (puisqu'on ne change pas la moyenne).

Pour même genre de raisons, pour tout choix de boule $B(x,r)$ comme ci-dessus, on a encore la formule $Tf_2(z) - Tf_2(x) = a_{x,r}(z)$ quand $z \in B(x, r/2)$ (et avec la décomposition plus haut).

Notons encore que la fonction Tf que l'on vient de définir dépend linéairement de f, et coïncide avec la définition précédente (modulo une constante additive) quand f est une fonction test. Bref, on a défini notre extension de T.

On a fait le plus dur. Vérifions encore que $Tf \in BMO$. On se donne une boule $B = B(x, r/2)$, et on utilise la décomposition ci-dessus. On pose $c_B = Tf_2(x)$, et alors

$$\begin{aligned} \frac{1}{|B|}\int_B |Tf - c_B| &\leq \frac{1}{|B|}\int_B |Tf_1| + \frac{1}{|B|}\int_B |Tf_2(z) - Tf_2(x)| \\ &\leq \left\{\frac{1}{|B|}\int_B |Tf_1|^2\right\}^{1/2} + \frac{1}{|B|}\int_B |a_{x,r}(z)| \\ &\leq |B|^{-1/2}||Tf_1||_2 + C||f||_{\infty} \\ &\leq C|B|^{-1/2}||f_1||_2 + C||f||_{\infty} \leq C||f||_{\infty} \end{aligned} \tag{18}$$

par (16), puis en utilisant la continuité de T sur L^2 et le fait que $||f_1||_\infty \leq ||f||_\infty$. On en déduit bien le résultat, en utilisant la caractérisation élémentaire de BMO avec les constantes c_B ou c_Q. $\square$

Le théorème 4 a une version duale, qui dit que si T est un CZO, alors il envoie l'espace de Hardy atomique H^1 land L^1. Je ne crois pas qu'on le fera. A la place, voyons un autre résultat extrême, le fait que T envoie L^1 dans L^1_{faible}, qui va aussi nous donner l'occasion de voir un autre outil simple, mais remarquable, la décomposition de Calderón-Zygmund d'une fonction.

Théorème 5. *Si T est un CZO, T a une extension continue $T : L^1(\mathbb{R}^n) \to L^1_{faible}(\mathbb{R}^n)$.*

Il existe de ce théorème une version à poids. Voir un vague énoncé plus bas.

Donc on se donne $f \in L^1$, et on veut à la fois définir Tf et surtout vérifier qu'il est dans L^1_{faible}. On va utiliser la très utile décomposition de Calderón-Zygmund de f au niveau $\lambda > 0$, dont voici la définition.

On note toujours $\mathcal{D}$ l'ensemble des cubes dyadiques, puis $\mathcal{D}(\lambda)$ l'ensemble des cubes $Q \in \mathcal{D}$ tels que

$$m_Q|f| > \lambda, \tag{19}$$

et qui sont maximaux parmi les cubes de $\mathcal{D}$ ayant cette propriété. Posons

$$\Omega = \Omega(\lambda) = \bigcup_{Q \in \mathcal{D}(\lambda)} Q\,; \tag{20}$$

c'est une union presque disjointe, comme d'habitude en tant qu'union de cubes dyadiques maximaux ayant une propriété donnée. Vérifions que

$$\big\{x \in \mathbb{R}^n\,;\, |f(x)| > \lambda\big\} \text{ est presque contenu dans } \Omega(\lambda). \tag{21}$$

A cause du théorème de densité de Lebesgue (version forte), on sait que pour presque tout x, $|f(x)|$ est la limite des $m_Q|f|$, où D est un cube dyadique contenant x et dont le diamètre tend vers 0. Si $|f(x)| > \lambda$, certains de ces cubes vérifient (19). Noter que pour un tel cube Q, $\lambda|Q| \leq \int_Q |f| \leq |Q|\,||f||_1$, de sorte que Q ne peut pas être trop grand. Et donc ces cubes sont donc contenus dans un cube maximal vérifiant (19). On en déduit (21).

Maintenant on découpe f en une bonne fonction g, qui est aussi bornée, et une mauvaise fonction b, qui en fait n'est pas si mal non plus parce qu'elle est composée d'atomes. On pose

$$g(x) = f(x) \ \text{ pour } x \in \mathbb{R}^n \setminus \Omega(\lambda) \tag{22}$$

et

$$g(x) = m_Q f \ \text{ pour } x \in Q \text{ quand } Q \in \Omega(\lambda), \tag{23}$$

ce qui définit donc bien g presque-partout, puis on pose $b = f - g$. Noter que b est nulle hors de $\Omega(\lambda)$, et

$$b(x) = f(x) - m_Q f \ \text{ pour } x \in Q \text{ quand } Q \in \Omega(\lambda), \tag{24}$$

ce qui implique en particulier que

$$m_Q b = 0 \ \text{ pour } Q \in \Omega(\lambda). \tag{25}$$

On a annoncé que g est bornée. En fait, (21) et (22) disent que $|g(x)| = |f(x)| \leq \lambda$ presque-partout hors de $\Omega(\lambda)$, et sur $\Omega(\lambda)$ on a $|g(x)| \leq 2^n\lambda$ parce que, comme pour $Q \in \mathcal{D}(\lambda)$, le parent $\widehat{Q}$ de Q ne vérifie pas (19), on a

$$|m_Q f| \leq m_Q|f| = \frac{1}{|Q|}\int_Q |f| \leq \frac{1}{|Q|}\int_{\widetilde{Q}} |f| = 2^n m_{\widetilde{Q}}|f| \leq 2^n\lambda. \tag{26}$$

On aura aussi besoin de la taille de $\Omega(\lambda)$:

$$|\Omega(\lambda)| = \sum_{Q\in\mathcal{D}(\lambda)} |Q| \leq \lambda^{-1} \sum_{Q\in\mathcal{D}(\lambda)} \int_Q |f| = \lambda^{-1}\int_{\Omega(\lambda)} |f| \leq \lambda^{-1}||f||_1 \tag{27}$$

puisque les cubes sont presque disjoints et par (19).

Voici tout ce qu'il est utile de savoir sur la décomposition $f = g + b$. Voyons maintenant comment s'en servir pour définir et estimer Tf. On va d'abord supposer que $f \in L^2 \cap L^1$, ce qui permet de ne pas avoir de problème de définition.

On commence par noter que $g \in L^2$, avec

(28)
$$||g||_2^2 = \int_{\mathbb{R}^n\setminus\Omega(\lambda)} |f|^2 + \int_{\Omega(\lambda)} |g|^2 \leq \lambda \int_{\mathbb{R}^n\setminus\Omega(\lambda)} |f| + 2^{2n}\lambda^2|\Omega(\lambda)| \leq C\lambda||f||_1$$

à cause de nos bornes sur g et par (27). Comme T est maintenant défini et borné sur L^2, on trouve que $Tg \in L^2$, avec $||Tg||_2^2 \leq C\lambda ||f||_1$, donc

$$\big|\big\{x \in \mathbb{R}^n \, ; \, |Tg(x)| > \lambda\big\}\big| \leq C\lambda^{-1}||f||_1 \, . \tag{29}$$

Notons $\widehat{\Omega}$ l'union des doubles des cubes $Q \in \mathcal{D}(\lambda)$ (ce ne sont pas des cubes dyadiques, mais tant pis). Noter que

$$|\widehat{\Omega}| \leq C \sum_{Q \in \mathcal{D}(\lambda)} |Q| \leq C|\Omega(\lambda)| \leq C\lambda^{-1}||f||_1 \, , \tag{30}$$

et contentons-nous d'estimer Tb hors de $\widehat{\Omega}$. Pour chaque cube Q, notons $b_Q = \mathbf{1}_Q b = \mathbf{1}_Q[f - m_Q f]$. Pour $x \notin 2Q$, on peut utiliser le noyau K pour calculer $Tb_Q(x)$, et on trouve

$$Tb_Q(x) = \int_Q K(x,y)[f(y) - m_Q f] dy. \tag{31}$$

En fait, ceci serait vrai directement si l'on savait que b_Q est une fonction test, mais on en déduit le cas général facilement par un passage à la limite.

Notons y_Q le centre de Q, et observons que

$$\int_Q K(x,y_Q)[f(y) - m_Q f] dy = K(x,y_Q) \int_Q [f(y) - m_Q f] dy = 0$$

puisque la moyenne est nulle, de sorte que l'on peut soustraire dans (31) pour obtenir une meilleure décroissance:

$$\begin{aligned} |Tb_Q(x)| &= \Big| \int_Q [K(x,y) - K(x,y_Q)][f(y) - m_Q f] dy \Big| \\ &\leq C \int_Q \frac{|y-y_Q|^\delta}{|x-x_Q|^{n+\delta}} |f(y) - m_Q f| dy \leq C d(Q)^\delta |x-y_Q|^{-n-\delta} \int_Q |f|. \end{aligned} \tag{32}$$

On intègre en x et on trouve

$$\int_{\mathbb{R}^n \setminus 2Q} |Tb_Q(x)| dx \leq C \int_Q |f| \leq C\lambda |Q| \tag{33}$$

(puisque $\int_Q |f| \leq 2^n|Q|$). Et ensuite on somme en Q et il vient
(34)
$$\int_{\mathbb{R}^n\setminus\widehat{\Omega}} |Tb(x)|dx \leq \sum_Q \int_{\mathbb{R}^n\setminus\widehat{\Omega}} |Tb_Q(x)|dx \leq C\sum_Q \lambda|Q| \leq C\lambda|\Omega(\lambda)| \leq C||f||_1.$$

Bref, b n'est pas si mauvaise parce qu'elle est composée d'"atomes" b_Q dont l'intégrale est nulle, ce qui donne assez de décroissance aux $T(b_Q)$.

Finalement, notons $Z = \{x\,;\,|Tf(x)| > 2\lambda\}$. Cet ensemble est contenu dans l'union de $\widehat{\Omega}$, de l'ensemble estimé en (29) où $|Tg| > \lambda$, et de l'ensemble des $x \in \mathbb{R}^n \setminus \widehat{\Omega}$ où $|Tb(x)| > \lambda$, dont la mesure est également $\leq C\lambda^{-1}||f||_1$.

Bref, $|Z| \leq C\lambda^{-1}||f||_1$, et on a montré que $Tf \in L^1_{faible}$, avec une norme $\leq C||f||_1$.

Il reste pour conclure à dire que l'on peut étendre T en un opérateur qui va de L^1 dans L^1_{faible}, par l'argument standard de densité qui marche encore même si L^1_{faible} n'est pas exactement normé.

On pourrait aussi noter que l'on peut définir assez correctement $Tf = Tg+Tb$ hors de $\widehat{\Omega}$, en regardant l'argument ci-dessus, et comme en prenant λ aussi grand qu'on veut, $\mathbb{R}^n \setminus \widehat{\Omega}$ est aussi petit qu'on veut, ceci permettrait de définit $Tf(x)$ presque partout, après une ennuyeuse vérification de cohérence.
□

Corollaire 6. *Si T est un CZO, T a une extension continue $T : L^p(\mathbb{R}^n) \to L^p(\mathbb{R}^n)$ pour $1 < p < +\infty$.*

D'abord, par interpolation réelle à partir du théorème 5, le résultat vaut pour $1 < p \leq 2$. Pour $p > 2$, on va procéder par dualité. Soit q l'esposant conjugué. On cherche à montrer que

$$|\langle Tf, g\rangle| \leq C||f||_p||g||_q \tag{35}$$

pour $f, g \in C^\infty_c$, donc que l'opérateur adjoint T^t défini par

$$\langle T^t g, f\rangle = \langle Tf, g\rangle \tag{36}$$

est borné (a une extension bornée) sur L^q. Or il est borné sur L^2 (puisque T l'est), et c'est un SIO associé au noyau $(x, y) \to K(y, x)$ (vérification facile directe). Donc il est bien continu sur L^q, comme souhaité. □

Il y a un certain nombre de résultats supplémentaires dont je ne parlerai pas. J'imagine qu'ils sont tous dans l'union des deux livres de Stein, et sans doute aussi dans le Meyer en trois volumes et dans le Lecture notes de Journé.

Soit donc T un CZO. D'abord, T a pour $1 < p < +\infty$ une extension continue de $L^p(\omega dx)$ dans lui-même, à condition que ω soit un poids de la classe A_p de Muckenhoupt, qui a une dédinition précise, mais est aussi la classe des poids ω tels que la fonction maximale de Hardy-Littlewood envoie continument $L^p(\omega dx)$ dans lui-même. Je crois que c'est un théorème de Fefferman-Stein, basé sur l'inégalité de Fefferman-Stein qui domine la fonction dièse de Tf par une fonction maximale.

On pourrait aussi interpoler directement entre des continuités $H^1 \to L^1$ et $L^\infty \to BMO$, grâce à un argument simple mais puissant de contrôle des fonctions de répartition, les inégalités aux bons λ, dont on parlera peut-être.

Le cas extrême, qui dit que pour $\omega \in A^1$, T envoie $L^1(\omega dx)$ dans $L^1_{faible}(\omega dx)$ marche encore, et avec presque la même démonstration que le théorème ci-dessus (une fois qu'on sait que T envoie $L^2(\omega dx)$ dans lui-même).

On a aussi le résultat dual du théorème 4, qui dit que T evoie l'espace atomique $H^1 \subset L^1$ dans L^1.

L'opérateur maximal T_* aussi a une extension continue de $L^1(\omega dx)$ dans $L^1_{faible}(\omega dx)$ pour $\omega \in A_1$ et de $L^p(\omega dx)$ dans lui-même pour $\omega \in A_p$, résultat qui est basé sur un contrôle ponctuel de $T_*(x)$ par des fonctions maximales de Tf et de f, qu'on appelle inégalité de Cotlar).

Bref, la grande morale de tous ces résultats est qu'une fois qu'on sait que le SIO T est continu sur L^2, beaucoup d'autres inégalités s'ensuivent. Il reste à savoir quand c'est le cas.

11. CRITÈRES DE CONTINUITÉ SUR L^2 : $T(1)$ ET $T(b)$

11.a. Le noyau de Cauchy

D'abord, il y a des opérateurs d'intégrale singulière qui ne sont pas des CZO. Pour commencer, il faut se méfier des opérateurs dont le noyau-distribution est porté par la diagonale (en clair, associés au noyau $K(x, y) = 0$): les opérateur différentiels, ou même l'opérateur de multiplication ponctuelle par une fonction m non bornée, sont des SIO non bornés.

Pour éviter cet ennui, on peut se donner un noyau standard K, et demander quand les opérateurs tronqués T_ε, $\varepsilon > 0$, sont bornés sur L^2 avec une borne indépendante de ε.

Pour $n = 1$ déjà, on voit que $K(x, y) = |x - y|^{-1}$ ne marche pas (calculer $T_\varepsilon(\mathbf{1}_{[0,1]})$). Donc il est utile d'avoir une certaine forme de cancellation (où K change de signe et certaines contributions en annulent d'autres).

Quand K est un noyau de convolution (i.e., quand $K(x, y) = k(x-y)$ pour une fonction k), ou plus généralement quand T est l'opérateur de convolution avec une distribution k qui coïncide avec une fonction localement intégrable hors de l'origine, on paut calculer en transformée de Fourier, et logiquement T est borné sur L^2 si et seulement si $\widehat{k}$ est bornée.

Dans le cas général (où T n'est pas un opérateur de convolution), la situation n'est pas si simple, et on s'est demandé un certain temps ce qui rendait les T_ε uniformément bornés sur L^2 (ou pas). Par exemple, il aurait été très pratique que ce soit le cas dès que K est antisymétrique $K(y, x) = -K(x, y)$. Mais ça n'est pas vrai (contrexemple de Y. Meyer je crois).

Un exemple qui a beaucoup motivé est celui du noyau de Cauchy sur un graphe Lipschitzien de dimension 1. Plutôt que de travailler sur le graphe de la fonction Lipschitzienne $A : \mathbb{R} \to \mathbb{R}$, paramétrons-le de la manière éviente, et considérons le noyau K_A défini par

$$K_A(x, y) = \frac{1}{x + iA(x) - y - iA(y)}. \tag{1}$$

C'est un beau noyau standard antisymétrique, et on avait à l'époque toutes sortes de bonnes raisons de se demander si l'opérateur de valeur principale associé (voir ci-dessus) est borné. La réponse est oui, obtenue en plusieurs étapes retentissantes, dont le cas particulier où $||A'||_\infty$ est assez petit (Calderón, 1977) et le cas général (Coifman, McIntosh, Meyer, 1982). C'est aussi (a posteriori) un bel exemple de trivialisation (relative) des maths.

11.b. Faible bornitude

On définit maintenent une propriété en principe notablement plus faible que la continuité L^2, destinée surtout à vérifier que T se comporte avec la bonne homogénétié vis-à-vis des translations et dilatations. On définit des semi-normes $N_{x,r,m}$ destinées à contrôler des fonctions sur $B(x, r)$ par

$$N_{x,r,m}(f) = \sum_{0 \leq j \leq m} \sum_D r^{\frac{n}{2}+j} ||Df||_{L^\infty(B(x,r))}, \tag{2}$$

où la somme se fait sur les opérateurs différentiels de la forme $D = \partial_1^{\alpha_1} \cdots \partial_n^{\alpha_n}$, avec $\alpha_1 + \cdots \alpha_n = j$ (des dérivées partielles d'ordre j). On a calculé l'homogénéité de manière à obtenir $||f||_{L^2(B(x,r))} \leq CN_{x,r,m}(f)$ trivialement, et que chaque terme ait la même homogénétié vis-à-vis des dilatations.

Definition 1. *On dira que l'opérateur T, défini sur $C_c^\infty(\mathbb{R}^n)$ et à valeurs dans son dual, est faiblement borné s'il existe $m \geq 0$ et $C \geq 0$ tels que*
(3)
$|\langle Tf, g\rangle| \leq CN_{x,r,m}(f)N_{x,r,m}(g)$ *pour $f, g \in C_c^\infty(\mathbb{R}^n)$ à supports dans $B(x,r)$.*

Bref, comme dans la définition de T, on s'autorise à utiliser autant de dérivées de f et g qu'on veut, mais cette fois on exige des estimations uniforme (avec l'homogénéité qui marcherait pour L^2) en fonction des supports.

En général, prendre $m = 1$ est largement suffisant, mais on ne sait jamais. On garde m quelconque ici comme on a gardé $C_c^\infty(\mathbb{R}^n)$ plus haut.

Exemple: si T est borné sur L^2, il est faiblement borné, et on peut pendre $m = 0$. Vérification triviale puisque $||f||_{L^2(B(x,r))} \leq CN_{x,r,0}(f)$.

Par contre, l'opérateur de dérivation (défini par $Tf = f'$) sur $\mathbb{R}$ n'est pas faiblement borné. Prendre une fonction $\varphi \in C^\infty$, à support dans $B(0,1)$, puis essayer d'appliquer (3) avec $x = 0$, $r > 0$, et $f = g = \varphi(\cdot/t)$. Le membre de gauche est $t^{n-1}\int\varphi\varphi'$, et le membre de droite est $t^n N_{0,1,m}(\varphi)$, ce qui ne colle pas.

Exercice Vérifier que pour $m \in L^1_{loc}(\mathbb{R}^n)$, l'opérateur de multiplication ponctuelle par m est faiblement borné si et seulement si m est (égale presque partout à une fonction) bornée. Faites comme ci-dessus, et appliquez encore le théorème de différentiation de Lebesgue).

Exercice Vérifier que si T est l'opérateur de valeur principale associé au noyau standard antisymétrique K, alors K est faiblement borné. Notez au passage qu'on utilise seulement l'estimation (9.2) sur la taille du noyau, et pas la régularité donnée par (9.3) et (9.4). L'exercice est facile: il s'agit simplement de noter que la démonstration d'existence qui mène à (9.11) donne aussi (3) avec $m = 1$.

11.c. Le théorème $T(1)$

Théorème 2 [D, Journé]. *Soit T in SIO. Alors T est un CZO si et seulement si T est faiblement borné, $T(1) \in BMO$, et $T^t(1) \in BMO$.*

On a vu que T^t, défini par (9.36), est aussi in SIO, et que les conditions sont nécessaires.

Pour bien faire, il faut encore définir $T(1)$ et $T^t(1)$ (comme des distributions définies modulo une constante additive, puisqu'on n'a défini Tf, $f \in L^\infty$, que quand T est un CZO. Et une distribution définie modulo une constante additive, c'est juste une forme linéaire continue sur les fonctions test d'intégrale nulle. Donc il faut dire qui est $\langle Tf, g\rangle$ pour $f \in C^\infty$ bornée et $g \in C_c^\infty$ d'intégrale nulle. On choisit $\varphi \in C_c^\infty$, égale à 1 sur une grosse boule contenant le support de g, et on écrit

$$\langle Tf, g\rangle = \langle T(\varphi f), g\rangle + \langle T((1-\varphi)f, g\rangle, \tag{4}$$

où $\langle T(\varphi f), g\rangle$ est bien défini parce que φf est une fonction test, et

$$\langle T((1-\varphi)f), g\rangle = \int\int [K(x,y) - K(x_0,y)](1-\varphi(y))f(y)g(x)dxdy \tag{5}$$

On s'est permis de choisir x_0 dans le support de g (de sorte que $1-\varphi(y) = 0$ près de x) et de soustraire $K(x_0, y)$ pour faire converger l'intégrale, parce que $\int K(x_0,y)g(x)dx = 0$ pour tout y tel que $1-\varphi(y) \neq 0$. Maintenant l'intégrale converge, puisque pour tout x dans le support de g,

(6)

$$\int \big|[K(x,y) - K(x_0,y)](1-\varphi(y))f(y)\big|dy$$
$$\leq C||f||_\infty |x-x_0|^\delta \int |y-x_0|^{-n-\delta}(1-\varphi(y))dy \leq C$$

où C dépend de $||f||_\infty$ et de la distance du support de g à celui de $1-\varphi$, mais ça n'est pas grave.

Exercice Vérifier que la définition de $\langle Tf, g\rangle$ ne dépend pas du choix de φ, et qu'elle n'est pas incompatible avec celle de Tf donnée plus haut quand T est un CZO et $f \in L^\infty$. Noter en tout cas qu'on n'a pas changé Tf quand f est une foncton test.

La démonstration va prendre un certain temps, et peut-être d'ailleurs qu'on ne va pas vérifier tous les détails. Le plan est de démontrer d'abord le théorème dans un cas particulier plus simple, quand $T(1) = T^t(1) = 0$, puis de s'y ramener en soustrayant des opérateurs bornés connus qu'on aura construits pour l'occasion.

A mon avis, il y a une démonstraition meilleure que toutes les autres, parce qu'elle est simple et s'adapte très bien. C'est celle de Coifman et Meyer présentée ci-dessous.

11.d. Lemme de Schur et une variante

On utilisera peut-être le lemme suivant sur la continuité des opérateurs définis par des matrices à décroissance rapide loin de la diagonale.

Théorème (Schur). *On se donne un ensemble d'indices I, une matrice $A = ((a_{i,j}))$ (avec $i, j \in I$), et on suppose qu'il existe des nombres $\omega_i > 0$, $i \in I$ et une constante $C \geq 0$ tels que*

$$\sum_{j \in I} |a_{i,j}| \omega_j \leq C \omega_i \quad \textit{pour tout } i \in I \tag{7}$$

$$\sum_{i \in I} |a_{i,j}| \omega_i \leq C \omega_j \quad \textit{pour tout } j \in I. \tag{8}$$

Alors A définit un poérateur borné sur $\ell^2(I)$, de norme au plus C, par $Ax = y$, où $x = \{x_j\}, j \in I$, $y = \{y_i\}, i \in I$, et $y_i = \sum_{j \in I} a_{i,j} x_j$.

Remarques avant la preuve: c'est un critère qui ne fait intervenir que la taille des $a_{i,j}$; on pourrait donc supposer qu'ils sont tous ≥ 0. Alors l'hypothèse est l'existence d'un sous-vecteur propre à coordonnées positives pour A et son transposé.

La première condition donne une continuité sur L^∞, la seconde sur L^1, et on pourrait interpoler. On peut aussi procéder directement, comme suit.

Donc on se donne x. On écrit $y_i \leq \sum_j |a_{i,j} x_j| = \sum_j (|a_{i,j} \omega_j|)^{1/2} (|a_{i,j} x_j^2 / \omega_j|)^{1/2}$ et on applique Cauchy-Schwarz:

$$|y_i|^2 \leq \{\sum_j |a_{i,j} \omega_j|\}\{\sum_j |a_{i,j} x_j^2 / \omega_j|\} \leq C \omega_i \{\sum_j |a_{i,j} x_j^2 / \omega_j|\}.$$

On somme et on trouve

$$|y|^2 = \sum_i |y_i|^2 \leq C \sum_i \sum_j |a_{i,j}| |x_j^2| \omega_i / \omega_j = C^2 \sum_j |x_j^2|$$

en sommant sur i d'abord. □

Vraissemblablement, on utilisera à la place du lemme de Schur le résultat asymétrique suivant, que je dois à Nazarov, Treil, Volberg, qui prétendent que c'est bien classique.

Théorème (?). *On se donne toujours un ensemble d'indices I, une matrice $A = ((a_{i,j}))$ (avec $i, j \in I$), et on suppose qu'il existe des constantes C_1 et C_2 telles que*

$$(a) \qquad \text{pour tout } i, \text{ il existe au plus } C_1 \text{ indices } j \in I \text{ tels que } a_{i,j} \neq 0,$$

et

$$(b) \qquad \text{pour tout } j \in J, \ \sum_{i \in I} |a_{i,j}|^2 \leq C_2^2.$$

Alors A définit (par la même formule que plus haut) un opérateur borné sur $\ell^2(I)$, de norme au plus $C_1 C_2$.

Pour la démonstration, on commence par couper A en au plus C_1 matrices qui vérifient toujours (b) avec la même constante, mais vérifient (a) avec $C_1 = 1$. On se réduit ainsi au cas de $C_1 = 1$.

Ensuite, notons v_j le vecteur de $\ell^2(I)$ dont les coordonnées sont les $a_{i,j}$. C'est donc le j-ième vecteur colonne de la matrice. Les vecteurs v_j sont orthogonaux, parce qu'ils ont des supports disjoints: si pour un certain i, $a_{i,j}$ et $a_{i,k}$ sont tous les deux non nuls, (a) dit que $j = k$.

Notons encore $x = \{x_j\}$. Alors $y = Ax = \sum_j x_j v_j$ (combinaison linéaire des vecteurs colonne), et ensuite $||y||^2 = \sum_j |x_j|^2 ||v_j||^2$ par orthogonalité. Comme $||v_j||^2 \leq C_2^2$ par (b), on trouve $||y||^2 \leq C_2^2 ||x||^2$, et on en déduit le théorème. □

11.e. Définition de T sur les fonctions simples

On va supposer que T est un SIO faiblement borné, et on va estimer les coefficients de sa matrice dans la base de Haar. Après on appliquera le lemme de Schur.

Commencons par prouver que pour $Q, R \in \mathcal{D}$ (deux cubes dyadiques) disjoints, on peut définir

$$\langle T\mathbf{1}_Q, \mathbf{1}_R \rangle = \int_{y \in Q} \int_{x \in R} K(x, y) dx dy \tag{9}$$

avec convergence absolue de l'intégrale. En fait, pour $x \notin \overline{Q}$, on peut définir

$$T\mathbf{1}_Q(x) = \int_{y \in Q} K(x, y) dy \tag{10}$$

et en plus, si l'on note $d(x,Q)$ la distance de x à Q,
(11)

$$|T\mathbf{1}_Q(x)| \leq C \int_{y\in Q} |x-y|^{-n}dy \leq C \int_{d(x,Q)\leq|x-y|\leq d(x,Q)+\mathrm{diam}(Q)} |x-y|^{-n}dy$$
$$\leq C \log\big(1+\frac{\mathrm{diam}(Q)}{d(x,Q)}\big).$$

Ensuite, il ne reste plus qu'à intégrer en x pour tomber sur (9).

Notons encore que

$$(12) \qquad |\langle T\mathbf{1}_Q, \mathbf{1}_R\rangle| \leq C|Q|$$

quand Q et R sont des cubes distincts de même génération. On intègre (11) et c'est ce qui tombe.

Toutes ces formules ont un sens, au sens où si l'on approche $T\mathbf{1}_Q$ et $\mathbf{1}_R$ par des fonctions tests f_k et g_k (disons, à support dans Q et R, puis qu'on passe à la limite après avoir noté que $\langle Tf_k, g_k\rangle = \int\int f_k(y)g_k(x)dxdy$, on trouve toujours la même limite qui est le membre de droite de (9).

Ceci nous permet de définir $\langle Tf, g\rangle$ quand f et g sont des combinaisons linéaires de fonctions caractéristiques de cubes dyadiques, et ont des supports disjoints. On va maintenant donner une définition de $\langle T\mathbf{1}_Q, \mathbf{1}_Q\rangle$ quand $Q \in \mathcal{D}$, qui nous permettra d'omettre cette hypothèse de support. En même temps, on montrera que

$$(13) \qquad |\langle T\mathbf{1}_Q, \mathbf{1}_Q\rangle| \leq C|Q| \ \text{ pour } Q \in \mathcal{D}.$$

Faisons comme si on avait déjà traité le chapitre 12 sur les cubes de Whitney, donnons nous un cube Q, et notons $\{Q_j\}$, $j \in J$, la collection des cubes de whitney associés à l'ouvert $\mathrm{int}(Q)$. Ainsi, $\mathrm{diam}(3Q_j) \sim \mathrm{dist}(Q_j, \mathbb{R}^n \setminus Q)$ Notons encore $\{\varphi_j\}$ la partition de l'unité associée, choisie de sorte que

$$(14) \qquad \sum_j \varphi_j = \mathbf{1}_Q,$$

chaque φ_j est à support dans $2Q_j$, et

$$(15) \qquad |\nabla^k \varphi_j(x)| \leq C\mathrm{diam}(Q_j)^{-k} \ \text{ pour } k \leq m,$$

où m est comme dans la condition de faible bornitude. On va écrire

$$\langle T\mathbf{1}_Q, \mathbf{1}_Q\rangle = \sum_{i\in J}\sum_{j\in J}\langle T\varphi_i, T\varphi_j\rangle \tag{16}$$

avec convergence absolue.

On commence par les termes les plus intéressants pour lesquels $3Q_i \cap 3Q_j \neq \emptyset$. Pour chaque i, il y a au plus C indices j comme ceci (par définition des cubes de Whitney), et pour chacun Q_j est de diamètre comparable à celui de Q_i. On va vérifier que pour chacun,

$$|\langle T\varphi_i, T\varphi_j\rangle| \leq C|Q_i|. \tag{17}$$

Comme nos deux fonctions sont supportées dans une même boule $B(x,r)$ de rayon $C\mathrm{diam}(Q_j)$, on utilise la faible bornitude dans cette boule, et il s'agit donc d'estimer $N_{x,r,m}(\varphi_i)$ et $N_{x,r,m}(\varphi_j)$. On trouve $N_{x,r,m}(\varphi_i) \leq C\mathrm{diam}(Q_j)^{n/2}$ et pareil pour φ_j. Alors (17) s'en déduit aussitôt.

La contribution totale de tous les termes de ce type est au plus $C\sum_i |Q_i| \leq C|Q|$, ce qui est compatible avec (13).

On passe aux termes rectangles ou $3Q_i$ ne rencontre pas $3Q_j$. Pour ces termes, on utilisera le noyau, et le fait que

$$|\langle T\varphi_i, T\varphi_j\rangle| \leq C\int_{2Q_i}\int_{2Q_j}|K(x,y)|dxdy \leq C\int_{2Q_i}\int_{2Q_j}|x-y|^{-n}dxdy. \tag{18}$$

On fixe i, et on somme sur j. Il vient

$$\begin{aligned}\sum_{j;3Q_i\cap 3Q_j=\emptyset}|\langle T\varphi_i, T\varphi_j\rangle| &\leq C\int_{3Q_i}\int_{Q\setminus 4Q_i}|x-y|^{-n}dxdy \\ &\leq C\int_{Q_i}\int_{C^{-1}\,\mathrm{dist}(x,\mathbb{R}^n\setminus Q)\leq|x-y|\leq\mathrm{diam}(Q)}|x-y|^{-n}dydx \\ &\leq C\int_{Q_i}\log\Big(\frac{C\mathrm{diam}(Q)}{\mathrm{dist}(x,\mathbb{R}^n\setminus Q)}\Big)dx.\end{aligned} \tag{19}$$

Il reste à sommer sur i, mais on note que les Q_i sont disjoints et contenus dans Q, de sorte que

$$\sum_i\sum_{j;3Q_i\cap 3Q_j=\emptyset}|\langle T\varphi_i, T\varphi_j\rangle| \leq C\int_Q\log\Big(\frac{C\mathrm{diam}(Q)}{\mathrm{dist}(x,\mathbb{R}^n\setminus Q)}\Big)dx \leq C|Q|. \tag{20}$$

Ceci termine la vérification de la convergence de la grande somme dans (16), avec un résultat inférieur à $C|Q|$. Quand on approxime $\mathbf{1}_Q$ par des sommes finies $\sum_{i\in I_0} \varphi_i$, les objets $\langle T\sum_{i\in I_0} \varphi_i, \sum_{i\in I_0} \varphi_i\rangle$ sont bien définis, majorés comme plus haut, et convergent. D'où une définition relativement cohérente de $\langle T\mathbf{1}_Q, \mathbf{1}_Q\rangle$, avec l'estimation (13).

A cause de tout ce qui précède, on peut définir $\langle Tf, g\rangle$ lorsque f et g sont des fontions simple (combinaisons liéaires de fonctions caractéristiques de cubes dyadiques). Et en plus, le résultat est linéaire en f et en g (vérification agréable laissée en exercice). On va montrer dans le paragraphe suivant que

$$|\langle Tf, g\rangle| \le C||f||_2||g||_2 \tag{21}$$

pour f et g simple, ce qui donnera le prolongement continu sur L^2 souhaité (nouvel exercice: vérifier que ce prolongement coïncidera automatiquement avec T sur les fonctions test), et ceci sous la condition supplémentaire que

$$T1 = T^t 1 = 0. \tag{22}$$

11.f. La matrice de T quand $T1 = T^t(1) = 0$

Noter que les fonctions $h_{Q,\varepsilon}$ de la base de Haar sont des fonctions simples. On va calculer la matrice de (notre première extension de) T dans cette base. Noter que si f et g sont des fonctions simples d'intégrale nulle (et cette classe aussi est dense, donc on peut s'en contenter), elles ont des développements finis dans la base, de la forme $f = \sum_{Q,\varepsilon} f_{Q,\varepsilon} h_{Q,\varepsilon}$ et $g = \sum_{Q,\varepsilon} g_{Q,\varepsilon} h_{Q,\varepsilon}$, et alors

$$\langle Tf, g\rangle = \sum_{Q,\varepsilon}\sum_{Q',\varepsilon'} c_{Q,\varepsilon,Q',\varepsilon'}\, f_{Q,\varepsilon}\, g_{Q',\varepsilon'}, \tag{23}$$

avec

$$c_{Q,\varepsilon,Q',\varepsilon'} = \langle Th_{Q,\varepsilon}, h_{Q',\varepsilon'}\rangle \tag{24}$$

Comme $||f||_2^2 = \sum |f_{Q,\varepsilon}|^2$ et pareil pour g, on voit qu'il va suffire de prouver que la matrice des coefficents $c_{Q,\varepsilon,Q',\varepsilon'}$ définit un opérateur borné sur $\ell^2(I)$, où $I = \mathcal{D}\times\mathcal{E}'$ est l'ensemble des paires Q, ε.

Et bien sûr on va utiliser le lemme de Schur ou de NTV pour faire ceci. L'idée sera donc de majorer les $c_{Q,\varepsilon,Q',\varepsilon'}$, en montrant qu'ils décroissent quand on s'éloigne de la diagonale, et après de sommer contre des poids convenables.

Aucun de nos caluls ne dépendra de ε ou ε'. Donc on va noter h_Q et $h_{Q'}$ au lieu de $h_{Q,\varepsilon}$ et $h_{Q',\varepsilon'}$, et $c_{Q,Q'}$ au lieu de $c_{Q,\varepsilon,Q',\varepsilon'}$.

Et il va rester à distinguer plein de cas et estimer $c_{Q,Q'}$ dans chacun. En fait, on va poser $Q' = R$ et estimer les $c_{Q,R}$.

On note $d(Q)$ la taille de Q, et pareil pour $d(R)$.

Noter que si T vérifie nos hypothèses, T^t aussi, et les coefficients $c_{Q,R}$ associés à T^t sont exactement les $c_{R,Q}$. Donc on va estimer $c_{Q,R}$ quand $d(R) \geq d(Q)$, et on aura automatiquement la même estimation pour $c_{R,Q}$.

Cas 1. On suppose que $d(R) \geq d(Q)$ (comme toujours) et que Q et R sont disjoints. On veut montrer qu'alors

$$(25) \quad |c_{Q,R}| \leq C|Q|^{1/2}|R|^{-1/2}\Big\{\frac{d(Q)}{d(Q)+\mathrm{dist}(Q,R)}\Big\}^{\delta} \ \text{ si } \ \mathrm{dist}(Q,r) \leq d(R),$$

$$(26) \quad |c_{Q,R}| \leq C|Q|^{1/2}|R|^{1/2}d(Q)^{\delta}\,\mathrm{dist}(Q,R)^{-n-\delta} \ \text{ si } \ \mathrm{dist}(Q,R) \geq d(R),$$

Dans tout ce calcul, on va utiliser le noyau, et dire que

$$(27) \qquad c_{Q,R} = \langle Th_Q, h_R\rangle = \int_Q\int_R K(x,y)h_Q(y)h_R(x)dydx.$$

Pour l'intégrale sur $R\cap 2Q$, on utilise juste la taille du noyau, le fait que $||h_Q||_\infty \leq C|Q|^{-1}$, pareil pour R, et alors

(28)

$$\int_Q\int_{R\cap 2Q}|K(x,y)h_Q(y)h_R(x)|dydx \leq C|Q|^{-1/2}|R|^{-1/2}\int_Q\int_{R\cap 2Q}|x-y|^{-n}dxdy$$
$$\leq C|Q|^{-1/2}|R|^{-1/2}\int_Q \log\Big(1+\frac{d(Q)}{\mathrm{dist}(x,\mathbb{R}^n\setminus Q)}\Big) \leq C|Q|^{1/2}|R|^{-1/2}$$

puisque Q et R sont disjoints et par un calcul fait plus haut. Ceci rentre dans le cadre de (25), puisque si $2Q$ rencontre R, $\mathrm{dist}(Q,R) \leq Cd(Q)$. Evidemment (26) n'est pas concerné puisque cans ce cas $R\cap 2Q = \emptyset$.

Pour l'intégrale hors de $2Q$, on estime d'abord $Th_Q(x)$, en utilisant le fait que h_Q est à support dans Q et d'intégrale nulle:

(29)

$$|Th_Q(x)| = |\int_Q K(x,y)h_Q(y)|dy = |\int_Q [K(x,y_Q)-K(x,y)]h_Q(y)|dy$$
$$\leq C|Q|^{1/2}d(Q)^{\delta}\int_Q \frac{1}{|x-y_Q|^{n+\delta}}dy \leq C|Q|^{1/2}d(Q)^{\delta}\,\mathrm{dist}(x,Q)^{-n-\delta}$$

en notant y_Q le centre de Q. Il reste à intégrer ceci sur $R \setminus 2Q$, multiplier par $|R|^{-1/2}$ (pour la taille de h_R, et on trouve

$$|c_{Q,R}| \leq C|Q|^{1/2}|R|^{-1/2}d(Q)^{\delta} \int_{R\setminus 2Q} \mathrm{dist}(x,Q)^{-n-\delta}dx \tag{30}$$

Dans le cas de (26), $\mathrm{dist}(x,Q)$ reste équivalent à $\mathrm{dist}(Q,R)$ et on trouve (26) directement. Dans le cas de (25), l'intégrale sur $R \setminus 2Q$ porte sur la région où $\mathrm{dist}(x, y_Q) \geq \min(d(Q), \mathrm{dist}(Q,R))$ donc

(31)
$$\begin{aligned}\int_{R\setminus 2Q} \mathrm{dist}(x,Q)^{-n-\delta}dx &\leq C \int_{|x-y_Q|\geq \min(d(Q),\mathrm{dist}(Q,R))} |x-y_Q|^{-n-\delta}dx \\ &\leq C[d(Q)+\mathrm{dist}(Q,R)]^{-\delta},\end{aligned}$$

ce qui donne (25). [Fin du premier cas.]

Cas 2. On suppose toujours que $d(R) \geq d(Q)$, et maintenant que $Q \subset R$ (c'est le cas qui reste, puisqu'on parle de cubes dyadiques). Si $Q = R$, on trouve

$$|c_{Q,Q}| \leq C \tag{32}$$

en découpant h_Q en somme de fonctions caractéristiques, et en appliquant (12) et (13). [On récupère le facteur manquant $|Q|^{-1}$ à cause de la norme L^∞ de h_Q.]

Reste donc le cas oà Q est strictement inclus dans R. On ne peut pas directement utiliser le noyau, mais on va s'y ramener en utilisant le fait que $T^t 1 = 0$. On note $\mathcal{F}(R)$ la collection des enfants de R, et $S(Q)$ l'enfant de R qui contient Q (donc $S(Q) \in \mathcal{F}(R)$). Ensuite on écrit

$$h_R = \sum_{S\in\mathcal{F}(R)} \alpha_S \mathbf{1}_S \tag{33}$$

avec des coefficents α_S tels que $|\alpha_S| \leq C|R|^{-1/2}$. Et donc

$$c_{Q,R} = \sum_{S\in\mathcal{F}(R)} \alpha_S \langle Th_Q, \mathbf{1}_S\rangle. \tag{34}$$

Tous les termes pour lesquels $S \neq S(Q)$ pourront être traités directement, avec le noyau et comme précédemment, sauf celui de $S(Q)$, puisque $Q \subset S(Q)$.

Donc occupons nous de $S = S(Q)$. On utilise le fait que $T^t 1 = 0$ pour écrire que

$$(35) \qquad \langle Th_Q, \mathbf{1}_{S(Q)} \rangle + \langle h_Q, 1 - \mathbf{1}_{S(Q)} \rangle = 0.$$

On a défini $T^t(1)$ un peu différemment, donc il faut faire une petite vérification. D'abord, on obtiendra (35) dès qu'on aura prouvé que $\langle Th_Q, \varphi \rangle + \langle h_Q, 1-\varphi \rangle = 0$, où φ est une fonction test qui égale à 1 dans un voisinage de $S(Q)$, parce que la différence entre les deux formules consiste à ajouter et retirer la même intégrale $\int_Q \int_{\mathbb{R}^n \setminus S(Q)} K(x,y) h_Q(x)(\varphi(y) - \mathbf{1}_{S(Q)}(y)) dydx$. Et cette nouvelle formule est presque la définition du fait que $T^t 1 = 0$, la seule différence étant que h_Q n'est pas régulière. On règle le problème comme plus haut, en écrivant h_Q comme somme de fonctions test (qu'on n'a même pas besoin de prendre d'intégrale nulle sauf erreur).

Bref, on a (35), et on écrit

$$(36) \qquad c_{Q,R} = \sum_{S \in \mathcal{F}(R), S \neq S(Q)} \alpha_S \langle Th_Q, \mathbf{1}_S \rangle - \alpha_{S(Q)} \langle h_Q, 1 - \mathbf{1}_{S(Q)} \rangle.$$

Et maintenant on traite ces termes comme ci-dessus; celui qui donne la plus grosse contribution est le dernier, où la preuve de (25) donne

$$(37) \quad |\alpha_{S(Q)} \langle h_Q, 1 - \mathbf{1}_{S(Q)} \rangle| \leq C |Q|^{1/2} |R|^{-1/2} \left\{ \frac{d(Q)}{d(Q) + \mathrm{dist}(Q, \mathbb{R}^2 \setminus S(Q))} \right\}^{\delta}.$$

[On reprend juste les mêmes estimations, en notant que dans (31) on pouvait aussi intégrer sur tout $\mathbb{R}^n \setminus 2Q$].

Voilà pour les estimations des $c_{Q,R}$. Il reste à voir que ces estimations donnent la continuité sur $\ell^2(I)$ de la matrice dont les coefficients sont les $c_{Q,R}$ (ou plus précisément, les $c_{Q,\varepsilon,R,\varepsilon'}$).

Une méthode consiste à appliquer le lemme de Schur, avec les poids $\omega(Q) = |Q|^{1/2} d(Q)^{-\delta/2}$. On va laisser la vérification en exercice.

Une autre méthode plus jolie consiste à utiliser le critère classique de NTV.

On ne l'applique pas directement à la matrice toute entière. D'abord, on coupe cette matrice en deux, en ne gardant que les termes où $|Q| \leq |R|$. Et ensuite, on coupe la matrice en tranche $A_{k,l}$, pour lesquelles $d(R) = 2^k d(Q)$ (avec donc $k \geq 0$) et $Q \subset R + ld(R)$, $l \in \mathbb{R}^n$, et on va appliquer le théotème à chaque $A_{k,l}$.

Donc on fixe $k \geq 0$ et $l \in \mathbb{Z}^n$. On note que pour chaque cube $Q \in \mathcal{D}$, il y a un seul cube $R \in \mathcal{D}$ tel que $d(R) = 2^k d(Q)$ et $Q \subset R + l d(R)$, donc on a la condition (a) avec $C_1 = (2^n - 1)^2$ (à cause des choix de ε et ε' possibles). Il reste à fixer R et à sommer $|c_{Q,R}|^2$ sur Q.

On commence par le cas où $l = 0$ et $Q \subset R$ (notre second cas). On applique (37) et il sort

$$C|R|^{-1} \sum_{Q \subset R} |Q| \Big\{ \frac{d(Q)}{d(Q) + \mathrm{dist}(Q, \mathbb{R}^2 \setminus S(Q))} \Big\}^{2\delta}$$

On redécoupe ceci en tranches en fonction de la distance de Q à $\mathbb{R}^2 \setminus S(Q)$. Quand celle-ci est de l'ordre de $2^m d(Q)$, $0 \leq m \leq k$, les cubes Q concernés ont une mesure totale $\leq C 2^{m-k}|R|$ (une fine couche annulaire d'épaisseur environ $2^{m-k} d(R)$ autour des frontières des enfants de R), de sorte que la contribution totale est au plus $C \sum_{0 \leq m \leq k} 2^{m-k}\, 2^{-2\delta m}$ puisque $\frac{d(Q)}{d(Q)+\mathrm{dist}(Q,\mathbb{R}^2 \setminus S(Q))} \sim 2^{-m}$ pour ces cubes. En supposant $\delta < 1/2$, ce qui ne coûte rien, le plus gros terme est pour $m = k$, mais quand on somme sur m il reste $C 2^{-2\delta k}$. On reprend la racine, et on trouve que la norme de $A_{k,0}$ est $\leq C 2^{-\delta k}$, ce qui donne un résultat agréablement sommable.

Le cas où $l \neq 0$, mais $|l| \leq C$ se traite pareillement, sauf qu'on est maintenant dans le cas 2, et qu'on peut utiliser (26) (ou, le cas échéant, (27) qui est encore meilleur, mais à peine). On trouve encore que la norme de $A_{k,0}$ est $\leq C 2^{-\delta k}$.

Pour les autres cas, Q est loin de R, on utilise (26), et la somme des carrés est dominée par

$$\begin{aligned} C \sum_{Q} |Q||R| d(Q)^{2\delta}\, \mathrm{dist}(Q,R)^{-n-\delta} &\leq C \sum_{Q} |Q||R|\ [2^{-k} d(R)]^{2\delta}\ [|l| d(R)]^{-2n-2\delta} \\ &\leq C 2^{-2\delta k} |l|^{-2n-2\delta}. \end{aligned}$$

Ensuite on prend la racine pour avoir C_2, mais ça reste sommmable en k et en l, et termine la démonstration de notre cas particulier: T a une extension bornée sur L^2 dans la cas où, en plus des hypothèses du théorème $T(1)$, on a que $T1 = T^t 1 = 0$.

L'air de rien, on a quand même utilisé ici une propriété importante des cubes dyadiques (la propriété dite de petite frontière). En effet, pour que la contribution des petits cubes Q situés près du bord de R soit assez petite, on a utilisé le fait que

(37.5)
$$\left|\{x \in R\,;\, \mathrm{dist}(x, \mathbb{R}^n \backslash R) \leq \tau d(R)\}\right| + \left|\{x \in \mathbb{R}^n \backslash R\,;\, \mathrm{dist}(x, \mathbb{R}) \leq \tau d(R)\}\right| \leq C\tau^\alpha |R|$$

pour $0 < \tau \leq 1$ et un certain $\alpha > 0$, qu'on peut d'ailleurs ici prendre égal à 1. On a en fait utilisé (37.5) deux ou trois fois (comme pour (12) ou vers (20)), par exemple pour faire converger des intégrales $\int_Q \int_R |K(x,y|$, mais c'est un peu la seule façon dont la géométrie intervient.

11.g. Les paraproduits

En fait il s'agira de forme assez shématisées des paraproduits introduits par J.-M. Bony (dans un cadre différent). L'idée d'utiliser ces opérateurs, autant que la version simplifiée, est due à Y. Meyer.

On se donne une fonction $\beta \in BMO$, et on va construire un CZO (ou presque) $P = P_\beta$ tel que $P1 = \beta$, et $P^t 1 = 0$. Quand ce sera fait, on pourra conclure en disant que si T est comme au théorème $T(1)$, et si on pose $\beta = T1$ et $\gamma = T^t 1$, alors $\widetilde{T} = T - P_\beta - P_\gamma^t$ est encore un SIO faiblement borné, et de plus $\widetilde{T}1 = \widetilde{T}^t 1 = 0$, donc $\widetilde{T}$ est borné sur L^2, donc T aussi.

Pour $Q \in \mathcal{D}$, notons $\alpha_{Q,\varepsilon}$ les coefficients de β dans la base de Haar. On se souvient qu'on a la condition de Carleson

$$\sum_{Q \subset R} \sum_{\varepsilon} |\alpha_{Q,\varepsilon}|^2 \leq C|R|\ ||\beta||_{BMO} \tag{38}$$

pour $R \in \mathcal{D}$. Posons aussi $\theta_Q = |Q|^{-1} \mathbf{1}_Q$ pour $Q \in \mathcal{D}$. On veut définir P par son noyau

$$P(x,y) = \sum_{Q,\varepsilon} \alpha_{Q,\varepsilon}\, h_{Q,\varepsilon}(x)\, \theta_Q(y) \tag{39}$$

ou encore

$$Pf = \sum_{Q,\varepsilon} \alpha_{Q,\varepsilon}\, \langle f, \theta_Q \rangle\, h_{Q,\varepsilon}, \tag{40}$$

en espérant que toutes ces sommes convergent.

En fait, si f est une fonction simple (une combinaison linéaire finie de fonctions caractéristiques de cubes), $\langle f, \theta_Q \rangle$ est nul sauf quand Q est contenu dans une boule fixe, et alors il est borné (car $f \in L^\infty$). Alors les coefficients

$\alpha_{Q,\varepsilon}\,\langle f,\theta_Q\rangle$ sont de carré sommable à cause de (38), donc dans ce cas la série de (40) converge dans L^2.

On commence par vérifier la taille du noyau, c.-à.-d. que pour x et $y \neq x$ fixés, la série dans (39) converge, et

$$P(x,y) \leq C|x-y|^{-n}. \tag{41}$$

En effet, le terme avec $h_{Q,\varepsilon}(x)\,\theta_Q(y)$ n'existe que si x et y sont dans Q. Donc il y a au plus un terme par échelle (par valeur de $d(Q)$) et par valeur de ε, et sa taille est au plus $C|\alpha_{Q,\varepsilon}||Q|^{-1}|h_{Q,\varepsilon}(x)| \leq C|\alpha_{Q,\varepsilon}||Q|^{-3/2} \leq C|Q|^{-1}$ à cause de (38) (avec $R = Q$). Comme $d(Q) \geq C^{-1}|x-y|$ car x et y sont dans Q, on peut sommer sur Q, et it reste $\leq C|x-y|^{-n}$. Donc on a (41).

Notons que P vérifie l'analogue de la condition de base pour les noyaux (le fait que $\langle Tf,g\rangle = \int\int K(x,y)f(y)g(x)$), avec les fonctions simples au lieu des fonctions test. On en déduit aisément la même chose pour les fonctions test, par un petit passage à la limite (en utilisant juste la densité des fonctions simples et l'estimation (41)).

Ensuite on devrait vérifier la régularité du noyau, mais mauvaise surprise, elle est juste un peu insuffisante. Par contre, on a la propriété suivante qui va remplacer:

$$Ph(x) = P^t h(x) = 0 \;\text{ pour } x \notin R. \tag{42}$$

dès que pour $R \in \mathcal{D}$ et h une fonction simple bornée supportée par R et telle que $\int h = 0$.

La vérification est facile. Commençons par P. Dans (40), à cause du coefficient $\langle f,\theta_Q\rangle$, seuls les cubes Q qui rencontrent R ont une contribution non nulle. Et même, ils ne peuvent contenir R puisque $\int_R h = 0$. Donc $Q \subset R$, et alors $h_{Q,\varepsilon}(x) = 0$, hors de R, comme souhaité.

Pour P^t, on commence par observer que pour f et h simples

$$\langle Pf,h\rangle = \sum_{Q,\varepsilon} \alpha_{Q,\varepsilon}\,\langle f,\theta_Q\rangle\,\langle h,h_{Q,\varepsilon}\rangle, \tag{43}$$

de sorte que

$$P^t h = \sum_{Q,\varepsilon} \alpha_{Q,\varepsilon}\,\langle h,h_{Q,\varepsilon}\rangle\,\theta_Q \tag{44}$$

(par densité des fonction simples). Ensuite, les seuls cubes Q qui contribuent dans (44) sont ceux pour lesquels Q rencontre R (à cause de $\langle h, h_{Q,\varepsilon}\rangle$, et sans être strictement plus grand (car alors $h_{Q,\varepsilon}$ serait constante sur R et $\int hh_{Q,\varepsilon}$ serait nulle), ce qui donne (42) dans ce cas aussi.

On pourra se contenter de (42), au lieu de la régularité du noyau, parce que dans la première partie de la démonstration, que l'on veut maintenant appliquer à $\widetilde{T} = T - P_\beta - P_\gamma^t$, on a uniquement utilisé la régularité du noyau dans (29) (et son analogue pour le cas 2, et maitenant on a juste une somme de trois termes, un qui est petit comme avant, et les deux autres qui sont nuls.

Maintenant on va s'occuper de la continuité L^2. On se donne f simple, et on note que (40) dit que les coefficients de Pf dans la base de Haar sont les $\alpha_{Q,\varepsilon}\,\langle f, \theta_Q\rangle$, de sorte que

$$||Pf||_2^2 = \sum_{Q,\varepsilon} |\alpha_{Q,\varepsilon}|^2\,|\langle f, \theta_Q\rangle|^2. \tag{45}$$

Et donc il s'agit simplement de majorer le membre de droite, ce qu'on fera en utilisant de théorème de Carleson. Comme on a discrétisé à outrance, on va faire ceci dans un cadre presque trop simple.

D'abord, il nous faut une mesure de Carleson. On prend

$$\mu = \sum_{Q,\varepsilon} |\alpha_{Q,\varepsilon}|^2 \delta_Q, \tag{46}$$

où δ_Q est la mesure de Dirac au point $(y_Q, d(Q))$, et où y_Q désigne le centre de Q. C'est bien une mesure de Carleson: la mesure de la boîte de Carleson $R \times (0, d(R)]$ est en fait le second membre de (38).

Ensuite on veut une fonction de $(x,t) \in \mathbb{R}^n_+$. On pose $F(x,t) = |Q|^{-1}\int_Q |f|$ au point $(y_Q, d(Q))$ quand $Q \in \mathcal{D}$. Ce qu'on met ailleurs (par exemple 0) n'a pas d'importance, puisque μ ne le verra pas.

Ensuite on applique le corollaire 3 page 35 du théorème de Carleson, avec $\varphi = \mathbf{1}_B$, où B est une grosse boule centrée à l'origine. Le corollaire dit que $G(x,t) = \varphi_t * |f|(x)$ est dans $L^2(d\mu)$, avec une norme au plus $C||\mu||_{CM}^{1/2}||f||_2 \leq C||\beta||_{BMO}||f||_2$.

Par ailleurs, notre fonction F est dominée par G, donc

$$\int F(x,t)^2 d\mu(x,t) \leq C||\beta||_{BMO}^2 ||f||_2^2.$$

Or $\int F(x,t)^2 d\mu(x,t) = \sum_{Q,\varepsilon} |\alpha_{Q,\varepsilon}|^2 \big\{|Q|^{-1}\int_Q |f|\big\}^2 = \sum_{Q,\varepsilon} |\alpha_{Q,\varepsilon}|^2 |\langle f, \theta_Q\rangle|^2 = ||Pf||_2^2$, par (45). Bref,

$$||Pf||_2 \leq C||\beta||_{BMO}||f||_2 \tag{47}$$

comme souhaité.

Donc P a une extension continue sur L^2, et est presque un CZO (à cause de son petit défaut de régularité pour le noyau). Passons au calcul de $P1$ et $P^t 1$.

Faisons le calcul sans nous soucier de la convergence. On applique (40) avec $f = 1$, donc $\langle f, \theta_Q\rangle = 1$, donc $Pf = \sum_{Q,\varepsilon} \alpha_{Q,\varepsilon} h_{Q,\varepsilon} = \beta$. Et on laisse la érification que même avec notre définition formelle, il en est bien ainsi.

De même, quand on utilise (44) avec $h = 1$, et on tombe tout de suite sur 0 à cause de $\langle h, h_{Q,\varepsilon}\rangle$. A nouveau, le passage à la limite est laissé en exercice.

Fin de la démonstration de $T(1)$. $\square$

11.h. $T(b)$

D'abord on teste $T(1)$ sur le noyau de Cauchy $K_A(x,y) = (x+iA(x)-y-iA(y))^{-1}$ (où $A : \mathbb{R} \to \mathbb{R}$ est lipschitzienne), et l'opérateur de valeur principale associé.

Par antisymétrie, il suffit de prouver que $T1 \in BMO$ (puisque dans ce cas, T est automatiquement faiblement borné). Mauvaise nouvelle, on ne sait pas calculer $T1$. Cependant, on sait calculer $T(1+iA')$, et on trouve 0. Plus précisément, on trouve vraiment 0 si A est à support compact, ce qui en fait suffit pour obtenir l'estimation L^2 souhaitée, par un argument de passage à la limite. Sinon, on peut aussi se débrouiller pour vérifier que dans le cas général, $T(1+iA')$ est une constante (donc égal à 0 à une constante additive près).

Le genre de passage à la limite auquel je pense (et qui pourrait être utilisé plusieurs fois dans ce qui suit) est le suivant: on démontre d'abort la continuité sur L^2 de l'opérateur T souhaité, dans le cas particulier où A est à la fois régulière et à support compact, mais avec des estimations uniformes. Ce qui dit que $|\langle Tf, g\rangle| \leq C||f||\,||g||$ pour f, g des fonctions-test, et C ne dépendant que de $||A'||_\infty$, par exemple. Puis on se donne A générale, on l'écrit comme limite d'une suite $\{A_k\}$ tronquées et régularisées, et on note que pour f et g fixés, $\langle Tf, g\rangle$ est bien la limite des $\langle T_k f, g\rangle$. Par exemple, le passage à la limite doit bien se passer si l'on utilise le fait qu'on a des opérateurs de valeurs singulière, donnés par $2\langle Tf, g\rangle = \int\int K_A(x,y)[f(y)g(x) - f(x)g(y)]dxdy$, et pareil pour les T_k.

Si $||A'||_\infty$ est assez petit, on peut s'en trier par une inégalité priori. Supposons encore que A est de classe C^2 et à support compact, auquel cas on sait démontrer à la main que l'opérateur T est un CZO et que tout ce qu'on va écrire a un sens. On note $||K||$ la norme de noyau standard de K (la meilleure constante C qui intervient dans la définition de noyau standard de K, en prenant $\delta = 1$ par exemple). On note $||T||$ la norme de son extension sur L^2. Alors:

$$||T|| \leq C||K|| + C||T1||_{BMO}$$

à cause du théorème $T(1)$, et

$$||T1||_{BMO} = ||TA'||_{BMO} \leq C\big(||K|| + ||T||\big)\,||A'||_\infty$$

puisque T est un CZO. Quand on suppose que $||A'||_\infty \leq 1$, disons, alors $||K|| \leq C$, on peut mettre bout-à-bout les inégalités précédentes, et obtenir que

$$||T|| \leq C||K|| + C||K||\,||A'||_\infty + C||T||\,||A'||_\infty \leq C + C||T||\,||A'||_\infty \leq C + \frac{1}{2}\,||T||$$

si $||A'||_\infty$ est assez petit, ce qui permet de conclure après passage à la limite.

Autre calcul possible: écrire le développement de K_A en série, dont le terme n-linéaire est le noyau K_n donné par

$$K_n(x,y) = \Big(\frac{A(x) - A(y)}{x - y}\Big)^n \frac{1}{x-y} \tag{48}$$

L'opérateur T_n de valeur principale associé à K_n est appelé n-ième commutateur de Calderón. Quand $n = 0$, on retrouve (à une constante multiplicative près) la transformée de Hilbert. Et pour contrôler la norme $||T_n||$ par récurrence, on observe que, par un petit calcul avec une intégration par parties (et en supposant A à support compact pour simplifier)

$$T_{n+1}(1) = T_n(A'). \tag{49}$$

Ceci donne $||T_n|| \leq C^{n+1}||A'||_\infty^n$, et permet de conclure en sommant la série. Après un temps de reflexion, on peut même se convaincre que les deux calculs ci-dessus sont essentiellement pareils.

Donc $T(1)$ permet de démontrer le théorème de Calderón 77 (la continuité de l'opérateur de Cauchy pour $||A'||_\infty$ assez petit, mais pas celui de Coifman, McIntosh, et Meyer (sans restriction sur $||A'||_\infty$).

D'où l'introduction, à la demande générale, d'un théorème $T(b)$ qui marche directement dans le cas général, et qui permette de se contenter de calculer $T(1 + iA')$ dans le cas précédent.

Les notations pour $T(b)$ sont les suivantes. On se donne toujours un opérateur T, et aussi deux fonctions bornées à valeurs complexes b_1 et b_2.

Ici il est plus pratique de dire que T est défini et continu de $b_1 C_c^\infty(\mathbb{R}^n)$ et à valeurs dans le dual de $b_2 C_c^\infty(\mathbb{R}^n)$, ce qui signifie simplement que ce qui est bien défini est $\langle T(b_1 f), b_2 g\rangle$ pour f et g dans $C_c^\infty(\mathbb{R}^n)$.

On demande toujours qu'il existe un noyau standard K tel que

$$\langle Tf, g\rangle = \int\int K(x,y) f(y) g(x) dy dx \tag{50}$$

lorsque $f \in b_1 C_c^\infty(\mathbb{R}^n)$ et $g \in b_2 C_c^\infty(\mathbb{R}^n)$ ont des supports disjoints, et ceci nous donne une définition d'opérateur d'intégrale singulière associé à b_1 et b_2.

On aura besoin d'hypothèses sur b_1 et b_2. La plus simple est de demander que les deux soient bornées accrétives, c'est à dire que d'une part $b_i \in L^\infty$, et d'autre part il existe $c > 0$ telle que

$$\mathcal{R}(b_i(x)) \geq c \ \text{ pour presque-tout } x \in \mathbb{R}^n. \tag{51}$$

(la partie réelle). Au moins, $b(x) = 1 + iA'(x)$ ci-dessus vérifie (51). Mais il y a diverses conditions plus faibles qui marchent encore (para-accrétivité, pseudo-accrétivité, et d'autres encore). Par exemple, la démonstration suggérée ci-dessous marche très bien si l'on suppose qu'il existe $c > 0$ tel que

$$\left| \frac{1}{|Q|} \int_Q b_i(y) dy \right| \geq c \ \text{ pour tout cube dyadique } Q. \tag{52}$$

Noter que la valeur absolue est bien à l'extéreur de l'intégrale. Ceci implique bien que $|b(x)| \geq c$ presque-partout, par le théorème de densité de Lebesgue, mais la réciproque est fausse: $b(x) = e^{ix}$ ne marche pas, à cause des grands intervalles où la moyenne de b est trop petite.

Avec toutes ces notations, le résultat est le suivant.

Théorème $T(b)$**.** *On se donne b_1 et b_2 bornées et vérifiant (52), et un opérateur T d'intégrale singulière associé à b_1 et b_2. On suppose que $Tb_1 \in BMO$, $T^t b_2 \in BMO$, et que $M_{b_2} T M_{b_1}$ est faiblement borné, où l'on a noté*

M_{b_i} l'opérateur de multiplication ponctuelle par b_i. Alors T a une extension bornée sur L^2.

Quelques commentaires, mais pas de démontration.

D'abord, la réciproque est vraie, mais on le savait déjà: si T est borné, $M_{b_2}TM_{b_1}$ aussi, donc $M_{b_2}TM_{b_1}$ est faiblement borné, et T envoie L^∞ dans BMO.

Pour ceci et aussi la définition de Tb_1 et T^tb_2, il faudrait vérifier que ce qui a été fait plus haut marche encore avec nos définitions un peu différente, mais c'est vrai, et avec les mêmes démonstrations.

Ensuite, il faut adapter la démonstration de Coifman-Semmes au cas des fonctions b_i. On a en fait besoin d'une base de Haar adaptée à b_i, que l'on construit comme plus haut, mais en modifiant les opérateurs d'espérance conditionnelle (les projections E_k), et de projctions sur les W_k (les opérateurs $\Delta_k = E_{k+1} - E_k$ (voir (4) page 38)). Ici on utilise

$$F_k f = \frac{1}{E_k b}\, E_k(bf) \tag{53}$$

au lieu de $E_k f$, ou plus expicitement

$$F_k f(x) = (\int_Q b)^{-1} \int_Q bf \tag{54}$$

où Q est le cube de $\mathcal{D}_k$ qui contient x. Noter que $|E_k b| \geq c$ à cause de (52). Ensuite, on remplace Δ_k par $D_k = F_{k+1} - F_k$, et on vérifie consciencieusement que la construction de la base de haar, et ensuite la démonstration du théorème, passent encore sans ennuis. Pour le fait que $|f||_2^2 \sim \sum_Q ||D_k f||_2^2$, on a besoin quand même d'une petite estimation supplémentaire à base de mesure de Carleson (et du fait que b_i est bornée). Le fait qu'on utilise deux bases de Haar différentes, une pour décomposer f et l'autre pour décomposer g, ne pose pas de problème particulier. A la fin, on est juste amené à estimer des coefficients de matrice du type $\langle Tb_1 h_{Q,1,\varepsilon}, b_2 h_{R,1,\varepsilon}\rangle$, où les fonctions h viennent de deux systèmes de Haar différents.

Je crois que j'ai mis quelques informations supplémentaires dans mon Lecture notes in Mathematics, qu'on peut aussi consulter pour les deux remarques suivantes.

Il y a même une extension possible, avec quasiment la même démonstration, au cas où K, b_1 et b_2, sont à valeurs dans une algèbre de Clifford.

La démonstration de Coifman-Semmes est très bonne aussi parce qu'elle passe sans problème aux espaces de type homogènes. [Penser à des espaces métriques munis d'une mesure doublante, mais j'ai un peu menti en cours, c'est un poil plus général que ça, puisqu'on s'autorise aussi des presque-distances qui ne vérifient l'inégalité triangulaire qu'avec une constante.] La raison est simple: on peut construire sur ces espaces des décompositions de l'espace en sortes de cubes dyadiques, avec les propriétés utilisées dans la démonstration. En particulier, la propriété de petites frontières analogue à (37.5).

Voir M. Christ, puis d'autres auteurs, pour des théorèmes $T(b)$ avec b variable (en gros, une par cube dyadique, mais qui n'a plus besoin de vérifier (52) pour les autres cubes).

Enfin, signalons que même avec des mesures qui ne sont pas doublantes, on a encore de la théorie de Calderón-Zygmund possible. Résultats de Nazarov-Treil-Volberg et de X. Tolsa en particulier. Par contre, la technicité augmente assez brutalement.

12. CUBES ET EXTENSIONS DE WHITNEY

Pour ce chapitre aussi, la technique est plus important que le résultat, qu'on va choisir le plus simple possible pour illustrer.

12.a. Les cubes de Whitney

On commence par la description des cubes de Whitney. On se donne un ouvert $\Omega \subset \mathbb{R}^n$, avec $\Omega \neq \mathbb{R}^n$. On va faire un choix un peu arbitraire; d'autres sont possibles.

On note $\mathcal{W} = \mathcal{W}(\Omega)$ l'ensembles des cubes dyadiques Q de $\mathbb{R}^n$ qui sont contenus dans Ω, et même tels que

$$\mathrm{dist}(Q, \partial\Omega) \geq 10\,\mathrm{diam}(Q) \tag{1}$$

et qui sont maximaux parmi les cubes dyadiques vérifiant (1).

Bien sûr $\mathrm{diam}(Q)$ est le diamètre de Q, donc $\sqrt{n}$ fois son côté, qui est une puissance de 2. On va lister quelques propriétés simples de $\mathcal{W}$. D'abord, le père Q^* de Q ne vérifie pas (1), donc $\mathrm{dist}(Q^*, \partial\Omega) < 10\,\mathrm{diam}(Q^*) = 20\,\mathrm{diam}(Q)$. Et $\mathrm{dist}(Q, \partial\Omega) \leq \mathrm{dist}(Q^*, \Omega) + \mathrm{diam}(Q)$ (parceque tout point de Q^* se trouve à moins de $\mathrm{diam}(Q)$ d'un point de Q, donc $\mathrm{dist}(Q, \partial\Omega) \leq 21\,\mathrm{diam}(Q)$. Bref,

$$10\,\mathrm{diam}(Q) \leq \mathrm{dist}(Q, \partial\Omega) \leq 21\,\mathrm{diam}(Q) \quad \text{pour tout } Q \in \mathcal{W}. \tag{2}$$

Notons que

$$\Omega = \bigcup_{Q \in \mathcal{W}} Q \tag{3}$$

parce que si $x \in \Omega$, tous les petits cubes dyadiques qui contiennent x vérifient (1); un tel cube est forcément contenu dans un cube maximal avec cette propriété, parce que les cubes qui contiennent x et ont un diamètre supérieur à $\mathrm{dist}(x, \partial\Omega)$ ne vérifient pas (1). L'union dans (3) est presque disjointe:

$$\text{les cubes } Q,\ Q \in \mathcal{W} \text{ ont des intérieurs disjoints.} \tag{4}$$

En effet, si deux cubes dyadiques ont des intérieurs non disjoints, on sait que l'un des deux est contenu dans l'autre; si de plus ce sont des cubes maximaux ayant une certaine propriété (ici, (1)), ils sont égaux.

Il sera aussi utile de savoir que les cubes $3Q$, $Q \in \mathcal{W}$, sont à recouvrement borné. On note $3Q$ le cube de même centre et de rayon 3 fois plus grand. Ce n'est pas un cube dyadique, mais ça n'est pas grave. On va montrer que si Q et R sont deux cubes de $\mathcal{W}$ tels que $3Q \cap 3R \neq \emptyset$, alors

$$\frac{1}{3}\,\mathrm{diam}(Q) \leq \mathrm{diam}(R) \leq 3\,\mathrm{diam}(Q) \tag{6}$$

et

$$\mathrm{dist}(Q, R) \leq \mathrm{diam}(Q) + \mathrm{diam}(R) \leq 4\,\mathrm{diam}(Q). \tag{7}$$

Pour la démonstration, le mieux est de faire un petit dessin. Tout point de $3Q$ est à distance au plus $\mathrm{diam}(Q)$ de Q. Si un tel point est aussi dans $3R$, il est à distance au plus $\mathrm{diam}(R)$ de R, et forcément $\mathrm{dist}(Q, R) \leq \mathrm{diam}(Q) + \mathrm{diam}(R)$. Il reste à prouver (6). Soit $x \in 3Q \cap 3R$. Alors

$$\begin{aligned} \mathrm{dist}(x, \partial\Omega) &\leq \mathrm{dist}(Q, \partial\Omega) + \mathrm{dist}(x, Q) + \mathrm{diam}(Q) \\ &\leq \mathrm{dist}(Q, \partial\Omega) + 2\,\mathrm{diam}(Q) \leq 23\,\mathrm{diam}(Q) \end{aligned} \tag{8}$$

par (2), et aussi

$$\mathrm{dist}(x, \partial\Omega) \geq \mathrm{dist}(R, \partial\Omega) - \mathrm{diam}(R) \geq 9\,\mathrm{diam}(R); \tag{9}$$

on en déduit que $\mathrm{diam}(R) \leq (23/9)\,\mathrm{diam}(Q) \leq 3\,\mathrm{diam}(Q)$; l'autre inégalité se démontre en échangeant Q et R.

La conséquence de (6) et (7) est que $3Q$ ne rencontre pas plus de C cubes $3R$, $R \in \mathcal{W}$.

12.b. La partition de l'unité associée

Gardons Ω et les cubes de Whithey $Q \in \mathcal{W}$ comme ci-dessus. Pour $Q \in \mathcal{W}$, choisissons une fonction $\widetilde{\varphi}_Q$ telle que

$$0 \leq \widetilde{\varphi}_Q(x) \leq 1 \text{ pour tout } x \in \mathbb{R}^n, \tag{10}$$

$$\widetilde{\varphi}_Q(x) = 1 \text{ pour } x \in Q \quad \text{et} \quad \widetilde{\varphi}_Q(x) = 0 \text{ pour } x \notin 2Q, \tag{11}$$

$$|\nabla^l \widetilde{\varphi}_Q(x)| \leq C_l \operatorname{diam}(Q)^{-l} \text{ pour tout } x \in \mathbb{R}^n. \tag{12}$$

[En général, on n'a pas besoin de toutes ces dérivées, et $l = 1, 2$ suffisent, mais de toute façon cela ne nous coûte rien de les demander.] Notons que $\widetilde{\varphi}_Q(x) = 0$ hors de Ω par (11) et car $3Q \subset \Omega$ à cause de (1). Par ailleurs,

$$1 \leq \sum_{Q \in \mathcal{W}} \widetilde{\varphi}_Q \leq C \quad \text{sur } \Omega \tag{13}$$

puisque les $3Q$ sont de recouvrement borné (et même chaque $3Q$ ne rencontre pas plus de C cubes $3R$, $R \in \mathcal{W}$). Maintenant on pose

$$\varphi_Q = \frac{\widetilde{\varphi}_Q}{\sum_{Q \in \mathcal{W}} \widetilde{\varphi}_Q}, \tag{14}$$

et les φ_Q, $Q \in \mathcal{W}$, sont la partition de l'unité souhaitée. Elles ont les propriétés simples suivantes:

$$\sum_{Q \in \mathcal{W}} \varphi_Q = \mathbf{1}_\Omega, \tag{15}$$

$$0 \leq \varphi_Q(x) \leq 1 \text{ pour } x \in \mathbb{R}^n \quad \text{et} \quad \varphi_Q(x) = 0 \text{ pour } x \notin 2Q, \tag{16}$$

$$|\nabla^l \varphi_Q(x)| \leq C_l \operatorname{diam}(Q)^{-l} \text{ pour } x \in \mathbb{R}^n \text{ et } l \geq 0. \tag{17}$$

Noter au passage que diam(Q) est équivalent à dist$(x, \partial\Omega)$ sur le support de φ_Q, par (2) (ou (8) et (9) pour Q).

12.c. Un résultat d'extension

C'est surtout pour voir comment on écrit une formule avec la partition de l'unité ci-dessus. On va noter, pour $\alpha \in]0,1]$, $\Lambda^\alpha(F)$ l'espaces des fonctions Höldériennes f définies sur F et telles que

$$|f(x) - f(y)| \leq C|x-y|^\alpha \quad \text{pour } x, y \in F. \tag{18}$$

On peut munir $\Lambda^\alpha(F)$ le la norme $|| \, ||_{\Lambda^\alpha}$, où $||f||_{\Lambda^\alpha}$ est la plus petite constante C telle qu'on ait (18).

Theorem (Whitney). *Soient $F \subset \mathbb{R}^n$ un ensemble fermé et $f \in \Lambda^\alpha(F)$. Alors f a une extension à $\mathbb{R}^n$ qui est également Höldérienne, avec $||f||_{\Lambda^\alpha(\mathbb{R}^n)} \leq C||f||_{\Lambda^\alpha(F)}$. Et de plus f est de classe C^∞ sur $\Omega = \mathbb{R}^n \setminus F$, et*

$$|\nabla^k f(x)| \leq C_k ||f||_{\Lambda^\alpha(F)} \operatorname{dist}(x, F)^{\alpha - k} \quad \textit{pour } x \in \Omega. \tag{19}$$

En plus, C_k ne dépend que de k et de la dimension n.

Noter que l'on pouvait sans crainte supposer F fermé, puisque en général toute fonction Höldérienne sur F a une unique extension à $\overline{F}$ (par uniforme continuité), qui est Höldérienne avec la même norme.

C'est juste un exemple, Whitney donne d'autres théorèmes d'extension plus généraux et un peu plus délicats [jets de Whitney et dérivées d'ordres plus élevés], mais le même genre d'idées marche.

Quand $\alpha = 1$ (pour les fonctions Lipschitziennes), on peut trouver une formule plus simple qui donne une extension lipschitzienne (mais pas C^∞ hors de F); voir l'exercice ci-dessous. Et pour des fonctions à valeurs dans $\mathbb{R}^n$ (eucliden), Kirszbraun a démontré qu'on pouvait trouver une extension Lipschitzienne avec la même norme.

Quand même, la méthode ci-dessous a l'avantage de marcher comme une machine, et (sauf erreur de ma part aisément corrigée par le lecteur) de s'appliquer immédiatement à f à valeurs dans un espace de Banach.

Pour démontrer le théorème, on va écrire une formule pour l'extension (qu'on note encore f). Evidemment, on garde les mêmes valeurs sur F, donc il s'agit de définir f sur Ω. On pose

$$f(x) = \sum_{Q \in \mathcal{W}} a_Q \, \varphi_Q(x), \tag{20}$$

où $\mathcal{W}$ et les φ_Q sont définis comme ci-dessus, à partir de $\Omega = \mathbb{R}^n \setminus F$, et où il s'agit juste de choisir les constantes a_Q à peu près convenablement.

On ne se foule pas trop. Pour chaque Q, on choisit $\xi_Q \in F$ tel que $\mathrm{dist}(x, Q) \leq 10\,\mathrm{dist}(F, Q)$, et on pose $a_Q = f(\xi_Q)$. [J'ai mis 10 pour insister sur le fait qu'on a un peu le choix.] Il reste à vérifier toutes les propriétés. On va supposer que $||f||_{\Lambda^\alpha(F)} = 1$ pour simplifier. Ça n'est évidemment pas grave.

D'abord, f est clairement C^∞ sur Ω. Ensuite, vérifions que

$$|f(x) - f(\xi)| \leq C|x - \xi|^\alpha \quad \text{pour } x \in \Omega \text{ et } \xi \in F. \tag{21}$$

On commence par le cas où $\mathrm{dist}(x, \xi) \leq 10d(x)$ où l'on note $d(x) = \mathrm{dist}(x, F)$. Soit $Q \in \mathcal{W}$ tel que $x \in 2Q$. Noter que

$$d(x) \geq \mathrm{dist}(Q, F) - \mathrm{diam}(Q) \geq \mathrm{dist}(Q, F)/2 \tag{22}$$

à cause de (2), donc $\mathrm{dist}(Q, F) \leq 2d(x)$ et $|x - \xi_Q| \leq 10\,\mathrm{dist}(Q, F) + \mathrm{diam}(Q) \leq 30d(x)$ (en comptant large), et

$$|a_Q - f(\xi)| = |f(\xi_Q) - f(\xi)| \leq C|\xi_Q - \xi|^\alpha \leq Cd(x)^\alpha.$$

Ceci vaut pour tout Q tel que $\varphi_Q(x) \neq 0$. On somme et on trouve que

$$\begin{aligned} |f(x) - f(\xi)| &= \Big| f(\xi) - \sum_Q a_Q \varphi_Q(x) \Big| = \Big| \sum_Q [f(\xi) - a_Q] \varphi_Q(x) \Big| \\ &\leq Cd(x)^\alpha \sum_Q |\varphi_Q(x)| = Cd(x)^\alpha \end{aligned}$$

puisque $\sum_Q \varphi_Q(x) = 1$ (deux fois). Donc (20) vaut quand $\mathrm{dist}(x, \xi) \leq 10d(x)$. Autrement, on passe par un $\xi_1 \in F$ tel que $\mathrm{dist}(x, \xi_1) \leq 10d(x)$ et on trouve que

$$|f(x) - f(\xi)| \leq |f(x) - f(\xi_1)| + |f(\xi_1) - f(\xi)| \leq Cd(x)^\alpha + C|\xi_1 - \xi|^\alpha \leq C|x - \xi|^\alpha$$

puisque justement $|x - \xi| \geq 10d(x)$. Il ne reste plus qu'à montrer que

$$|f(x) - f(y)| \leq C|x - y|^\alpha \quad \text{pour } x, y \in \Omega. \tag{23}$$

Quand $|x-y| \geq d(x)/10$, on passe par $\xi \in F$ tel que $\mathrm{dist}(x,\xi) \leq 10d(x)$, et on trouve

$$\begin{aligned} |f(x)-f(y)| &\leq |f(x)-f(\xi)| + |f(\xi)-f(y)| \leq C|x-\xi|^\alpha + C|\xi-y|^\alpha \\ &\leq C|x-\xi|^\alpha + C(|\xi-x|+|x-y|)^\alpha \leq C|x-y|^\alpha. \end{aligned}$$

Quand $|x-y| \leq d(x)/10$, il nous suffira de démontrer (19) pour $k=1$, puisqu'alors il viendra $|f(x)-f(y)| \leq C|x-y|d(x)^{\alpha-1} \leq C|x-y|^\alpha$. Donc il reste à démontrer (19).

On va dériver dans (20). Noter que comme plus haut, si Q est un cube tel que $2Q$ recontre un tout petit voisinage de x, on a encore $\mathrm{diam}(Q) \leq 2d(y)$ avec y très proche de x, comme en (22). Donc $\mathrm{diam}(Q) \leq 3d(y)$. Et alors $\xi_Q \in B(x, 100d(x))$. Noter aussi qu'il n'y a qu'un nombre fini de cubes Q qui interviennent, donc la dérivation sous le signe somme ne pose pas de problème. On écrit

$$\nabla^k f(x) = \sum_Q a_Q \nabla^k \varphi_Q(x), \tag{24}$$

puis on observe que $\sum_Q \nabla^k \varphi_Q(x) = 0$, puisque c'est une dérivée de la fonction constante égale à 1. On se donne R tel que $x \in R$, et on retire la constante nulle $a_R \sum_Q \nabla^k \varphi_Q(x)$ de (24) pour obtenir

$$\begin{aligned} |\nabla^k f(x)| &= \Big|\sum_Q [a_Q - a_R]\nabla^k \varphi_Q(x)\Big| \leq \sum_Q |a_Q - a_R||\nabla^k \varphi_Q(x)| \\ &\leq \sum_Q |a_Q - a_R| \ \mathrm{diam}(Q)^{-k} \leq \sum_Q |a_Q - a_R| \, d(x)^{-k}, \end{aligned} \tag{25}$$

par (17), et puisque $\mathrm{diam}(Q) \geq Cd(x)$ quand $2Q$ est proche de x. On sait aussi que $|a_Q - a_R| = |f(\xi_Q) - f(\xi_R)| \leq C|\xi_Q - \xi_R|^\alpha \leq Cd(x)^\alpha$, et que la somme a au plus C termes, donc (25) donne $|\nabla^k f(x)| \leq Cd(x)^{\alpha-k}$, comme annoncé dans (19). Ceci termine la démonstration du théorème. □

Exercice. Soient $F \subset \mathbb{R}^n$ et $f : F \to \mathbb{R}$ une fonction Lipschitzienne, qu'on va supposer de norme Lipschitzienne au plus 1. Montrer que f a une extension 1-Lipschitzienne définie sur $\mathbb{R}^n$. Indication: la plus grande solution possible, essayer la formule $f(y) = \inf\{f(x) + |x-y| \, ; \, x \in F\}$. Faire un dessin.

Noter que même pour les fonctions à valeurs dans $\mathbb{R}^n$, il y a une extension Lipschitzienne de même norme, mais là, c'est plus délicat.

Corollaire (Whitney). *Il existe une fonction* $f : [0,1]^2 \to \mathbb{R}$ *de classe* C^1, *avec la propriété que* $f(z_1) \neq f(z_2)$, *où* $z_1 = (1/4, 0)$ *et* $z_2 = (3/4, 0)$, *mais pourtant il existe une courbe* Γ *dans* $]0,1[^2$ *qui va de* z_1 *à* z_2, *et surlaquelle* $Df = 0$.

C'est possible parce qu'on n'a pas dit que Γ est de classe C^1, ni même de longueur finie. Il y a même mieux (mais plus délicat): Körner sait trouver $f : [0,1]^2 \to \mathbb{R}$, non constante, mais telle que pour a, $b \in]0,1[^2$, il existe un chemin $\gamma_{a,b}$ qui va de a à b, tel qu'en tout point intérieur de $\gamma_{a,b}$, f est différentiable avec une différentielle nulle.

La démonstration du corollaire est laissée en exerice, avec l'indication suivante: d'abord tracer Γ, un flocon de neige, puis une fonction f sur Γ (qui monte régulièrement, de sorte que $|f(x) - f(y)| \leq C|x-y|^\alpha$ pour $x, y \in \Gamma$, avec un $\alpha > 1$), et appliquer la construction de Whitney. Le point est que (21) a lieu avec un $\alpha > 1$ (ce qui est rare!), d'où la différentielle nulle. Il faut aussi un petit calcul basé sur (21) pour montrer que la différentielle est continue jusque sur Γ. Ou alors, plus brutalement, utiliser le second résultat d'extension ci-dessous avec $G = 0$ sur Γ. $\square$

12.d. Un second résultat d'extension

On veut généraliser un peu le résultat d'extension plus haut, en autorisant des dérivées d'ordre supéreur (ici juste d'ordre 1). D'une part c'est intéressant de voir comment on complète d'autre part on se servira de quelque chose de proche (dans le pararaphe 18 sur Lusin).

Theorem (Whitney aussi). *Soient* $F \subset \mathbb{R}^n$ *un ensemble fermé et* $f : F \to \mathbb{R}$, *à étendre. On suppose qu'en tout point* $\xi \in F$, *on dispose d'un vecteur* $G(\xi) \in \mathbb{R}^n$, *tel que*

$$|f(y) - f(\xi) - G(\xi) \cdot (y - \xi)| \leq |y - \xi| \varepsilon(|y - \xi|) \tag{26}$$

pour $y \in F$, *et où* $\varepsilon(r)$ *est une fonction croissante de* r *donnée, qui tend vers* 0 *quand* r *tend vers* 0. *Par exemple,* $\varepsilon(r) = Cr^\alpha$ *pour un* $\alpha \in]0,1]$ *donné. On suppose qu'il existe* $C \geq 0$ *telle que*

$$\varepsilon(\lambda r) \geq C^{-1} \lambda \varepsilon(r) \quad \text{pour } r > 0 \text{ et } 0 < \lambda \leq 1. \tag{26.5}$$

On suppose en outre que

$$|G(\xi) - G(\zeta)| \leq \varepsilon(|\xi - \zeta|) \tag{27}$$

pour $\xi, \zeta \in F$. Alors il existe une extension g de f à $\mathbb{R}^n$, telle que f est de classe C^1 sur $\mathbb{R}^n$ (et C^∞ sur $\mathbb{R}^n \setminus F$), avec

$$|Dg(x) - Dg(y)| \leq C\varepsilon(11|x-y|) \tag{28}$$

pour $x, y \in \mathbb{R}^n$, et aussi, pour $k \geq 2$ et $x \in \mathbb{R}^n \setminus F$,

$$|D^k g(x)| \leq C_k \operatorname{dist}(x,F)^{-k+1}\varepsilon(6\operatorname{dist}(x,F)). \tag{29}$$

Donc $G(\xi)$ joue le rôle d'un gradient.

J'ai ajouté la contrainte (26.5), sans quoi la démonstration ne marche pas. Elle n'interviendra que tout à fait à la fin, dans la fin de démonstration de (41) qui se trouve sous (50), quand à partir de l'estimation (29) de la dérivée à l'intérieur de $\mathbb{R}^n \setminus F$, on en déduit la partie (28) de l'estimation du module de continuité de Dg. Autrement, sauf erreur de ma part, je me suis débrouillé en mettant des second membres comme $\varepsilon(11|x-y|)$, pour ne pas en avoir besoin.

Ceci dit, c'est raisonnable de n'utiliser que des fonctions ε qui vérifient (26.5), donc qui ne se mettent pas à tendre vers 0 extrêmement vite. Ce serait comme parler des fonctions 2-Lipschitziennes sur $\mathbb{R}$, qui vérifieraient $|f(x) - f(y)| \leq C|x-y|^2$ pour $x, y \in \mathbb{R}$ tels que $|x-y| \leq 1$; on vérifie aisément que seules les fonctions constantes sont comme cela, en joignant deux points de $\mathbb{R}$ par des subdivisions très fines.

Bien sûr (26.5) est vérifié avec $C = 1$ pour $\varepsilon(r) = Cr^\alpha$ quand $0 < \alpha < 1$. L'idée est que $\varepsilon(r)$ ne tende pas vers 0 trop vite quand r tend vers 0 (sinon on aura du mal à assurer (28). Dans la pratique, et en supposant qu'on travaille aux échelles $\leq r_0$ (par exemple, si on veut appliquer le théorème sur un compact de diamètre $\leq r_0/2$), si notre fonction initiale ε tend vers 0 trop vite, on appliquera le résultat en utilisant une autre fonction plus régulière, par exemple

$$\widetilde{\varepsilon}(r) = \sup_{r \leq t \leq r_0} \frac{r}{t}\,\varepsilon(t), \tag{29.5}$$

qui est fabriquée pour vérifier (26.5) avec $C = 1$, et qui tend quand même vers 0 quand r tend vers 0 si notre fonction initiale ε tend vers 0.

A cause de (26.5), $\varepsilon(11r) \leq C\varepsilon(r)$ et on pourrait se dispenser des constantes 11 et 6, qui d'ailleurs ne sont pas optimales, quitte à augmenter encore un peu les constantes C et C_k dans (28) et (29).

Ici on a pris f à valeurs dans $\mathbb{R}$ pour simplifier, mais à nouveau, quitte à faire plus attention et à prendre $G(\xi)$ à valeurs dans les applications linéaires de $\mathbb{R}^n$ dans H, on pourrait prendre f à valeurs dans un Banach H.

De même, on se contente d'une dérivée, mais on pourrait en demander plus, à condition de supposer l'existence de développements limités d'ordre plus grand, avec une certaine cohérence (on appelle ça des jets de Whitney).

Le sujet n'est pas mort. On peut se demander comment trouver les jets juste en regardant les valeurs de f sur F. Travaux assez récents et spectaculaires de C. Fefferman et coauteurs. Il faut voir que par exemple, bien que le meilleur choix a posteriori de fonction G est de prendre le gradient de l'extension; on ne le connaît pas au départ, plusieurs choix peuvent correspondre à (26), et en construire un qui vérifie (27) peut demander un certain travail (dans le cas où il en existe, ce qui n'est pas une conséquence de (26)).

Donc on veut modifier la construction ci-dessus en tenant compte des $G(\xi)$, et la formule qui vient naturellement est

$$g(x) = \sum_{Q \in \mathcal{W}} \varphi_Q(x)[f(\xi_Q) + G(\xi_Q).(x - \xi_Q)] \tag{30}$$

pour $x \in \Omega$, où $\Omega = \mathbb{R}^n \setminus F$, $\mathcal{W}$ est toujours l'ensemble des cubes de Whitney de Ω, et on a associé à chaque $Q \in \mathcal{W}$ une fonction φ_Q comme plus haut et un point $\xi_Q \in F$ tel que $\operatorname{dist}(\xi_Q, Q) \leq 2 \operatorname{dist}(Q, F)$.

On garde $g(x) = f(x)$ sur F, bien entendu, pour avoir une extension.

On veut déjà vérifier que g est de clase C^1.

Pour $x \in \Omega$, on rappelle que dans un petit voisinage de x, la somme dans (30) ne concerne qu'un nombre fini de cubes Q (ceux pour lesquels $2Q$ rencontre le petit voisinage), donc la formule se différentie sous le signe somme, et

$$Dg(x) = \sum_{Q \in \mathcal{W}} D\varphi_Q(x)[f(\xi_Q) + G(\xi_Q).(x - \xi_Q)] + \sum_{Q \in \mathcal{W}} \varphi_Q(x) G(\xi_Q) \tag{31}$$

Et bien sûr Dg est continue sur l'ouvert Ω.

Avant de se lancer dans les calculs, notons qu'on aura souvent besoin de savoir les mêmes choses. Par exemple, si $x \in \Omega$, notons systématiquement

$$d(x) = \operatorname{dist}(x, F) \quad \text{et aussi } \mathcal{W}(x) = \big\{Q \in \mathcal{W}\,;\, x \in 3Q\big\}. \tag{32}$$

Donc $\varphi_Q(x) = 0$ quand $Q \in \mathcal{W} \setminus \mathcal{W}(x)$ (par (16)). Vérifions que

$$9 \operatorname{diam}(Q) \leq d(x) \leq 23 \operatorname{diam}(Q) \quad \text{pour} \quad Q \in \mathcal{W}(x). \tag{33}$$

D'abord, $d(x) = \mathrm{dist}(x, F) \leq \mathrm{dist}(Q, F) + 2\,\mathrm{diam}(Q)$ (prendre un point $z \in Q$ qui réalise la distance à F et noter que $|x - z| \leq 2\,\mathrm{diam}(Q)$). On déduit la seconde inégalité de (33) de la la maximalité de Q dans la définition de $\mathcal{W}$ (voir (2)). Aussi,

$$\mathrm{dist}(Q, F) \leq d(x) + \mathrm{diam}(Q) \tag{34}$$

(car $\mathrm{dist}(x, Q) \leq \mathrm{diam}(Q)$), et comme $\mathrm{dist}(Q, F) \geq 10\,\mathrm{diam}(Q)$ par définition (voir (1)), on trouve la première partie de (33). Ensuite, vérifions que

$$|\xi_Q - x| \leq 3d(x) \quad \text{pour} \quad Q \in \mathcal{W}(x). \tag{35}$$

Soit z un point de Q qui réalise la distance à ξ; alors

$$\begin{aligned} |\xi_Q - x| \leq |\xi_Q - z| + |z - x| &\leq |\xi_Q - z| + 2\,\mathrm{diam}(Q) = \mathrm{dist}(\xi_Q, Q) + 2\,\mathrm{diam}(Q) \\ &\leq 2\,\mathrm{dist}(Q, F) + 2\,\mathrm{diam}(Q) \leq 2d(x) + 4\,\mathrm{diam}(Q) \end{aligned} \tag{36}$$

par (34); on en déduit (35) à cause de (33).

Vérifions maintenant que g est différentiable en $y \in F$, avec $Dg(y) = G(y)$ (je confonds dérivée et gradient, mille excuses). Il s'agit de prouver que près de y, on a le développement limité

$$g(x) = g(y) + G(y).(x - y) + o(|x - y|) = f(y) + G(y).(x - y) + o(|x - y|), \tag{37}$$

avec en fait $o(|x - y|) \leq C|x - y|\varepsilon(6|x - y|)$ (où je viens de mettre 6 un peu au hazard en relisant, à cause du $6d(x)$ plus bas).

Pour $x \in F$, ce développement est valide parceque $g(x) = f(x)$, et par (26). Donc on peut supposer que $x \in \Omega$ et utiliser (30). Il vient (puisque $\sum_{Q \in \mathcal{W}(x)} \varphi_Q(x) = 1$)

$$\begin{aligned} g(x) - g(y) - &G(y) \cdot (x - y) = \\ &\sum_{Q \in \mathcal{W}(x)} \varphi_Q(x)[f(\xi_Q) + G(\xi_Q).(x - \xi_Q) - f(y) - G(y) \cdot (x - y)] \end{aligned}$$

Si on remplace $f(\xi_Q)$ par $f(y) + G(y) \cdot (\xi_Q - y)$ et $G(\xi_Q)$ par $G(y)$, l'expression entre crochet devient

$$f(y) + G(y) \cdot (\xi_Q - y) + G(y).(x - \xi_Q) - f(y) - G(y) \cdot (x - y) = 0.$$

Mais on introduit deux erreurs, qu'on va estimer facilement. La première est

$$
\begin{aligned}
E_1(Q) &= |f(\xi_Q) - f(y) - G(y)\cdot(\xi_Q - y)| \le |\xi_Q - y|\varepsilon(|\xi_Q - y|) \\
&\le (3d(x) + |x-y|)\,\varepsilon(3d(x) + |x-y|) \le 4|x-y|\,\varepsilon(4|x-y|)
\end{aligned}
\tag{38}
$$

par (26) et (35), et puisque $d(x) = \mathrm{dist}(x, F) \le |x-y|$ (car $y \in F$) et ε est une fonction croissante. La seconde est

$$
\begin{aligned}
E_2(Q) &= \big|[G(\xi_Q) - G(y)]\cdot(x - \xi_Q)\big| \le \varepsilon(|\xi_Q - y|)\,|x - \xi_Q| \\
&\le 3d(x)\,\varepsilon(3d(x) + |x-y|) \le 3|x-y|\,\varepsilon(4|x-y|),
\end{aligned}
\tag{39}
$$

cette fois par (27). On additionne tout ceci et on trouve

$$
\begin{aligned}
|g(x) - g(y) - G(y)\cdot(x-y)| &\le \sum_{Q\in\mathcal{W}(x)} \varphi_Q(x)[E_1(Q) + E_2(Q)] \\
&\le 7|x-y|\,\varepsilon(4|x-y|) \sum_{Q\in\mathcal{W}(x)} \varphi_Q(x) = 7|x-y|\,\varepsilon(4|x-y|).
\end{aligned}
\tag{40}
$$

On en déduit bien (37) et la différentiabilité de g aux points de F (donc finalement partout, puisqu'on a (31)).

Maintenant on sait que Df existe, et est donnée soit par $G(y)$ (sur F), soit par la formule (31). On doit vérifier que Df est continue, avec le bon module de continuité.

Donc on veut montrer que

$$
|Df(x) - Df(y)| \le C\varepsilon(11|x-y|).
\tag{41}
$$

Quand $x, y \in F$, c'est juste (27) (et le coefficient 11 ne gêne pas puisque ε est croissante), car $Df = G$ sur F.

Prenons maintenant $y \in F$ et $x \in \Omega$. Posons $A(x) = \sum_{Q\in\mathcal{W}} \varphi_Q(x) G(\xi_Q)$. Alors

$$
\begin{aligned}
|G(y) - A(x)| &\le \sum_{Q\in\mathcal{W}} \varphi_Q(x)|G(\xi_Q) - G(y)| \\
&\le \sum_{Q\in\mathcal{W}(x)} \varphi_Q(x)\varepsilon(|\xi_Q - y|) \le 4\varepsilon(4|x-y|)
\end{aligned}
\tag{42}
$$

en utilisant encore (27), puis le fait que $|\xi_Q - y| \leq |\xi_Q - y| + |x-y| \leq 3d(x) + |x-y| \leq 4|x-y|$ par (35) et comme plus haut. Il nous reste à évaluer

$$B(x) = \sum_{Q\in\mathcal{W}} D\varphi_Q(x)[f(\xi_Q) + G(\xi_Q).(x-\xi_Q)] \tag{43}$$

Soit p l'un des ξ_Q tel que $x \in 2Q$ (on aurait pu aussi prendre n'importe quel des points de F tels que $|p-x| \leq 2\,\mathrm{dist}(x,F)$). Otons $f(p) + G(p).(x-p)$ de chaque terme; puisque $\sum_{Q\in\mathcal{W}} D\varphi_Q(x) = 0$, il vient

$$B(x) = \sum_{Q\in\mathcal{W}} D\varphi_Q(x)[f(\xi_Q) + G(\xi_Q).(x-\xi_Q) - f(p) - G(p).(x-p)] \tag{44}$$

A nouveau, si on remplace $f(\xi_Q)$ par $f(p) + G(p)\cdot(\xi_Q - p)$, et aussi $G(\xi_Q)$ par $G(p)$, la quantité entre crochets devient

$$f(p) + G(p)\cdot(\xi_Q - p) + G(p).(x-\xi_Q) - f(p) - G(p).(x-p) = 0$$

Et les deux erreurs commises sont

(45)

$$E_1'(Q) = |f(\xi_Q) - f(p) - G(p)\cdot(\xi_Q - p)| \leq |\xi_Q - p|\varepsilon(|\xi_Q - p|) \leq 6d(x)\varepsilon(6d(x))$$

puisque $|\xi_Q - p| \leq |\xi_Q - x| + |p-x| \leq 6d(x)$ par (35), et aussi

$$E_2'(Q) = |[G(\xi_Q) - G(p)]\cdot(x-\xi_Q)| \leq \varepsilon(|\xi_Q - p|)\,|x-\xi_Q| \leq 3d(x)\varepsilon(6d(x)). \tag{46}$$

On somme tout ceci sur $Q \in \mathcal{W}(x)$ et on trouve

$$\begin{aligned} |B(x)| &\leq \sum_{Q\in\mathcal{W}(x)} |D\varphi_Q(x)|[E_1'(Q) + E_2'(Q)] \\ &\leq C \sum_{Q\in\mathcal{W}(x)} \mathrm{diam}(Q)^{-1}[d(x)\varepsilon(6d(x))] \leq C\varepsilon(6d(x)), \end{aligned} \tag{47}$$

où l'on utilise maintenant l'estimation (17) sur la dérivée de φ_Q, le fait que $\mathrm{diam}(Q) \sim d(x)$ par (33), et le fait que $\mathcal{W}(x)$ a au plus C éléments. Comme $Df(x) - Df(y) = A(x) + B(x) - G(y)$, on déduit (41) de (42) et (47), dans notre cas où $y \in F$ et $x \in \Omega$, et avec la meilleure borne $C\varepsilon(4|x-y|)$ (ajouté en 2017: ou est-ce $\varepsilon(6|x-y|)$?).

Il reste pour finir à vérifier (41) quand $x, y \in \Omega$, et aussi à montrer (29). Pour cela, on commence par (29) et donc on dérive (30) k fois.

Le premier terme est qu'on va regarder est celui où toutes les dérivées portent sur φ_Q, c.-à-d.,

$$D_1(x) = \sum_{Q \in \mathcal{W}} D^k \varphi_Q(x)[f(\xi_Q) + G(\xi_Q).(x - \xi_Q)] \tag{48}$$

C'est pareil que $B(x)$ (voir (43)), sauf que $D\varphi_Q(x)$ est remplacé par $D^k \varphi_Q(x)$. On fait les mêmes calculs, sauf qu'au lieu de majorer $|D\varphi_Q(x)|$ par $\mathrm{diam}(Q)^{-1} \sim d(x)^{-1}$, on majore $|D^k \varphi_Q(x)|$ par $\mathrm{diam}(Q)^{-k} \sim d(x)^{-k}$. On trouve $|D_1(x)| \leq C_k d(x)^{-k+1} \varepsilon(6d(x))$, commme souhaité pour (29).

Les termes restants sont ceux où on a dérivé $[f(\xi_Q) + G(\xi_Q).(x - \xi_Q)]$ une fois (c'est le maximum, puisqu'elle est affine). On trouve

$$D_2(x) = \sum_{Q \in \mathcal{W}} D^{k-1} \varphi_Q(x) G(\xi_Q) = \sum_{Q \in \mathcal{W}} D^{k-1} \varphi_Q(x)[G(\xi_Q) - G(p)], \tag{49}$$

où p est l'un des ξ_Q, $Q \in \mathcal{W}(x)$, choisi au hasard, et où l'on a pu faire la soustraction parce que $k \geq 2$ et donc $\sum_{Q \in \mathcal{W}} D^{k-1} \varphi_Q(x) = 0$. On note que $|G(\xi_Q) - G(p)| \leq \varepsilon(|\xi_Q - p|) \leq \varepsilon(6d(x))$, et quand on somme on trouve encore $|D_2(x)| \leq C_k d(x)^{-k+1} \varepsilon(6d(x))$. Donc on a bien (29).

Finalement, passons à la démonstration de (41) pour $x, y \in \Omega$. Si $|x-y| \leq 4d(x)/5$, on intègre (29) (avec $k = 2$) entre x et y. Comme pour $z \in [x, y]$, on a que $d(x)/5 \leq d(z) \leq 9d(x)/5$ et donc

$$D^2 g(z) \leq C d(z)^{-1} \varepsilon(6d(z)) \leq C d(x)^{-1} \varepsilon(11d(x)) \tag{50}$$

quand on intègre on trouve que

$$|Dg(x) - Dg(y)| \leq C|x - y| d(x)^{-1} \varepsilon(11d(x)). \tag{51}$$

Maintenant on se sert enfin de (26.5), comme on l'avait annoncé, pour faire l'estimation restante la plus triviale a priori. On applique (26.5) avec $r = 11d(x)$ et $\lambda = \frac{|x-y|}{d(x)} \leq 1$, dit que

$$\varepsilon(11d(x)) = \varepsilon(r) \leq C\lambda^{-1} \varepsilon(\lambda r) = C \frac{d(x)}{|x - y|} \varepsilon(11|x - y|). \tag{51.5}$$

Du coup, (51) implique bien (41).

Sinon, si $|x-y| > 4d(x)/5$, on passe par un $z \in F$, à distance minimale de x, et on utilise le cas de (41) qu'on connaît déjà (où l'un des deux points est dans F), et on obtient que

(52)
$$|Dg(x)-Dg(y)| \le |Dg(x)-Dg(z)|+|Dg(z)-Dg(y)| \le C\varepsilon(4|z-x|)+C\varepsilon(4|z-y|)$$

Mais $|z-y| \le |z-x| + |x-y| \le d(x) + |x-y| \le \frac{5|x-y|}{4} + |x-y| = \frac{9|x-y|}{4}$, donc $4|z-y| \le 9|x-y|$, et aussi $4|z-x| \le 9|x-y|$ plus simplement, de sorte que (52) donne bien (41). □

12.e. Encore un résultat d'extension, le théorème de Kirszbraun

Théorème (Kirszbraun). *Soient $F \subset \mathbb{R}^n$ un sous-ensemble et $f : F \to H$ une fonction 1-Lipschitzienne. Alors f a une extension 1-Lipschitzienne à $\mathbb{R}^n$ tout entier.*

On va prendre la démonstration de Federer (2.10.43 page 201). Il donne aussi un contre-exemple quand la norme de l'espace d'arrivée n'est pas hilbertienne.

Ici, 1-Lipschitzienne signifie seulement que $|f(x) - f(y)| \le |x-y|$.

On pourrait supposer F fermé puisque de toute façon on pourrait commencer par étendre f à $\overline{F}$, sans rien changer.

En fait, il suffira de savoir, étant donné un point supplémentaire x_0 hors de F, étendre f à $F \cup \{x_0\}$ (toujours en restant 1-Lipschitzien). Disons pourquoi.

Si comme Federer on aime l'axiome du choix, on met sur l'ensemble des couples (A, g), où $A \subset \mathbb{R}^n$ contient F et $g : A \to H$ est une extension de f, l'ordre qu'on imagine: $(A, g) < (A', g')$ si $A \subset A'$ et g' est une extension 1-Lipschitzienne de g. Puis on choisit un élément maximal parmi ces paires, qui existe parce que l'ordre est inductif. C'est à dire, si l'on a une famille de couples qui est totalement ordonnée, alors elle a une borne supérieure, obtenue simplement en prenant l'union des ensembles, et la fonction sur cette union qui coïncide avec g sur A pour toute paire (A, g) de la famille. Noter que cette fonction est 1-lipschitzienne. Si pour cet élément maximal, A n'est pas $\mathbb{R}^n$, utiliser ce qu'on vient de dire que l'on sait faire pour ajouter un point x_0 et obtenir une contradiction.

Si on veut éviter Zorn, on se donne une suite de points à ajouter, de manière que la suite soit dense dans $\mathbb{R}^n \setminus F$, et on définit des extensions successives, toutes 1-Lipschitziennes. Et on vérifie qu'il existe une unique $g : \mathbb{R}^n \to H$ qui les prolonge toutes.

Donc on se donne $x_0 \in \mathbb{R}^n \backslash F$ et on veut définir $z = f(x_0)$. En translatant tout, on peut supposer que $x_0 = 0$.

La contrainte est que $|z - f(x)| \leq |x|$ pour tout $x \in F$, et on peut même se contenter de le vérifier pour $x \in F'$, une partie au plus dénombrable dense. Donc on veut juste montrer que

$$\bigcap_{x \in F'} \overline{B}(f(x), |x|) \neq \emptyset. \tag{53}$$

Le membre de gauche est l'intersection décroissante des ensembles $\bigcap_{x \in F_m} \overline{B}(f(x), |x|)$, où F_m est l'ensmble fini des m premiers éléments d'une énumération de F'. Puisque chacun des ensembles en question est compact, l'intersection est non vide si

$$\bigcap_{x \in G} \overline{B}(f(x), |x|) \neq \emptyset \tag{54}$$

pour tout G fini contenu dans F. On fixe donc G fini, et on va montrer (54) pour conclure. Pour tout $t > 0$, posons

$$Y_t = \bigcap_{x \in G} \overline{B}(f(x), t|x|) \tag{55}$$

Pour t assez grand, $0 \in Y_t$ donc Y_t n'est pas vide (se souvenir que $0 \notin F$, donc $x \neq 0$ pour $x \in G$). Soit $s = \inf\{t > 0\,;\, Y_t \neq \emptyset\}$. Noter d'abord, encore par compacité, que $Y_s \neq \emptyset$. Il suffit donc de montrer que $s \leq 1$ pour conclure.

Grâce au statut de minimum de s, l'ensemble Y_s a des propriétés particulières qui vont permettre d'en savoir plus.

Vérifions d'abord que Y_s est réduit à un point. En effet, si Y_s contient deux points distincts z_1 et z_2, $z = (z_1 + z_2)/2$ fait strictement mieux, au sens où pour chaque $x \in G$, par la stricte convexité de la boule $\overline{B}(f(x), s|x|)$ qui contient z_1 et z_2, le point z est dans une boule strictement plus petite. Et alors z est dans un Y_t avec $t < s$ (une contradiction).

Quitte à additionner une constante à f (donc à tous les $f(x)$, $x \in G$, et donc à Y_x aussi), on peut supposer que $Y_s = \{0\}$.

Soit G_0 l'ensemble des points de G tels que $|f(z)| = s|x|$. Noter que G_0 n'est pas vide, car sinon 0 serait dans un Y_t avec $t < s$. Si l'on pouvait séparer 0 de l'ensemble des $f(x)$, $x \in G_0$, par un hyperplan P, on pourrait remplacer 0 par un meilleur point, situé sur le segment qui va de 0 à sa projection sur P. En effet, cela diminuerait strictement les distances à chaque point de $f(G_0)$, et pour les autres ça n'est pas grave si on bouge peu l'origine. Donc 0 est dans l'enveloppe convexe des $f(x)$, $x \in G_0$.

En bref, $0 = \sum_{x \in G_0} \lambda_x f(x)$, avec des $\lambda_x \geq 0$ tels que $\sum \lambda_x = 1$. Et $|f(x)| = s|x|$ puisque $x \in G_0$. Maintenant on va utiliser l'identité $2u \cdot v = |u|^2 + |v|^2 - |u - v|^2$:

$$\begin{aligned} 0 = 2\Big|\sum_{x \in G_0} \lambda_x f(x)\Big|^2 &= 2 \sum_x \sum_y \lambda_x f(x) \cdot \lambda_y f(y) \\ &= \sum_x \sum_y \lambda_x \lambda_y \big[|f(x)|^2 + |f(y)|^2 - |f(x) - f(y)|^2\big] \end{aligned}$$

Maintenant on dit que $|f(x)| = s|x|$ et que $|f(x) - f(y)| \leq |x - y|$ et on trouve que

$$\begin{aligned} 0 &\geq \sum_x \sum_y \lambda_x \lambda_y \big[|sx|^2 + |sy|^2 - |x - y|^2\big] \\ &= s^2 \sum_x \sum_y \lambda_x \lambda_y \big[|x|^2 + |y|^2 - |x - y|^2\big] - (1 - s^2) \sum_x \sum_y \lambda_x \lambda_y |x - y|^2 \\ &= 2s^2 \Big|\sum_{x \in G_0} \lambda_x x\Big|^2 - (1 - s^2) \sum_x \sum_y \lambda_x \lambda_y |x - y|^2 \end{aligned}$$

(en revenant par le même calcul). Donc

$$2s^2 \Big|\sum_{x \in G_0} \lambda_x x\Big|^2 \leq (1 - s^2) \sum_x \sum_y \lambda_x \lambda_y |x - y|^2,$$

et on veut une contradiction quand $s > 1$ (ce qui tombe bien car le membre de droite a alors l'air négatif).

Si on peut trouver x, y distincts tels que λ_x et λ_y sont non nuls, alors forcément $1 - s^2 \geq 0$, $s \leq 1$, et on a gagné. Sinon, seul un coefficient est non nul, 0 est l'un des $f(x)$, et comme $|f(x)| = s|x|$ il vient $s = 0$ (puisque $0 \notin G$). C'est encore mieux. $\square$

13. DIFFERENTIABILITE PRESQUE-PARTOUT

On va démontrer le théorème de Rademacher-Calderón:

Theorème. *Soit $f \in W^{1,p}(V)$, où V est un domaine de $\mathbb{R}^n$ et $p > n$. Alors f est différentiable en x pour presque tout $x \in V$, et la différentielle est donnée par la dérivée au sens des distributions (elle aussi définie presque partout).*

Quelques remarques avant de faire la démonstration. Le cas des fonctions Lipschitzennes est le théorème de Rademacher; je suppose que la démonstration ci-dessous est due à Calderón.

C'est amusant que les hypothèses ci-dessus, pour obtenir la différentiabilité presque-partout, sont les mêmes que plus haut, pour dire que f est seulement (égale p.p. à une fonction) localement Höldérienne. Ceci prouve surtout, à mon avis, que quand on applique les injections de Sobolev, on perd souvent beaucoup d'information.

Le résultat est local; on pourra donc supposer que V est une petite boule, ou un petit cube ouvert, si ça nous aide.

Dans le cas Lipschitzien, on va d'abord vérifier que f est Λ-Lipschitzienne (sur un domaine V convexe) si et seulement si f est dans $W^{1,1}_{loc}(V)$, avec $|\nabla f| \leq \Lambda$ presque-partout. De sorte que la définition donnée plus haut de $W^{1,+\infty}$ donne bien les fonctions Lipschitziennes dans les bons cas. Et que le théorème ci-dessus s'applique aux fonctions lipschitziennes.

D'abord on suppose que $|f(x) - f(y)| \leq \Lambda |x - y|$ dans un domaine V, et on doit montrer que la dérivée de f, au sens des distributions, est une fonction bornée de norme $||\nabla f||_\infty \leq \Lambda$. Et il est facile à voir qu'il suffit de montrer ceci dans une boule B telle que $2B \subset V$.

On approxime f (localement) par une suite $\{f_k\}$, obtenue par convolution (c'est pour ceci qu'on suppose que $2B \subset V$, de sorte que f_k est définie dans B pour k assez grand, et $\{f_k\}$ converge vers f uniformément dans B (rappelons que f est Lipschitzienne dans $2B$). Les f_k sont elles aussi Λ-Lipschitziennes sur B (par calcul direct, ou en observant qu'une moyenne de fonctions Λ-Lipschitziennes est Λ-Lipschitzienne), et comme elles sont de classe C^1, un calcul brutal de la dérivée $\nabla f_k(x) \cdot v$ comme limite de $[f_k(x + tv) - f_k(x)]/t$ montre que $|\nabla f_k(x) \cdot v| \leq \Lambda |v|$ pour tout $v \in \mathbb{R}^n$, puis que $|\nabla f_k| \leq \Lambda$ partout.

On utilise maintenant le fait que L^∞, qui est le dual de L^1, a une boule unité faiblement compacte, où faiblement fait référence à la dualité avec L^1. On peut extraire une sous-suite pour laquelle les fonctions ∇f_k convergent faiblement (au sens de la dualité avec les fonctions de L^1), vers une limite g. Cette limite g définit une forme linéaire continue sur L^1, de norme au plus Λ aussi, donc c'est une fonction bornée de norme plus Λ. En plus, la démonstration du lemme 2 du paragraphe 6 (sur $W^{1,p}$) dit que $f \in W^{1,+\infty}(B)$, avec $\nabla f = g$ dans B (rappelons que f_k tend vers f uniformément dans B). Donc $f \in W^{1,+\infty}(B)$, avec $|\nabla f|(x) \leq \Lambda$ p.p. Et donc ceci reste vrai dans V, par un argument simple de localisation.

Réciproquement, vérifions que si $f \in W^{1,1}_{loc}(U)$ avec une dérivée (au sens des distributions) ∇f telle que $|\nabla f| \leq \Lambda$ presque-partout, alors f est (égale presque-partout à une fonction) Λ-Lipschitzienne dans toute boule B contenue dans V.

On peut se contenter de vérifier ceci pour toute boule $B' \subset B$ strictement plus petite (mais avec le même Λ), car après il suffira de prendre une union.

On définit f_k comme ci-dessus (par convolution avec des φ_k, et donc f_k est définie sur B'), et on note que puisque $\nabla f_k = \nabla f * \varphi_k$ (un calcul fait pour la démonstration du lemme 2 dans le paragraphe 6 sur $W^{1,p}$), $|\nabla f_k| \leq \Lambda$ (dans B' et pour k assez grand). Alors f_k est Λ-Lipschitzienne (appliquer le théorème des accroisements finis sur le segment $[x,y]$. Comme $\{f_k\}$ converge vers f dans $L^1(B')$ (par convolution et notre hypothèse que $f \in W^{1,1}(B)$), on peut extraire une sous-suite qui converge aussi presque-partout, et alors f est égale presque-partout sur B' (là où la suite converge) à une fonction localement Λ-Lipschitzienne.

Revenons au théorème, maintenant que nous savons qu'il s'applique aussi aux fonctions Lipschitziennes. Notons que sous nos hypothèses, Poincaré dit que f est (égale preque partout à une fonction) continue. Notons aussi que par Hölder, le résultat est sensé être d'autant plus difficile que p est petit. En particulier, on va maintenant pouvoir supposer que $p < +\infty$ sans perte de généralité.

Le point central sera la démonstration d'une inégalité maximale, qui donnera le résultat par passage à la limite à partir d'une classe dense.

Pour simplifier les calculs (et sans perdre de généralité), on va supposer que V contient la boule $B(0,3)$, et on va prouver la différentiabilité dans $B(0,1)$.

On se donne x dans $B(0,1)$ et $y \in B(0,2)$; donc le cône $C(x,y)$ est contenu dans $B(0,2) \subset V$, et on utilise la proposition 1 du chapitre 6 (sur Poincaré). On trouve que

$$|f(x) - f(y)| \leq C\,|x-y|^{\frac{p-n}{p}} \left\{ \int_{C(x,y)} |\nabla f|^p \right\}^{1/p}. \tag{1}$$

En supposant maintenant que $|x-y|<1$, et en posant $r=|x-y|$, il vient

$$\frac{|f(x)-f(y)|}{|x-y|} \leq C r^{-\frac{n}{p}} \Big\{ \int_{B(x,r)} |\nabla f|^p \Big\}^{1/p} \leq C M_p(|\nabla f|)(x) \tag{2}$$

où l'on a posé

$$M_p(g)(x) = \sup_{r\in]0,1]} \Big\{ \frac{1}{|B(x,r)|} \int_{B(x,r)} |g|^p \Big\}^{1/p} = M(|g|^p)^{1/p}(x) \tag{3}$$

pour g mesurable positive, et où $M=M_1$ est la fonction maximale centrée de Hardy-Littlewood, où l'on se restreint aux rayons ≤ 1.

Noter que le second membre dépend de x, mais plus de $y \in B(x,1)$. Comme on va avoir besoin d'estimer des normes dans L^p, on prend une sécurité supplémentaire. On choisit un p_1 tel que $n<p_1<p$, et on applique (2) avec l'exposant p_1. On trouve ainsi que

$$S_f(x) =: \sup_{y\in B(x,1)} \frac{|f(x)-f(y)|}{|x-y|} \leq C M_{p_1}(|\nabla f|)(x) \tag{4}$$

pour $x\in B(0,1)$. C'est bien, parce que $f\in W^{1,p}(B(0,2))$, donc $g=|\nabla f| \in L^p(B(0,2)$, donc $g^{p_1}\in L^{p/p_1}$, et (puisque maintenant $p/p_1>1$) le théorème maximal dit que $M(g^{p_1})\in L^{p/p_1}$ également. De sorte que

$$\begin{aligned} ||M_{p_1}(|\nabla f|)||_p^p &= \int M_{p_1}(g)^p = \int M(g^{p_1})^{p/p_1} = ||M(g^{p_1})||_{p/p_1}^{p/p_1} \\ &\leq C\, ||g^{p_1}||_{p/p_1}^{p/p_1} = ||g||_p^p = ||\nabla f||_p^p. \end{aligned} \tag{5}$$

Donc on a montré que $S_f\in L^p(B(0,1))$, avec

$$||S_f||_{L^p(B(0,1))} \leq C||M_{p_1}(|\nabla f|)||_{L^p(B(0,2))} \leq C||\nabla f||_{L^p(B(0,2))}.$$

Maintenant, on pose

$$S'_f(x) = \sup_{y\in B(x,1)} \frac{1}{|x-y|}\, |f(x)-f(y)-\nabla f(x)\cdot(x-y)| \tag{6}$$

et bien sûr $||S'_f||_{L^p(B(0,1))} \leq C||\nabla f||_p$, puisque $S'_f(x)\leq S_f(x)+|\nabla f(x)|$.

Maintenant, on est prêt à démontrer que si $f \in W^{1,p}(B(0,3))$, alors f est différentiable en x pour presque tout $x \in B(0,1)$. On pose

$$(7) \qquad L_f(x) = \lim_{r\to 0}\Big\{ \sup_{y\in B(x,r)} \frac{1}{|x-y|}\,|f(x)-f(y)-\nabla f(x)\cdot(x-y)|\Big\},$$

on constate que $L_f(x)=0$ partout lorsque f est de classe C^1. Soit maintenant $f \in W^{1,p}(B(0,3))$ et $\{f_k\}$ est une suite de fonctions de classe C^1 qui convergent vers f dans $W^{1,p}(B(0,2))$ (ce qui est facile à obtenir, par convolution). Alors, pour $x \in B(0,1)$ $L_f(x) \le L_{f_k}(x) + S'_{f-f_k}(x) = S'_{f-f_k}(x)$, donc

$$||L_f||_{L^p(B(0,1))} \le ||S'_{f-f_k}||_{L^p(B(0,1))} \le C||f-f_k||_{L^p(B(0,2))},$$

qui est aussi petit qu'on veut, donc $L_f(x)=0$ presque-partout sur $B(0,1)$, ce qui signifie que f est différentiable en x presque partout, comme annoncé. □

Pour conclure ce chapitre, donnons une variante de l'inégalité (2) ou (4), qui vaut pour $f \in W^{1,p}$ mais cette fois pour tout $p \ge 1$.

Proposition. *Soient $f \in W^{1,1}(B)$, où B est une boule de $\mathbb{R}^n$. Alors pour presque-tout $x \in B$ et presque-tout $y \in B$,*

$$(8) \qquad |f(x)-f(y)| \le C|x-y|\,\Big\{M\big(\mathbf{1}_B|\nabla f|\big)(x) + M\big(\mathbf{1}_B|\nabla f|\big)(y)\Big\},$$

où M est encore la fonction maximale centrée de Hardy-Littlewood.

Il semble que ceci ne soit pas trop ancien, et dû à Edberg ou Bojarski (et sans doute retouvé par de nombreux autres).

Bien sûr comme d'habitude le presque-partout dépend aussi du choix de représentant de f qu'on aura fait.

On peut même remplacer $\mathbf{1}_B$ dans (8) par $\mathbf{1}_{B_0}$, où B_0 est n'importe quelle boule contenue dans B qui contient x et y. C'est à la fois une conséquence de l'énoncé initial (parce que $f \in W^{1,1}(B_0)$ si $f \in W^{1,1}(B)$, et avec une norme plus petite) et de la démonstration.

Démontrons la proposition. On va en fait montrer que (8) a lieu dès que x et y sont des points de Lebesgue pour ∇f. Et pour comparer $f(x)$ et $f(y)$, on passera par la moyenne de f sur diverses boules B_k.

Commençons par choisir trois boules B_0, $B_{0,x}$, et $B_{0,y}$ contenues dans B et ayant les propriétés suivantes: $x \in B_{0,x}$, $y \in B_{0,y}$, $B_{0,x}$ et $B_{0,y}$ sont de rayon $r_0 = |x-y|/2$, le rayon de B_0 est compris entre r_0 et $10r_0$, et enfin

$$(9) \qquad |B_{0,x}\cap B_0| \ge cr_0^n \;\text{ et }\; |B_{0,y}\cap B_0| \ge cr_0^n$$

pour une constante $c > 0$ qu'on n'aura pas besoin de calculer. Pour l'existence de ces boules, il est recommandé de faire un petit dessin, et de corriger les constantes si je me suis encore trompé.

Ensuite, pour $k \geq 1$, on choisit des boules $B_{k,x}$ et $B_{k,y}$, $k \geq 1$, de rayon $r_k = 2^{-k} r_0$, telles que $x \in B_{k+1,x} \subset B_{k,x}$ et $y \in B_{k+1,y} \subset B_{k,y}$ pour $k \geq 0$. Très faciles à choisir par récurrence.

On s'intéresse aux moyennes de f, par exemple sur les $B_{k,x}$. Posons $m_k = \oint_{B_{k,x}} f(y)dy$, où on va noter $\oint$ les moyennes (je n'ai pas trouvé le symbole pour les intégrales barrées). Alors

(10)

$$\begin{aligned} |m_{k+1} - m_k| = \Big| \oint_{B_{k+1,x}} f - m_k \Big| &\leq \oint_{B_{k+1,x}} |f - m_k| \leq 2^n \oint_{B_{k,x}} |f - m_k| \\ &\leq 2^n C r_n \oint_{B_{k,x}} |\nabla f| \leq 4^n C r_n \oint_{B(x,2r_k)} \mathbf{1}_B | \nabla f| \\ &\leq 4^n C r_n M\big(\mathbf{1}_B |\nabla f|\big)(x) \end{aligned}$$

juste parce que $B_{k+1,x} \subset B_{k,x}$, que $|B_{k+1,x}| = 2^{-n} |B_{k,x}|$, puis par l'inégalité de Poincaré avec les exposants 1, et enfin en appliquant la définition de la fonction maximale de Hardy-Littlewood. On additionne tout ceci en se souvenant que $r_k = 2^{-k} r_0 = 2^{-k-1}|x-y|$, et on trouve que

$$|m_k - m_0| \leq C'|x-y| M\big(\mathbf{1}_B |\nabla f|\big)(x). \tag{11}$$

Evidemment, les mêmes estimations sur les $B_{k,y}$ donnent

$$\Big| \oint_{B_{k,y}} f - \oint_{B_{0,y}} f \Big| \leq C'|x-y| M\big(\mathbf{1}_B |\nabla f|\big)(y). \tag{12}$$

Pour finir, on va comparer m_0 et $\oint_{B_{0,y}} f$ à $m = \oint_{B_0} f$. Par exemple,

$$\begin{aligned} \int_{B_{0,x} \cap B_0} |f(z) - m| &\leq \int_{B_0} |f(z) - m| \leq C r_0 \int_{B_0} |\nabla f| \\ &\leq C r_0 \int_{B(x,20r_0)} \mathbf{1}_B \, |\nabla f| \leq C r_0^{n+1} M\big(\mathbf{1}_B |\nabla f|\big)(x) \end{aligned} \tag{13}$$

par Poincaré et comme en (10), et de même

$$\begin{aligned} \int_{B_{0,x} \cap B_0} |f(z) - m_0| &\leq \int_{B_{0,x}} |f(z) - m_0| \leq C r_0 \int_{B_{0,x}} |\nabla f| \\ &\leq C r_0 \int_{B(x,2r_0)} \mathbf{1}_B \, |\nabla f| \leq C r_0^{n+1} M\big(\mathbf{1}_B |\nabla f|\big)(x) \end{aligned} \tag{14}$$

puis en additionnant
(15)

$$|m-m_0| = |B_{0,x}\cap B_0|^{-1}\int_{B_{0,x}\cap B_0}|m-m_0| \leq Cr_0^{-n}\int_{B_{0,x}\cap B_0}|m-m_0|$$

$$\leq Cr_0^{-n}\int_{B_{0,x}\cap B_0}|f(z)-m|+|f(z)-m_0| \leq Cr_0 M\big(\mathbf{1}_B|\nabla f|\big)(x)$$

à cause de (9). On a une estimation semblable de $m - \fint_{B_{0,y}} f$, on additionne tout, et on trouve

$$|\fint_{B_{k,x}} f - \fint_{B_{k,y}} f| \leq C|x-y|\Big\{M\big(\mathbf{1}_B|\nabla f|\big)(x) + M\big(\mathbf{1}_B|\nabla f|\big)(y)\Big\} \tag{16}$$

grâce à (11) et (12) en particulier.

Et maintenant $f(x) = \lim_{k\to+\infty} m_k$, car

$$|f(x)-m_k| = |\fint_{B_{k,x}} f(x)-f(y)dy| \leq \fint_{B_{k,x}}|f(x)-f(y)| \leq 2^n\fint_{B(x,2r_k)}|f(x)-f(y)|$$

puisque $x\in B_k$, donc $B_k\subset B(x,2r_k)$, et le membre de droite tend vers 0 par définition d'un point de Lebesgue. De même, $f(y) = \lim_{k\to+\infty}\fint_{B_{k,y}} f$, et $|f(x)-f(y)|$ est bien dominé par le membre de droite de (16), comme annoncé. $\square$

A ce stade, il est raisonnable de continuer avec le chapitre 18 sur Lusin et Egoroff (la première partie est indépendante, et la seconde utilise un peu de Rademacher et la seconde extension de Whitney). Mais je n'ai pas osé changer l'ordre de mes chapitres.

14. MESURE DE HAUSDORFF

14.a. Définitions

Le mieux est de commencer par la définition. Pour plus de précisions sur ce qui suit, voir par exemple le livre de Mattila. On se donne un ensemble (quelconque!) $E \subset \mathbb{R}^n$, et un nombre réel $d \in [0, n]$. On commence par poser, pour $\delta > 0$,

(1)
$$\mathcal{H}^d_\delta(E) = \inf \Big\{ \sum_{j \in \mathbb{N}} \mathrm{diam}(A_j)^d \,;\, E \subset \bigcup_{j \in \mathbb{N}} A_j \text{ et } \mathrm{diam}(A_j) \leq \delta \text{ pour tout } j \Big\}.$$

Ceci est une fonction décroissante de δ, ce qui permet de poser

$$\mathcal{H}^d(E) = \lim_{\delta \to 0} \mathcal{H}^d_\delta(E) = \sup_{\delta > 0} \mathcal{H}^d_\delta(E). \tag{2}$$

Quelques commentaires et énoncés sans démonstration.

Comme toujours en théorie de le mesure, il est important de ne prendre que des familles au plus dénombrables d'ensembles A_j. On aurait aussi prendre des boules, on aurait obtenu une mesure légèrement différente (mais quand même comprise entre $C^{-1}\mathcal{H}^d$ et $C\mathcal{H}^d$). Il y a plein d'autres variantes possibles.

Il est facile de vérifier que $\mathcal{H}^d$ est une "mesure extérieure métrique" sur $\mathbb{R}^n$ (muni de la distance euclidienne). Mesure extérieure signifie que $\mathcal{H}^d(\emptyset) = 0$, que $\mathcal{H}^d(E) \leq \mathcal{H}^d(F)$ quand $E \subset F$, et surtout que $\mathcal{H}^d\big(\cup_{j \in \mathbb{N}} E_j\big) \leq \sum_{j \in \mathbb{N}} \mathcal{H}^d(E_j)$ pour les unions dénombrables. Métrique signifie que $\mathcal{H}^d(E \cup F) = \mathcal{H}^d(E) + \mathcal{H}^d(F)$ quand $\mathrm{dist}(E, F) > 0$. Ce qui n'est d'ailleurs pas le cas de $\mathcal{H}^d_\delta$, parce que si E et F sont proche, le même A_j peut servir efficacement à recouvrir à la fois E et F.

Par rapport aux mesures sur une tribu, l'avantage des mesures extérieures est que $\mu(E)$ est défini pour tout E. L'inconvénient est que l'on n'a pas forcément l'additivité pour les ensembles disjoints. En fait, si μ^* est une mesure extérieure sur Ω, on note $\mathcal{A}$ l'ensemble des $A \subset \Omega$ tels que $\mu^*(E) = \mu^*(E \cap A) + \mu^*(E \setminus A)$ pour tout $E \subset \Omega$. On appelle ça des ensembles mesurables pour μ^*, et c'est une tribu surlaquelle μ^* définit une mesure (conventionnelle). Dans l'autre sens, si μ est une mesure sur l'espace mesurable $(\Omega, \mathcal{A})$, on peut poser $\mu^*(E) = \inf\big\{\mu(A)\,;\, A \in \mathcal{A} \text{ et } E \subset A\big\}$, et μ^* est une mesure extérieure qui prolonge μ et pour laquelle tous les $A \in \mathcal{A}$ sont mesurables.

L'avantage d'avoir une mesure extérieure métrique est qu'alors tous les boréliens sont mesurables. C'est donc le cas pour $\mathcal{H}^d$. Sinon, on serait ennuyés par l'additivité, et c'est pour ça qu'on évite souvent de jouer avec les $\mathcal{H}^d_\delta$. Attention quand même: $\mathcal{H}^d$ n'est pas sigma-finie, ce qui nous force parfois à nous restreindre à une partie de $\mathbb{R}^n$ de $\mathcal{H}^d$-mesure finie ou sigma-finie avant d'appliquer certains théorèmes.

Signalons aussi que, essentiellement par invariance par translations, la restriction à $\mathbb{R}^d$ (ou à un plan de dimension d) de la mesure $\mathcal{H}^d$ est un multiple (avec une constante fixe et connue) de la mesure de Lebesgue. La vérification demande quand même une petite partie pénible, qui est de vérifier que la constante n'est pas 0. En fait, on en reparlera un peu plus loin, au moment où l'on dira que la densité est écrite pour normaliser de façon que pour un plan, la densité soit 1.

Au cas où l'on voudrait intégrer des fonctions positives qui ne sont pas mesurables, on peut toujours utiliser l'intégrale supérieure $\int^* f d\mu = \inf_g\{\int g\mu\}$, où l'inf est pris sur les fonctions g mesurables telles que $g \geq f$. Quand f est la fonction caractéristique de $E \subset \Omega$, on retrouve la mesure extérieure de E.

La mesure de Hausdorff $\mathcal{H}^d(E)$ est une fonction décroissante de d. En fait, il est assez facile de vérifier que pour $E \subset \mathbb{R}^n$ donné, il existe d_0 tel que $\mathcal{H}^d(E) = +\infty$ pour $d < d_0$ et $\mathcal{H}^d(E) = 0$ pour $d > d_0$. On appelle d_0 la dimension de Hausdorff de E; elle est supérieure à la dimension topologique (pour ceux qui connaissent!). Les trois cas, $\mathcal{H}^{d_0}(E) = 0, \mathcal{H}^{d_0}(E) = +\infty$, et $0 < \mathcal{H}^{d_0}(E) < \infty$ sont possibles.

On renvoie à Mattila (et aux TD) pour de nombreux exemples et compléments.

14.b. Densités

Sauf mention du contraire, on travaille à n et d fixés, donc on les oublie un peu des notations. Pour $A \subset \mathbb{R}^n$ et $a \in \mathbb{R}^n$, notons

$$\theta^*(A, a) = \limsup_{r \to 0} (2r)^{-d} \mathcal{H}^d(A \cap B(a, r)),$$

et

$$\theta_*(A, a) = \liminf_{r \to 0} (2r)^{-d} \mathcal{H}^d(A \cap B(a, r)),$$

et $\theta(A, a) = \theta^*(A, a) = \theta_*(A, a)$ quand ces deux derniers (les densités supérieure et inférieure) coïncident. Le facteur de normalisation 2 est destiné à faire que la densité soit 1 quand A est un d-plan et $a \in A$. Autrement dit, on a choisi 2 parce que

$$\mathcal{H}^d(\mathbb{R}^d \cap B(0, 1)) = \operatorname{diam}(B(0, 1))^d = 2^d. \tag{3}$$

Ceci demande quand même une bonne petite vérification bien calculatoire qu'on passera sous silence, dont l'essence est que les recouvrements de la boule unité de $\mathbb{R}^d$ par des boules est essentiellement optimal dans la définition de $\mathcal{H}^d(B(0,1))$; c'est la même histoire que pour calculer la constante α telle que $\mathcal{H}^d = \alpha\lambda$ (la mesure de lebesgue) sur $\mathbb{R}^d$. Et les calculs se basent sur l'inégalité isodiamétrique, qui dit qu'à diamètre donné, les boules maximisent la mesure de Lebesgue.

Le résultat suivant est d'usage presque constant.

Théorème. *Si $A \subset \mathbb{R}^n$ est tel que $\mathcal{H}^d(A) < +\infty$, alors*

$$2^{-d} \leq \theta^*(A,a) \leq 1 \textit{ pour } \mathcal{H}^d\textit{-presque tout } a \in A. \tag{4}$$

Si de plus A est $\mathcal{H}^d$-mesurable,

$$\theta^*(A,a) = 0 \textit{ pour } \mathcal{H}^d\textit{-presque tout } a \in \mathbb{R}^n \setminus A. \tag{5}$$

Noter qu'il n'est pas trop difficile de trouver des ensembles boréliens A tels que $0 < \mathcal{H}^d(A) < +\infty$, et tels que $\theta_*(A,a) = 0$ pour tout a. On se débrouille juste avoir de petits anneaux très épais qui ne rencontrent pas A, on utilise (4) pour dire que la densité supérieure n'est pas trop grande, et on en déduit que $\theta_*(A,a) = 0$ pour tout a. La partie plus désagréable de la démonstration consiste à s'assurer que $\mathcal{H}^d(A) > 0$, typiquement en sortant une mesure de notre manche (destinée à être équivalente à la mesure de Hausdorff). Voir un exercice plus bas.

Pour démontrer la première partie de (4), posons

$$B_k = \{a \in A\,;\, \mathcal{H}^d(A \cap B(a,r)) < \frac{k}{k+1}\, 2^{-d}(2r)^d \text{ pour } 0 < r < 2^{-k}\} \tag{6}$$

pour tout entier $k \geq 1$, et notons que si a viole la première inégalité de (4), donc si $\theta^*(A,a) < 2^{-d}$, alors $a \in B_k$ pour k assez grand. Donc il suffira de montrer que $\mathcal{H}^d(B_k) = 0$ pour tout k.

Fixons k, et soit $\varepsilon > 0$. Posons $\delta = 2^{-k-1}$. Par définition de $\mathcal{H}^d_\delta(B_k)$, on peut recouvrir B_k par des ensembles E_j, $j \in \mathbb{N}$, avec $\mathrm{diam}(E_j) \leq 2^{-k-1}$ pour tout j, tels que

$$\sum_j \mathrm{diam}(E_j)^d \leq \mathcal{H}^d_\delta(B_k) + \varepsilon \leq \mathcal{H}^d(B_k) + \varepsilon. \tag{7}$$

Ne gardons que les j pour lesquels E_j rencontre B_k en (au moins) un point qu'on appelle x_j. Posons aussi $r_j = \mathrm{diam}(E_j)$. Ainsi $B_k \cap E_j \subset A \cap \overline{B}(x_j, r_j)$. Mais on a décidé de travailler avec des boules ouvertes, donc on choisit aussi un $\lambda > 1$, très proche de 1, et on note que $B_k \cap E_j \subset A \cap B(x_j, \lambda r_j)$. Donc

$$
\begin{aligned}
\mathcal{H}^d(B_k) &\leq \sum_j \mathcal{H}^d(B_k \cap E_j) \leq \sum_j \mathcal{H}^d(A \cap B(x_j, \lambda r_j)) \leq \sum_j \frac{k\lambda^d}{k+1} r_j^d \\
&= \frac{k\lambda^d}{k+1} \sum_j \mathrm{diam}(E_j)^d \leq \frac{k\lambda^d}{k+1} [\mathcal{H}^d(B_k) + \varepsilon].
\end{aligned}
\tag{8}
$$

On obtient $\mathcal{H}^d(B_k) = 0$ en prenant $\lambda > 1$ tel que $\frac{k\lambda^d}{k+1} < 1$, puis ε assez petit (noter que $\mathcal{H}^d(B_k) \leq \mathcal{H}^d(A) < +\infty$). On en déduit la première moitié de (4).

Pour la seconde moitié, on va remplacer A par un ensemble borélien A' qui le contient et dont la mesure est finie. Pour l'existence d'un tel ensemble, on peut sans doute utiliser les critères généraux, mais le plus simple est peut-être d'utiliser une suite $\{\delta_k\}$ qui tend vers 0, puis pour chaque k un recouvrement de A par une famille dénombrable d'ensembles compacts qui minimise pratiquement le second membre de (1) avec δ_k, et de prendre l'intersection dénombrable des ensembles ainsi construits. Si la seconde inégalité de (4) est vraie pour A', elle l'est pour A. Donc supposons que A est borélien.

Fixons $t > 1$ et posons $Z_t = \{a \in A; \theta^*(A, a) > t\}$. Il suffit de voir que $\mathcal{H}^d(Z_t) = 0$. Notons μ la restriction de $\mathcal{H}^d$ à A. C'est une mesure borélienne finie, et par régularité, pour tout $\varepsilon > 0$ on peut trouver un ouvert V qui contient Z_t tel que $\mu(V) \leq \mu(Z_t) + \varepsilon$. Autrement dit, $\mathcal{H}^d(A \cap V) \leq \mathcal{H}^d(Z_t) + \varepsilon$.

Par définition de Z_t, pour tout $x \in Z_t$, il existe des rayons $r_x > 0$ arbitrairement petits (donc en particulier tels que $B(x, r_x) \subset V$), tels que $\mu(B(x, r_x)) = \mathcal{H}^d(A \cap B(x, r_x)) \geq t(2r_x)^d$. Par le lemme de presque-recouvrement de Besicovitch (fin du chapitre 1 sur les recouvrements), on peut trouver une famille au plus dénombrable des fermetures de ces boules, de manière qu'elles soient disjointes et recouvrent μ-presque tout Z_t. On peut d'ailleurs choisir toutes ces boules de diamètre $< \delta$, où $\delta > 0$ est donné quelconque. Notons $\overline{B}(x, r_x)$, $x \in X$, la collection de ces boules, et S leur union.

Noter que $\mathcal{H}^d_\delta(Z_t \setminus S) \leq \mathcal{H}^d(Z_t \setminus S) = 0$ (par définition de $\mathcal{H}^d$ comme borne supérieure), donc

$$\mathcal{H}^d_\delta(Z_t) \leq \mathcal{H}^d_\delta(S) + \mathcal{H}^d_\delta(Z_t \setminus S) = \mathcal{H}^d_\delta(S) \leq \sum_{x\in X} (2r_x)^d \leq t^{-1} \sum_{x\in X} \mathcal{H}^d(A \cap B(x, r_x))$$
$$\leq t^{-1} \sum_{x\in X} \mathcal{H}^d(A \cap \overline{B}(x, r_x)) \leq t^{-1}\mathcal{H}^d(A \cap V) \leq t^{-1}[\mathcal{H}^d(Z_t) + \varepsilon]$$

par ce que l'on vient de dire et puisque les $\overline{B}(x, r_x)$ recouvrent S. Ceci vaut pour tout $\delta > 0$, donc $\mathcal{H}^d(Z_t) \leq t^{-1}[\mathcal{H}^d(Z_t) + \varepsilon]$. Si $\mathcal{H}^d(Z_t) > 0$, on obtient une contradiction en prenant ε assez petit. Donc (4) est vraie. Si ceci vous rappelle les démonstrations de densités pour les mesures, c'est normal.

Passons à (5). Fixons A mesurable. Il suffit de montrer que pour tout $t > 0$, l'ensemble $Z_t = \{x \in \mathbb{R}^n \setminus A; \theta^*(A, x) > t\}$ est de $\mathcal{H}^d$-mesure nulle.

Notons encore μ la mesure définie par $\mu(E) = \mathcal{H}^d(E \cap A)$. C'est une mesure borélienne finie, et de plus $\mu(\mathbb{R}^n \setminus A) = 0$ (parce que A est mesurable). En particulier, μ est régulière, donc il existe un ouvert V qui contient $\mathbb{R}^n \setminus A$ et tel que $\mu(V) \leq \varepsilon$. Pour $x \in Z_t$, on peut trouver $r_x > 0$ tel que $B(x, r_x) \subset V$ et $\mu(B(x, r_x)) = \mathcal{H}^d(A \cap B(x, r_x)) \geq t(2r_x)^d$. Cette fois, on utilise le lemme de recouvrement le plus simple, pour recouvrir Z_t avec des boules $B(x, 5r_x)$, $x \in X$, telles que les $B(x, r_x)$ sont disjointes. Alors

$$\mathcal{H}^d_\infty(Z_t) \leq \sum_{x\in X} \operatorname{diam}(B(x, 5r_x))^d = \sum_{x\in X} (10r_x)^d$$
$$\leq 5^d t^{-1} \sum_{x\in X} \mu(B(x, r_x)) \leq 5^d t^{-1}\mu(V) \leq 5^d t^{-1}\varepsilon$$

qui est aussi petit qu'on veut. Autrement dit, $\mathcal{H}^d_\infty(Z_t) = 0$, et on peut recouvrir Z_t avec des boules B_i telles que $\sum_i r_i^d$ soit aussi petit qu'on veut. Le sup des diamètres est également aussi petit qu'on veut. On en déduit que $\mathcal{H}^d(Z_t) = 0$, comme souhaité. Ceci termine la démonstration du théorème. □

Corollaire. *Si A et B sont $\mathcal{H}^d$-mesurables, avec $B \subset A$ et $\mathcal{H}^d(A) < +\infty$, alors, pour $\mathcal{H}^d$-presque tout $x \in B$, on a que $\theta^*(B, x) = \theta^*(A, x)$ et $\theta_*(B, x) = \theta_*(A, x)$.*

En effet, la différence est contrôlée par la densité supérieure de $A \setminus B$ au point x, qui est nulle presque-partout hors de $A \setminus B$. □

14.c. Effet des fonctions lipschitziennes

On commence par le plus facile: si $f : \mathbb{R}^n \to \mathbb{R}^m$ est C-Lipschitzienne alors (trivialement à partir des définitions)

$$\mathcal{H}^d(f(A)) \leq C^d \mathcal{H}^d(A) \text{ pour tout } A \subset \mathbb{R}^n. \tag{9}$$

Rappelons également que si $f : \mathbb{R}^n \to \mathbb{R}^m$ est Lipschitzienne, elle est différentiable presque-partout (vu en TD, avec des hypothèses plus faibles, sous le nom de théorème de Rademacher-Calderón).

Il se trouve que si $f : \mathbb{R}^n \to \mathbb{R}^m$ est Lipschitzienne, alors pour tout cube $Q_0 \subset \mathbb{R}^n$ et tout $\varepsilon > 0$, il existe $g : \mathbb{R}^n \to \mathbb{R}^m$ de classe C^1 telle que la mesure de Lebesgue (dans $\mathbb{R}^n$) de $\{x \in Q_0; f(x) \neq g(x)\}$ est inférieure à ε. Voir le chapitre 18.

Exercice 1. (Donné, comme le suivant, sans garantie. La dernière fois que j'ai regardé, c'était encore peu clair avec des fautes.) On se donne un ensemble $E \subset \mathbb{R}^n$ borélien, et $0 \leq d \leq n$.

1. On suppose qu'il existe une constante C_0 et une suite de rayons r_k tendant vers 0, tels que pour tout k, E est recouvert par $C_0 r_k^{-d}$ boules de rayon r_k. Montrer que $\mathcal{H}^d(E) \leq CC_0$.

2. On suppose qu'il existe une mesure (borélienne) finie μ sur E (autrement dit, une mesure borélienne finie μ telle que $\mu(\mathbb{R}^n \setminus E) = 0$), et des constantes $C_1 \geq 0$ et $r_0 > 0$ telles que $\mu(B(x,r)) \geq C_1^{-1} r^d$ pour $x \in E$ et $0 < r < r_0$. Montrer que $\mathcal{H}^d(A) \leq CC_1\mu(A) < +\infty$ pour $A \subset E$ borélien.

Exercice 2. On se donne $0 < \lambda < 1/2$, et on veut construire un ensemble de Cantor $K \subset [0,1]$ tel que $0 < \mathcal{H}^d(K) < +\infty$, où d est tel que $\lambda^d = 1/2$, et avec une densité inférieure nulle partout.

1. On prend λ' tel que $\lambda < \lambda' < 1/2$. Construire une suite $\{r_n\}$ avec les propriétés suivantes. D'abord, $r_0 = 1$ et $r_{n+1} \leq \lambda' r_n$ pour $n \geq 0$. Ensuite, $r_n \geq \lambda^n$ pour tout $n \geq 0$. Enfin, il existe une suite $\{n_j\}$ tendant vers $+\infty$ pour laquelle

$$\lim_{j \to +\infty} \frac{r_{n_j}}{\lambda^{n_j}} = 1 \text{ et } \lim_{j \to +\infty} \frac{r_{n_j}}{r_{n_j - 1}} = 0. \tag{1}$$

Vous trouverez peut-être plus facile de travailler sur $a_n = \log(\lambda^{-n} r_n)$.

2. Construire un ensemble de Cantor $K = \cap_n K_n$, où chaque K_n est composé de 2^n intervalles $I_{n,k}$ de longueur r_n, situés à distance $\geq c r_{n-1}$ les uns des autres (pour une constante $c > 0$ qui dépend de λ').

3. Dire pourquoi il existe une mesure naturelle sur K, telle que $\mu(I_{n,k}) = 2^{-n}$.

4. Utiliser l'exercice précédent pour montrer que $0 < \mathcal{H}^d(K) < +\infty$. Noter au passage que $C^{-1}\mu(I_{n,k}) \leq \mathcal{H}^d(I_{n,k}) \leq C\mu(I_{n,k})$ pour tout intervalle $I_{n,k}$.

5. On note $v(r) = r^{-d} \sup \mu(I)$, où la borne supérieure est prise sur les intervalles I de longueur r. Montrer que $\liminf_{r\to 0} v(r) = 0$ en utilisant la seconde partie de (1) (et le fait que la distance mentionnée en **2.** est bien cr_{n-1}). Pour la suite, je conseille même de faire un rapide dessin du graphe de $v(r)$.

6. En déduire que pour $x \in K$, la densité inférieure $\liminf_{r\to 0} r^{-d}\mathcal{H}^d(K \cap B(x,r))$ est nulle.

7. Quelle est la dimension de Hausdorff de $K \times K \subset \mathbb{R}^2$?

8. [un peu plus compliqué?] Montrer qu'il existe deux suites $\{r_k\}$ et $\{s_k\}$ comme ci-dessus telles que si $v(r)$ est associé come ci-dessus à $\{r_k\}$, et $v(r)$ à la suite $\{s_k\}$, alors

$$\lim_{r\to 0} \min(v(r), w(r)) = 0.$$

Utiliser ceci pour montrer que $\mathcal{H}^{2d}(K \times K') = +\infty$, où K et K' sont construits comme ci-dessus, à partir des suites $\{r_k\}$ et $\{s_k\}$, et naturellement $K \times K' \subset \mathbb{R}^2$. Sauf erreur de ma part, la construction donne même K et K' de dimension λ tels que $K \times K'$ est de dimension de Hausdorff $> 2\lambda$.

15. ENSEMBLES RECTIFIABLES

15.a. Les définitions équivalentes

On commence par une définition qui peut paraître bizarre. Comme souvent, d est fixé dans la suite. Ce qui suit n'a de sens que si d est un entier.

Definition 1. *Soit $E \subset \mathbb{R}^n$ un ensemble. On dira qu'il est rectifiable (ou d-rectifiable pour préciser) s'il existe une collection (au plus) dénombrable d'ensembles $A_j \subset \mathbb{R}^d$ et des applications lipschitziennes $f_j : A_j \to \mathbb{R}^n$, telles que*

$$\mathcal{H}^d\big(E \setminus \bigcup_j f_j(A_j)\big) = 0. \tag{2}$$

La terminologie n'est pas entièrement standard; ici je suis Mattila mais Federer disait "dénombrablement rectifiable", et réservait la rectifiabilité à des ensembles de mesure finie. Par définition, donc, la mesure de Hausdorff de dimension d de E est sigma-finie. On n'a pas supposé E mesurable (et alors le membre de gauche de (2) est une mesure extérieure), mais en tout cas il est contenu dans un ensemble mesurable.

Avec un peu de chance vous avez déjà entendu parler de courbes rectifiables (comprendre aussi, de longueur finie). Evidemment les courbes rectifiables sont des ensembles rectifiables (cela revient à en trouver un paramétrage lipschitzien). Et sauf erreur de ma part, le mot rectifiable vient de ce qu'on imagine les redresser pour en calculer la longueur.

Le fait que, par Rademacher, le paramétrage est dérivable presque partout (plus Sard pour dire que les points stationnaires donnent un ensemble de mesure nulle) amène presque à la conclusion que les courbes rectifiables ont une droite tangente presque partout. Je dis "presque" parce qu'il faudrait aussi s'occuper des points doubles ou multiples où la courbe se croise, avec deux ou plusieurs tangentes différentes, et montrer que l'ensemble correspondant est de mesure nulle. Exercice possible à faire, demandant je crois le même genre de technique de séparation des ensembles à base de (15.5), que l'on va utiliser plus bas.

Très facile: l'union au plus dénombrable d'ensembles rectifiables est encore rectifiable, ainsi qu'un sous-ensemble d'ensemble rectifiable.

Noter qu'on peut prolonger les applications lipschitziennes f_j à $\mathbb{R}^d$ tout entier, et comme E est encore contenu dans l'union des $f_j(\mathbb{R}^d)$, plus l'ensemble négligeable de (2). Autrement dit, dans la définition 1, on peut supposer que $A_j = \mathbb{R}^d$.

Ou encore, on peut supposer que chaque A_j est une boule, ou un cube (puisqu'on peut recouvrir $\mathbb{R}^d$ par des boules ou des cubes), et même est la boule ou le cube unité (puisqu'on peut toujours remplacer f_j par $f_j(a \cdot + b)$).

Vérifions maintenant qu'on peut se débrouiller avec des fonctions de classe C^1, avec des dérivées de rang maximal.

Pour ce qui suit, on suppose connu le chapitre sur Lusin. On verra plus tard comment on pourrait se débrouiller sans ceci, mais en attendant c'est plus pratique.

Lemme 3. *Si $E \subset \mathbb{R}^n$ est d-rectifiable, il existe des boules ouvertes $B_j \subset \mathbb{R}^d$, $j \in \mathbb{N}$, et des applications $g_j : B_j \to \mathbb{R}^n$ de classe C^1, telles que $Dg_j(x)$ est de rang d pour tout $x \in B_j$, et pour lesquelles*

$$\mathcal{H}^d\big(E \setminus \bigcup_j g_j(B_j)\big) = 0. \tag{4}$$

On pourrait d'ailleurs prendre chaque B_j égale à la boule ouverte unité de $\mathbb{R}^d$. Pour la démonstration, et à cause de ce qui précède, il suffira de recouvrir $f(B)$, quand $B \subset \mathbb{R}^d$ est une boule (disons, fermée) et $f : B \to \mathbb{R}^n$ est Lipschitzienne.

On applique le (second) corollaire du chapitre sur Lusin, qui dit que pour tout $\varepsilon > 0$ il existe un compact $E \subset B$ tel que $|B \setminus E| \leq \varepsilon$, et une fonction g de classe C^1 (carrément définie sur $\mathbb{R}^d$, mais on n'aurait besoin que de B) telle que $f = g$ sur E. Et par conséquent

$$\mathcal{H}^d(f(B) \setminus g(B)) \leq \mathcal{H}^d(f(B \setminus E)) \leq C^d \mathcal{H}^d(B \setminus E) \leq C^d \varepsilon, \tag{5}$$

où C est la constante de Lipschitz de f.

Donc en prenant les valeurs successives $\varepsilon = 2^{-k}$, on trouve des fonctions $g_k : B \to \mathbb{R}^m$, de classe C^1, telles que $\mathcal{H}^d\big(f(E) \setminus \bigcup_k g_k(B)\big) = 0$.

C'est presque ce qu'on voulait, mais maintenant il faut encore s'assurer que $Dg_j(x)$ est de rang d.

On prend une fonction $g_j = B_j \to \mathbb{R}^m$ de classe C^1, et on note Z_j l'ensemble des points critiques de g_j, c'est-à-dire des points x où $Dg_j(x)$ n'est pas de rang d. Par le **Théorème de Sard** (en principe vu en exercice comme application simple des lemmes de recouvrement, et qui marche encore avec la même démonstration quand g est à valeurs dans $\mathbb{R}^n$ avec $n \neq d$), $\mathcal{H}^d(g_j(Z_j)) = 0$. Donc il ne nous reste plus qu'à recouvrir $B_j \setminus Z_j$, qui est ouvert (parce que Dg_j est continue, et son injectivité est une condition ouverte), par de petites boules $B_{j,k}$, et à recouvrir $g_j(B_j)$ par les $g_j(B_{j,k})$. On oublie donc juste $g_j(Z_j)$, mais c'est un ensemble négligeable. Bref, la collection des $g_j : B_{j,k} \to \mathbb{R}^n$ convient pour le lemme 3. □

Encore un petit effort, pour ceux qui préfèrent les graphes aux variétés. Mais ce qui suit est juste un lemme standard sur les variétés. On suppose que $d < n$ sans perdre de généralité, puisque autrement $\mathbb{R}^n$ tout entier est d-rectifiable.

Lemme 6. *Soient $E \subset \mathbb{R}^n$ un ensemble d-rectifiable, et $\varepsilon > 0$. Il existe des plans P_j de dimension d dans $\mathbb{R}^n$, $j \in \mathbb{N}$, et des fonctions $f_j : P_j \to P_j^\perp$ de classe C^1 et qui sont également ε-Lipschitziennes, telles que si Γ_j désigne le graphe de f_i,*

$$\mathcal{H}^d\big(E \setminus \bigcup_j \Gamma_j\big) = 0. \tag{7}$$

On a noté $P_j^\perp$ le complémentaire orthogonal de P_j dans $\mathbb{R}^n$, et logiquement $\Gamma_j = \{x + f_j(x)\,;\, x \in P_j\}$. En bref, les Γ_j sont des images par des isométries de $\mathbb{R}^n$ de graphes de fonctions C^1 et ε-Lipschitziennes (de $\mathbb{R}^d$ dans $\mathbb{R}^{n-d}$).

Donnons nous $B \subset \mathbb{R}^d$ (une boule ouverte) et $g : B \to \mathbb{R}^n$, de classe C^1 et avec une dérivée de rang d en chaque point. Compte tenu du lemme précédent, il suffira de recouvrir $g(B)$.

Soit $x \in B$. Notons P l'image de $Dg(x)$; c'est donc un plan de dimension d. Notons π la projection orthogonale sur P. On peut appliquer le théorème d'inversion locale à $\pi \circ g : B \to P$, puisque sa différentielle en x est inversible. On trouve qu'il existe $r > 0$ tel que la restriction de $\pi \circ g$ à $B(x, r)$ est injective, son image est un voisinage ouvert V de $\pi \circ g(x)$ dans P, et la réciproque $\varphi : V \to B(x, r)$ est de classe C^1 aussi. On note $f = \pi^\perp \circ g \circ \varphi$, définie sur V et à valeurs dans $P^\perp$, où l'on a noté $\pi^\perp = I - \pi$ la projection sur $P^\perp$. C'est bien une fonction de classe C^1, ainsi que

$$h = I + f = I + \pi^\perp \circ g \circ \varphi = \pi \circ g \circ \varphi + \pi^\perp \circ g \circ \varphi = g \circ \varphi$$

qui va de V dans $\mathbb{R}^n$. On note que $g(B(x, r)) = g \circ \varphi(V) = h(V)$, donc le graphe de f au-dessus de V recouvre bien $g(B(x, r))$.

Posons encore $t = \pi \circ g(x)$, et notons que $Df(t) = D\pi^\perp \circ Dg(x) \circ D\varphi(t) = 0$ parce que l'image de $Dg(x)$ est P, surlequel $D\pi^\perp$ est nul. Donc en fait $|Df| \leq \varepsilon$ sur un voisinage V_0 de t dans P. Et si l'on choisit $r_1 < r$ assez petit, alors $V_1 = \pi \circ g(B(x, r_1))$ est contenu dans V_0. Par le même raisonnement que plus haut, $g(B(x, r_1)) = g \circ \varphi(V_1) = h(V_1)$, donc le graphe de f au-dessus de V_1 recouvre bien $g(B(x, r_1))$.

On laisse vérifier par le lecteur que la restriction de f à V_1 a une extension qui est encore de classe C^1 et est ε-Lipschitzienne (ou $C\varepsilon$-Lipschitzienne, si c'est plus facile à vérifer). Sinon, on pourrait aussi se contenter du graphe Γ de f au-dessus d'une boule centrée en t, qui reouvre un $g(B(x, r_2))$ et qui produit le même effet à la fin.

On a donc réussi, pour chaque $x \in B$, à recouvrir l'image par g d'un voisinage de x par un graphe Γ comme dans l'énoncé du lemme. Par compacité, on recouvre aussi $g(K)$, pour tout compact $K \subset B$, par une union finie de tels graphes. Et, en recouvrant B par une union dénombrable de compacts K, on trouve un recouvrement de $g(B)$ par des graphes qui conviennent pour le lemme 6. $\square$

Signalons avant de continuer que les réciproques des lemmes 3 et 6 sont triviales: un ensemble E qui vérifie la conclusion d'un de ces lemmes est bien rectifiable.

Et encore une remarque qu'on ne va énoncer que pour le lemme 6, mais qui aurait un analogue pour le lemme 3. Notons

$$E_j = E \cap \Gamma_j \setminus \bigcup_{k<j} \Gamma_k \,; \tag{8}$$

alors les E_j sont disjoints, et leur union est aussi l'union des $E \cap \Gamma_j$. Donc, si on pose $E' = E \setminus \big(\cup_j E_j \big)$, alors $\mathcal{H}^d(E') = 0$, et

$$E = E' \cup \bigcup_j E_j, \quad \text{une union disjointe.} \tag{9}$$

Et E_j est le graphe de la fonction f_j au-dessus de l'ensemble $A_j = \pi_j(E_j) \subset P_j$, où π_j est la projection orthogonale sur P_j.

Notons aussi que si E est mesurable (ce qui sera le cas à peu près systématiquement), chacun des morceaux est mesurable aussi, puisque les Γ_j sont boréliens (et ça serait pareil avec les $g_j(B_j)$).

Avant de passer aux propriétés des ensembles rectifiables, voici quelques exemples, qui sont utiles pour se rendre compte qu'il faudra être un peu prudent dans les énoncés de régularité ci-dessous.

Exemple 1. L'image d'une courbe rectifiable (de longueur finie) est 1-rectifiable (utiliser un paramétrage par la longueur d'arc).

Exemple 2. Les rationnels, ou n'importe quelle suite dense. (Tout est dans la partie de mesure nulle).

Exemple 3. Prenons $d = n - 1$ pour simplifier, et une suite $\{x_j\}$ dans $\mathbb{R}^n$. Alors $E = \cup_j \partial B(x_j, r_j)$ est rectifiable, à condition que $\sum_j r_j^d < +\infty$.

Contrexemple 4. On verra sans doute plus tard que l'ensemble de Cantor à 4 coins obtenu en remplaçant chaque carré Q de chaque génération par 4 carrés 4 fois plus petits situés dans Q aux 4 coins de Q est de dimension 1, avec une mesure de Hausdorff finie, mais totalement non rectifiable.

15.b. Densité, projections, et plans tangents approchés

Voyons maintenant quelques propriétés agréables des ensembles rectifiables. On va parler de densité, de projections sur des d-plans, et de plans tangents approximant.

Théorème 10 (sur la densité). *Si E est rectifiable et $\mathcal{H}^d$-mesurable, et si $\mathcal{H}^d(E) < +\infty$, alors la densité $\theta(E,x)$ existe et vaut 1 pour $\mathcal{H}^d$-presque tout $x \in E$. Par ailleurs, $\theta(E,x) = 0$ pour $\mathcal{H}^d$-presque tout $x \in \mathbb{R}^n \setminus E$.*

On rappelle que les densités sont définies par

$$\theta^*(E,x) = \limsup_{r\to 0}(2r)^{-d}\mathcal{H}^d(E \cap B(x,r)),$$

et

$$\theta_*(E,r) = \liminf_{r\to 0}(2r)^{-d}\mathcal{H}^d(E \cap B(x,r)),$$

et $\theta(E,x) = \theta^*(E,a) = \theta_*(E,x)$ quand ces deux derniers (les densités supérieure et inférieure) coïncident.

Noter au passage que si on oublie de dire que $\mathcal{H}^d(E) < +\infty$, le théorème devient faux. Prendre une union dénombrable dense de droites parallèles.

On sait déjà (à cause d'un théorème du chapitre 14.b précédent) que

$$\theta^*(E,x) \leq 1 \text{ pour } \mathcal{H}^d\text{-presque tout } x \in E \tag{11}$$

et

$$\theta^*(E,x) = 0 \text{ pour } \mathcal{H}^d\text{-presque tout } x \in \mathbb{R}^n \setminus E. \tag{12}$$

Donc il nous reste juste à démontrer que pour tout $\varepsilon > 0$,

$$\theta_*(E,x) \geq 1 - C\varepsilon \text{ pour } \mathcal{H}^d\text{-presque tout } x \in E, \tag{13}$$

où C est une constante géométrique, et puisqu'alors on aura aussi $\theta_*(E,x) \geq 1$ pour $\mathcal{H}^d$ presque-tout $x \in E$).

On va appliquer le lemme 6, et la décomposition de (9). Peut-être faudrait-il dire, à ce point, que la densité est calculable et égale à 1 en tout point quand E est un graphe de fonction C^1. Ce qui va suivre est juste basé sur cette observation (et peut-être, j'ai pas bien relu, contient une démonstration un poil trop compliquée de ce fait). Il suffit de prouver que pour tout j,

$$\theta_*(E,x) \geq 1 - C\varepsilon \text{ pour } \mathcal{H}^d\text{-presque tout } x \in E_j \tag{14}$$

On va prouver qu'en fait

$$\theta_*(\Gamma_j,x) \geq 1 - C\varepsilon \text{ pour } \mathcal{H}^d\text{-presque tout } x \in \Gamma_j \tag{15}$$

(et donc aussi pour $\mathcal{H}^d$-presque tout $x \in E_j$). Maintenant, et comme pour (12),

$$\theta^*(\Gamma_j \setminus E_j, x) = 0 \text{ pour } \mathcal{H}^d\text{-presque tout } x \in E_j \tag{16}$$

parce que E_j est mesurable, de $\mathcal{H}^d$-mesure finie, et disjoint de $\Gamma_j \setminus E_j$. En soustrayant ceci de (15), on trouve que pour $\mathcal{H}^d$-presque tout $x \in E_j$

(17)

$$\begin{aligned}
\theta_*(E, x) \geq \theta_*(E_j, x) &= \liminf_{r\to 0} (2r)^{-d}\mathcal{H}^d(E_j \cap B(x,r)) \\
&\geq \liminf_{r\to 0} (2r)^{-d}\big[\mathcal{H}^d(\Gamma_j \cap B(x,r)) - \mathcal{H}^d(\Gamma_j \setminus E_j \cap B(x,r))\big] \\
&\geq \liminf_{r\to 0} (2r)^{-d}\mathcal{H}^d(\Gamma_j \cap B(x,r)) - \limsup_{r\to 0} (2r)^{-d}\mathcal{H}^d(\Gamma_j \setminus E_j \cap B(x,r)) \\
&= \theta_*(\Gamma_j, x) - \theta^*(\Gamma_j \setminus E_j, x) \geq 1 - C\varepsilon,
\end{aligned}$$

ce dont on déduira (14) dès qu'on aura vérifié (15).

Donc on s'occupe de (14). En fait on pourrait même montrer que $\theta(\Gamma_j, x) = 1$ pour $x \in \Gamma_j$, parce que Γ_j est un graphe C^1, mais pour se simplifier la vie, on va plutôt utiliser le fait que c'est un graphe lipschitzien et prouver (14). Ca nous évitera de changer encore de coordonnées.

Donc on se donne $x \in \Gamma_j$, on pose $t = \pi_j(x)$, et on observe que pour $s \in P_j \cap B(t, (1-\varepsilon)r)$, et $h_j = I + f_j$ comme avant,

$$|h_j(s) - x| = |h_j(s) - h_j(t)| \leq |s-t| + |f_j(s) - f_j(t)| \leq (1+\varepsilon)|s-t| < r \tag{18}$$

donc $h_j(s) \in B(x,r)$. Ce qui signifie que $s = \pi_j(h_j(s)) \in \pi_j(\Gamma_j \cap B(x,r)$, et alors

(19)

$$\mathcal{H}^d(\Gamma_j \cap B(x,r)) \geq \mathcal{H}^d(\pi_j(\Gamma_j \cap B(x,r))) \geq \mathcal{H}^d(P_j \cap B(t,(1-\varepsilon)r)) = 2^d(1-\varepsilon)^d r^d$$

par normalisation de $\mathcal{H}^d$. On laisse tendre r vers 0 et on trouve que $\theta_*(\Gamma_j, x) \geq (1-\varepsilon)^d$, ce qui donne (14).

On en déduit le théorème sur la densité. $\square$

Passons maintenant aux projections sur des d-plans.

Théorème 20 (sur les projections). *Si E est rectifiable et mesurable et si $\mathcal{H}^d(E) > 0$, alors $\mathcal{H}^d(\pi_W(E)) > 0$ pour tout d-plan vectoriel W, sauf peut-être ceux qui sont "perpendiculaires" à un certain d-plan V.*

On a noté π_W la projection orthogonale sur W. Perpendiculaire signifie qu'au moins un vecteur non nul de V est orthogonal à W. Ou encore, que la restriction de π_W à V n'est pas injective (c'est comme ça que cela arrivera dans la démonstration). C'est sans doute un mauvais terme. En tout cas, $\mathcal{H}^d(\pi_W(E)) > 0$ pour presque tout W, si on met une mesure correcte sur la grassmanienne des d-plans.

On ne peut pas espérer un bien meilleur résultat, en tout cas avec ce type de formulation. Les exceptions possibles sont justifiées par le cas où E est (contenu dans) un d-plan V. Le fait qu'on ne donne aucune borne inférieure sur $\mathcal{H}^d(\pi_W(E))$ pour aucun W vient de ce que, même avec la normalisation $\mathcal{H}^d(E) = 1$, E peut très bien être froissé et recroquevillé dans une minuscule boule, avec donc de toutes petites projections. C'est dommage mais avec la définition de la rectifiabilité (juste qualitative et invariante par unions dénombrables), ce genre de choses est prévisible.

Malgré tout, si l'on tient compte des multiplicités dans le calcul des projection, on arrive à une estimation des projections en moyenne qui est plus correcte. Une formule du genre suivant:

$$\mathcal{H}^d(E) = C(n,d) \int_{W \in \mathcal{G}(n,d)} \int_W \sharp[\pi_W^{-1}(y)]\, d\mathcal{H}^d(y) dW \tag{21}$$

où $\mathcal{G}(n,d)$ est la grassmannienne des d-plans vectoriels de $\mathbb{R}^n$, et dW une mesure de probabilité sur $\mathcal{G}(n,d)$ invariante par rotations. Supposez quand même que $\mathcal{H}^d(E) < +\infty$ pour me rassurer. Noter la ressemblance avec la formule de la coaire ci-dessous, qu'il faudrait alors appliquer sur un ensemble rectifiable. Et il n'est pas si surprenant que ça marche; il s'agit en gros de découper E en bouts de graphes (comme dans le lemme 6), et de regarder la contribution de chaque morceau. On moyennise sur la grassmannienne en constatant que la moyenne du jacobien de π_W le long du plan tangent à E (voir ci-dessous) est une constante. On verra des calculs comme ceci (pour l'aire et la coaire), mais on ne les fera pas ici.

Voir Mattila ou Federer, par exemple, sous la rubrique "integralgeometric measure".

La démonstration se fait comme pour le théorème 10. Pour simplifier, on découpe E comme au lemme 6, et on choisit un j tel que $\mathcal{H}^d(E_j \cap \Gamma_j) > 0$.

Ensuite on choisit $x \in E_j \cap \Gamma_j$ tel que $\theta^*(\Gamma_j \setminus E_j, x) = 0$ (qui existe à cause de (16), peut-être ajouter une petite étape supplémentaire: se restreindre à l'intersection de E avec une partie bornée de Γ_j, pour avoir une mesure finie avant d'appliquer le théorème du chapitre précédent).

On va montrer que $\mathcal{H}^d(\pi_W(E)) > 0$ dès que π_W est injective sur P, où P est le plan tangent à Γ_j en x. Il nous reste à appliquer encore le théorème d'inversion locale. Rappelons que Γ_j est le graphe d'une fonction $f_j : P_j \to P_j^\perp$, de classe C^1. On note aussi $h_j = f_j + I$, et donc Γ_j est paramétré par $h : P \to \mathbb{R}^n$. Enfin, posons $t = \pi_j(x)$ (la projection sur P_j). La différentielle de $\pi \circ h$ en t est injective, comme composée de $Dh(t)$ (injective à cause de sa composante sur P_j) et $D\pi(x)$, injective sur l'image P de $Dh(t)$. Le théorème d'inversion locale s'applique, et donne un inverse φ de classe C^1 à $\pi \circ h$, défini dans un voisinage de $\pi(x)$.

Soit M la constante de Lipschitz de $h \circ \varphi$, restreinte à une petite boule $B = B(\pi(x), r_0)$. Pour $r < 0$ et $y \in B(\pi(x), r)$, noter $w = h \circ \varphi(y)$; alors $y \in B(x, Cr)$ et $y = \pi(w)$. Donc $B(\pi(x), r) \subset \pi(\Gamma_j \cap B(x, Cr))$ et maintenant il ne reste plus qu'à dire que pratiquement tous les points de $\Gamma_j \cap B(x, Cr)$ viennent de E:

(22)
$$\begin{aligned}\mathcal{H}^d(\pi(E \cap B(x, Cr)) &\geq \mathcal{H}^d(\pi(\Gamma_j \cap B(x, Cr)) - \mathcal{H}^d(\pi(\Gamma_j \setminus E \cap B(x, Cr)))\\ &\geq \mathcal{H}^d(P \cap B(\pi(x), r)) - \mathcal{H}^d(\Gamma_j \setminus E \cap B(x, Cr))\\ &\geq 2^d r^d - \mathcal{H}^d(\Gamma_j \setminus E \cap B(x, Cr)).\end{aligned}$$

Puisque $\theta^*(\Gamma_j \setminus E_j, x) = 0$, $\mathcal{H}^d(\Gamma_j \setminus E \cap B(x, Cr))$ est négliegable par rapport à r^d quand r tend vers 0, donc $\mathcal{H}^d(\pi(E \cap B(x, Cr)) > 0$ pour r assez petit, ce qui suffit pour notre énoncé. C'est un peu vexant de perdre autant dans les estimées, mais d'un autre coté il n'y a pas d'estimées plus quantitatives, et aussi il y a une sorte de réciproque: les ensembles totalement non rectifiables (voir plus loin) ont des projections de mesure nulle dans presque toute direction. □

Passons aux plans tangents approchés (sans doute le plus amusant?). On va noter, lorsque V est un d-plan vectoriel de dimension d et pour $\varepsilon > 0$,

$$V_\varepsilon = \{x \in \mathbb{R}^n\,;\, \mathrm{dist}(x, V) \leq \varepsilon\} \tag{23}$$

(un ε-voisinage de V) et

$$C(V, \varepsilon) = \{x \in \mathbb{R}^n\,;\, \mathrm{dist}(x, V) \leq \varepsilon |x|\} \tag{24}$$

(un voisinage conique).

Un (vrai) plan tangent à E en x est un plan (affine) $x + V$, où V est un d-plan vectoriel de dimension d, tel que pour tout $\varepsilon > 0$, il existe $r > 0$ tel que

$$E \cap B(x, r) \subset x + C(V, \varepsilon). \tag{25}$$

Ou de manière équivalente (exercice!), pour tout $\varepsilon > 0$, il existe $r_0 > 0$ tel que

$$E \cap B(x,r) \subset x + V_{\varepsilon r} \quad \text{pour } r \leq r_0. \tag{26}$$

Evidemment, pour un graphe lipschitzien, l'existence d'un plan tangent en un point correspond à la différentiabilité de la fonction au point correspondant. (J'ai dit lipschitzien pour éviter de parler de tangentes verticales).

Malheureusement, l'exemple 3 montre que même si on se permet de retirer de E un ensemble de mesure nulle, il n'y a pas forcément de plan tangent. On sera donc obligé d'utiliser une définition un peu plus compliquée, qui est plus stable quand on ajoute des ensembles de densité nulle en un point. [Densité nulle est donc mieux ici que mesure nulle.]

Définition 27. *On dira que $x+V$ est un plan tangent approché (approximate tangent plane) à E en x si*

$$\theta^*(E,x) > 0 \tag{28}$$

et, pour tout $\varepsilon > 0$,

$$\lim_{r\to 0} r^{-d}\mathcal{H}^d(E \cap B(x,r) \setminus [x + C(V,\varepsilon)]) = 0. \tag{29}$$

Bref, la densité en x de $E \setminus [x + C(V,\varepsilon)]$ est nulle. La première partie est destinée à éviter des indéterminations; elle ne nous coûtera pas cher de toute façon.

Cette définition a l'avantage insigne que si l'on ajoute à E un ensemble dont la densité en x est nulle, alors $x+V$ est encore un plan tangent approché à l'union en x. On s'en servira dans les preuves, pour localiser.

Exercice. Vérifier que l'on a (29) pour tout $\varepsilon > 0$ si et seulement si, pour tout $\varepsilon > 0$, on a

$$\lim_{r\to 0} r^{-d}\mathcal{H}^d(E \cap B(x,r) \setminus [x + V_{\varepsilon r}]) = 0. \tag{30}$$

Exercice. Donner un exemple où E a plus d'un d-plan tangent approché. Prendre $d = 2$ et $n = 3$. On peut même se débrouiller pour que E soit de densité 1 en x, et ait une (vraie) droite tangente.

Exercice. On se donne un ensemble E Ahlfors-régulier de dimension d dans $\mathbb{R}^n$. Montrer que si E a un d-plan tangent aproché $x+V$ en x, alors $x+V$ est vrai plan tangent à E. Montrer aussi que V est unique.

Theorème 31. *Soit $E \subset \mathbb{R}^n$ un ensemble mesurable d-rectifiable, avec $\mathcal{H}^d(E) < +\infty$. Alors pour $\mathcal{H}^d$-presque tout $x \in E$, E a un unique plan tangent approché en x.*

En oubliant de dire que $\mathcal{H}^d(E) < +\infty$, une union dénombrable dense de droites parallèles est un contre-exemple. C'est exclu dans la démonstration quand on applique le théorème du paragraphe 14.b pour dire que la densité de $E \setminus f_j(A_j)$ est nulle en presque tout point de $f_j(A_j)$.

On commence la démonstration comme d'habitude, en utilisant le lemme 6. On démontre d'abord l'existence. On se moque des points de la partie négligeable qui n'est pas dans les Γ_j, donc il suffit de vérifier qu'en presque tout $x \in E \cap \Gamma_j$, le plan tangent à Γ_j est aussi un plan tangent approché à E.

Et, à nouveau comme en (16), $\theta_*(\Gamma_j \setminus E, x) = 0$ pour presque tout $x \in E \cap \Gamma_j$. Il est facile de voir qu'en chacun de ces points x, le plan tangent à Γ_j est un plan tangent approché à E. La condition (28) sur la densité vient par exemple du fait que $\theta_*(\Gamma_j \setminus E, x) = 0$, et $\theta_*(\Gamma_j, x) > 0$, et pour ce qui est de (29), on note qu'ajouter un ensemble de densité nulle à $\Gamma_j \cap E$ a un effet nul sur (29). Voilà pour l'existence.

Reste l'unicité. On garde le même point $x \in E \cap \Gamma_j$ tel que $\theta_*(\Gamma_j \setminus E, x) = 0$, et on essaie un autre d-plan $x+Q$ que le plan tangent $x+P$ qu'on a trouvé (à cause de (28), les plans tangents approchés passent toujours par x).

Soit v un vecteur unitaire qui est dans $P \setminus Q$. Un tel vecteur existe, car P et Q sont différents et de même dimension. Notons $d = \mathrm{dist}(v, Q) > 0$.

Posons, pour r petit, $B_r = B(x+rv, dr/2)$, et notons que $\mathcal{H}^d(\Gamma_j \cap B_r) \geq cr^d$ pour r asser petit, pour une certaine constante $c > 0$ qui dépend de d. C'est facile; faire un dessin.

Comme $r^{-d}\mathcal{H}^d(B_r \cap \Gamma_j \setminus E) \leq r^{-d}\mathcal{H}^d(B(x, 2r) \cap \Gamma_j \setminus E)$, qui tend vers 0, on trouve aussi que $\mathcal{H}^d(\Gamma_j \cap E \cap B_r) \geq cr^d/2$ pour r petit. Mais pour ε assez petit, B_r ne rencontre pas $x + C(Q, \varepsilon)$, donc $r^{-d}\mathcal{H}^d(E \cap B(x, 2r) \setminus [x + C(Q, \varepsilon)]) \geq \mathcal{H}^d(\Gamma_j \cap E \cap B_r) \geq c/2$, qui ne tend pas vers 0. Donc Q n'est pas un plan tangent approché. $\square$

15.c. Comment fait-on sans Lusin? Un petit lemme avec du degré

On a bien utilisé la description donnée dans les lemmes 3 et 6, parce que les fonctions de classe C^1 et leur graphe sont plus agréables à utiliser que leurs analogues lipschitziens.

Mais la démonstration à base de Lusin et Whitney prend un certain temps, et on aurait pu à la place donner des démonstrations directes en restant dans le cadre C^1.

Voici juste une liste des ingrédients qui permettent de se débrouiller, mais je ne ferai pas (plus) l'argument, qui est quand même plus lourd que ci-dessus.

On garde quand même le fait que, par Rademacher, les fonctions lipschitziennes sont différentiables presque partout (utile pour choisir des points où la densité inférieure de $f(B)$ est 1 (ou $1-\varepsilon$ pour se faciliter la vie), et de même pour choisir un point avec plan tangent pour une bonne partie de $f(B)$.

De même, le théorème de Sard (sur l'image de l'ensemble des points où la différentielle n'est pas de rang d) vaut encore, avec la même démonstration, et c'est bien utile pour permettre d'étudier $f(B)$ en en un point $f(x)$ tel que $Df(x)$ est de rang d.

Par contre, le théorème d'inversion locale doit être remplacé, et en particulier il faut un moyen de montrer que si par exemple $g : \mathbb{R}^d \to \mathbb{R}^d$ est Lipschitzienne, et si $Dg(0)$ existe et est de rang d, alors pour r assez petit, $f(B(0,r))$ contient $B(f(0),cr)$, où $c > 0$ est par exemple une borne inférieure pour $\frac{1}{2}Df(0)\cdot v$, lorsque v est un vecteur unitaire.

Le lemme suivant permet de démontrer ceci.

Lemme 32. *Soient B la boule unité fermée dans $\mathbb{R}^n$, $0 < \tau < 1$, et et $f : B \to \mathbb{R}^n$ une application continue telle que $|f(x) - x| \leq \tau$ pour tout $x \in \partial B$. Alors $f(B)$ contient $B(0, 1-\tau)$.*

Il va s'agir d'utiliser un peu de topologie (par exemple, de théorie du degré), cachée dans le lemme suivant qu'on ne démontrera pas.

Lemme 33. *L'application identique sur la sphère unité S n'est pas homotope, parmi les applications continues de S dans S, à une application constante.*

Ou de manière équivalente, il n'existe pas de fonction continue $h : B \to S$, qui soit égale à l'identité sur S.

Quand $n = 2$, vous avez sans doute vu un moyen de faire ceci, si vous avez vu l'indice d'un point a par rapport à une courbe γ, qui compte le nombre algébrique de fois que la courbe tourne autour de a, et qui se calcule en nombres complexes par une formule du type $n(a) = \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-a}$. Ce nombre a le bon goût de ne pas changer lorsqu'on fait varier a et γ continûment, à condition que a ne traverse jamais γ. S'il existait h continue (du disque fermé à valeurs dans le cercle), ou pourrait s'en servir pour déformer un lacet γ égal au cercle (parcouru une fois) en lacet constant, tout en restant à valeurs dans le cercle (je sais, ça paraît difficile à faire, mais c'est normal), et alors l'indice de $a = 0$ serait 1 au départ mais 0 à l'arrivée.

Donc on admet le lemme 33, et on en déduit le lemme 32. On procède par contradiction, en se donnant $z \in B(0, 1-\tau)$, et en supposant que z n'est pas dans l'image de f. On va construire une homotopie impossible entre l'identité et une constante, et on en déduira la contradiction souhaitée.

Première étape (sur trois). Puisque $z \notin f(B)$, on définit une fonction continue $h : B \to \partial B$, par $h(y) = [f(y) - z]/|f(y) - z|$ pour $y \in B$. L'application $(y, t) \to h(ty)$ donne une homotopie qui va de la constante $h(0)$ à la restriction de h à ∂B, en restant comme souhaité parmi les applications continues qui vont de ∂B dans ∂B.

Ensuite, on construit une homotopie qui part de la restriction de h à ∂B et va à la restriction de $f(y)/|f(y)|$ à ∂B, en utilisant l'homotopie $(y, t) \to [f(y) - tz]/|f(y) - tz|$, qui est bien définie et continue parce qu'on va vérifier que $f(y) \neq tz$ pour $y \in \partial B$ et $t \in [0, 1]$. En effet, $|tz| < 1 - \tau$ puisque $z \in B(0, 1-\tau)$, alors que $|f(y)| \geq |y| - \tau = 1 - \tau$ puisque $|f(y) - y| \leq \tau$ et $y \in \partial B$.

Enfin, la restriction de $f(y)/|f(y)|$ est homotope à $y/|y|$, en utilisant l'homotopie $(y, t) \to [(1-t)f(y)+ty]/|(1-t)f(y)+ty|$. A nouveau, il faut juste vérifier que le dénominateur ne s'annule pas, et c'est vrai car $(1-t)f(y)+ty = y+(1-t)(f(y)-y)$ donc $|(1-t)f(y)+ty| \geq |y|-(1-t)|f(y)-y| \geq |y|-\tau > 0$.

En composant tout ça, on trouve l'homotopie interdite annoncée ci-dessus; la contradiction prouve le lemme 32. □

15.d. Ensembles purement non rectifiables.

Passons maintenant à l'étude rapide, sans presque de démonstrations, des ensembles purement non rectifiables.

Definition. *On dit que l'ensemble $E \subset \mathbb{R}^n$ est totalement non rectifiable (en anglais, "purely unrectifiable", et Besicovitch disait "irregular") si $\mathcal{H}^d(E \cap f(\mathbb{R}^d)) = 0$ pour toute application Lipschitzienne $f : \mathbb{R}^d \to \mathbb{R}^n$.*

Ou, de manière équivalente, si $\mathcal{H}^d(E \cap F) = 0$ pour tout F rectifiable. Noter que F est totalement non rectifiable quand $F \subset E$ et E est totalement non rectifiable. De même, une union dénombrable d'ensembles totalement non rectifiables est totalement non rectifiable. Ajouter un ensemble $\mathcal{H}^d$-négligeable ne change pas la totale non rectifiabilité non plus.

Exemples: l'ensemble de Cantor à 4 coins, un flocon de neige. Voir les TD. Et aussi essayer de suivre ce qui suit en pensant à l'ensemble de Cantor notamment.

Proposition. *Soit $E \subset \mathbb{R}^n$, tel que E soit union dénombrable d'ensembles $\mathcal{H}^d$-mesurables de $\mathcal{H}^d$-mesure finie. Il existe A rectifiable tel que $E \setminus A$ est totalement non rectifiable.*

Autrement dit, comme on peut remplacer A par $A \cap E$ (qui est également rectifiable), E a la décomposition $E = A \cup B$, avec A rectifiable et $B = E \setminus A$ totalement non rectifiable. Une telle décomposition est d'ailleurs unique module les ensembles de mesure nulle. En effet, si $E = A' \cup B'$ est une autre décomposition, $\mathcal{H}^d(B' \cap A) = 0$ puisque B' est totalement non rectifiable, donc $A \subset A \cap (A' \cup B')$ est $\mathcal{H}^d$-presque contenu dans A'. De même, A' est $\mathcal{H}^d$-presque contenu dans A. Comme $E = A' \cup B'$ est $\mathcal{H}^d$-presque une partition ($A' \cap B'$ est à la fois rectifiable et purement non rectifiable), on en déduit la presque unicité.

Pour démontrer la proposition, on commence par le cas où $\mathcal{H}^d(E) < +\infty$, on pose $M = \sup\{\mathcal{H}^d(E \cap A)\,;\, A \text{ rectifiable}\}$, et on se donne une suite $\{A_k\}$ d'ensembles rectifiables, avec $\mathcal{H}^d(E \cap A_k) \geq M - 2^{-k}$. On prend $A = \cup_k A_k$; c'est encore un ensemble rectifiable, et s'il existait B rectifiable tel que $\mathcal{H}^d(B \cap E \setminus A) > 0$, on aurait aussi que $\mathcal{H}^d(E \cap (A_k \cup B)) = \mathcal{H}^d(E \cap A_k) + \mathcal{H}^d(E \cap B) \geq M - 2^{-k} + \mathcal{H}^d(E \cap B \setminus A) > M$ pour k grand, une contradiction.

Quand E est seulement union dénombrable d'ensembles $\mathcal{H}^d$-mesurables de $\mathcal{H}^d$-mesure finie, on décompose chacun et on prend les unions.

Je n'ai pas eu le courage de vérifier ce qui se passe quand E n'est pas mesurable; peut-être que ça marche encore quand même. □

D'abord, noter que les bonnes propriétés signalées ci-dessus pour un ensemble rectifiable restent vraies $\mathcal{H}^d$-presque-partout sur la partie rectifiable d'un ensemble (mesurable de mesure finie) quelconque. C.-à.-d., E a une densité égale à 1 et un d-plan tangent approché en $\mathcal{H}^d$-presque tout $x \in A$ (où A est la partie rectifiable, comme ci-dessus). Vérification facile, puisque $E \setminus A$ a une densité nulle presque-partout sur A, et que nos propriétés sont insensibles à l'ajout d'un ensemble de densité nulle au point considéré.

Et maintenant il est assez facile de vérifier que l'ensemble de Cantor à 4 coins est totalement non rectifiable, puisque il n'y a, presque partout, ni plan tangent approché ni densité. (Aisément vérifiable à la main, cette fois.).

Les ensembles totalement non rectifiables ont toutes les "mauvaises" propriété qui correspondent aux bonnes propriétés des ensembles rectifiables. Voyons plus en détails. Ce qui suit est une combinaison de résultats de Marstrand, Mattila, Preiss, et d'autres; voir le livre de Mattila pour plus de détails et d'autres résultats du même type. Le lecteur est prévenu que les démonstrations sont intéressantes, mais délicates. La notion de mesure tangente y est fort utile.

On commence par les densités. Soit $E \subset \mathbb{R}^n$ un ensemble $\mathcal{H}^d$-mesurable, avec $\mathcal{H}^d(E) < +\infty$. Si E est totalement non rectifiable, alors $\theta_*(E,x) < \theta^*(E,x) \leq 1$ pour $\mathcal{H}^d$-presque tout $x \in E$.

Réciproquement, si $\theta_*(E,x) < 1$ pour $\mathcal{H}^d$-presque tout $x \in E$, alors E est totalement non rectifiable. En effet, on a vu que si E a une partie rectifiable, alors la densité inférieure y est p.p. ≥ 1.

On a aussi une version uniforme de la partie directe: on sait même qu'il existe une constante $c(d) < 1$ telle que si $\theta_*(E,x) \geq c(d)$ pour $\mathcal{H}^d$-presque tout $x \in E$, alors E est rectifiable. Du coup, $\theta_*(E,x) \leq c_d$ sur la partie non rectifiable de tout ensemble $\mathcal{H}^d$-mesurable tel que $\mathcal{H}^d(E) < +\infty$. En effet, si c'était faux sur un morceau Z, Z serait rectifiable.

Quand $d = 1$, Besicovitch a conjecturé que la constante $c(1)$ optimale est $1/2$, mais les résultats partiels les plus précis (Preiss-Tišer, Farag) n'y sont pas encore.

Pour les projections: si E est $\mathcal{H}^d$-mesurable, avec $\mathcal{H}^d(E) < +\infty$, et totalement non rectifiable, alors pour presque-tout d-plan vectoriel V, $\mathcal{H}^d(\pi_V \cap E) = 0$. On a noté π_V la projection orthogonale sur V, et on a muni la Grassmannienne $\mathcal{G}(n,d)$ des d-plans dans $\mathbb{R}^n$ de n'importe quelle mesure raisonnable (qui serait localement équivalente à la mesure de Lebesgue, quand on voit $\mathcal{G}(n,d)$ comme une variété C^1 avec des cartes; mais en fait il y a une sympathique mesure invariante).

Ceci n'empêche pas que la projection soit de mesure strictement positive pour certaines directions: les projections sur les droites de pente $1/2$ et -2 de l'ensemble de Cantor à 4 coins K envoient, de façon injective sauf en un ensemble dénombrable de points, l'ensemble K sur un segment. Et d'ailleurs, par cette projection, le codage des points de K par $4^{\mathbb{N}}$ correspond au développement d'un point de $[0,1]$ en base 4. On ne sait pas trop ce qu'il en est pour les autres directions (sauf, donc, que la mesure de la projection est nulle pour presque toute direction).

Pour les plans tangents, si E est $\mathcal{H}^d$-mesurable, avec $\mathcal{H}^d(E) < +\infty$, et totalement non rectifiable, alors pour presque-tout $x \in E$, E n'a aucun d-plan tangent approché en x.

Une dernière petite remarque. Dans beaucoup d'arguments précédents, on a utilisé le fait qu'en un sens, la décomposition de E en deux parties (mesurables) disjointes (par exemple, la partie rectifiable et la partie non rectifiable, ou avant $E \cap \Gamma_j$ et $E \setminus \Gamma_j$) se faisait proprement, au sens ou presque-partout sur un morceau, la densité supérieure de l'autre morceau est nulle. Mais si ceci est vrai au sens de la mesure, c'est faux du point de vue topologique. Il est assez facile de construire un ensemble 1-rectifiable E, disons de mesure $\mathcal{H}^1(E) = 1$, tel que $\overline{E} \setminus E$ est purement non rectifiable, et de mesure 1 aussi.

Dans le même ordre d'idées, la rectifiabilité ne passe pas bien à la limite: les aproximations par des unions de (frontières de) carrés de l'ensemble de Cantor de Garnett (4 coins) sont rectifiables, mais la limite ne l'est pas.

16. AIRE ET COAIRE

Pour ce qu'on va en dire, on aurait sans doute pu le mettre dans le chapitre sur les fonctions Lipschitziennes, et avant de parler de rectifiabilité. On commence par la formule de l'aire.

Théorème 1. *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction Lipschitzienne. Notons $J_f(x)$ le jacobien de Df en x (explication plus tard). Notons aussi, pour $y \in \mathbb{R}^m$ et $A \subset \mathbb{R}^n$ borélien, $N_{f,A}(y) \in [0, +\infty]$ le nombre de points de $A \cap f^{-1}(y)$. Alors $N_{f,A}$ est borélienne et*

$$\int_{R^m} N_{f,A}(y)\, d\mathcal{H}^n(y) = \int_A J_f(x)\, d\mathcal{H}^n(x) \tag{2}$$

pour tout $A \subset \mathbb{R}$ borélien. On a aussi, pour $g : \mathbb{R}^n \to \mathbb{R}$ borélienne positive ou intégrable sur $\mathbb{R}^n$ contre J_f,

$$\int_{R^m} \sum_{x \in f^{-1}(y)} g(x)\, d\mathcal{H}^n(y) = \int_{\mathbb{R}^d} g(x)\, J_f(x)\, d\mathcal{H}^n(x). \tag{3}$$

Quelques commentaires avant de faire la démonstration pour de vrai. D'abord, $J_f(x)$ est défini partout où f est différentiable. Si $m = n$, c'est juste la valeur absolue du déterminant de $Df(x)$, et la proposition est une généralisation du théorème de changement de variable que vous connaissez. Si $Df(x)$ est de rang $< n$, $J_f(x) = 0$. Dans le cas restant, disons qu'on note V l'image de $\mathbb{R}^n$ par $Df(x)$, qu'on identifie V à $\mathbb{R}^n$ par une isométrie φ, et que $J_f(x)$ est la valeur absolue du déterminant de $\varphi \circ Df(x)$. Ca n'est pas la définition officielle (qui est plus invariante et fait par exemple intervenir des formes n-linéaires alternées), mais on va se contenter de ça et renvoyer à Federer pour les définitions précises.

Quand f est injective, le membre de gauche est juste $\mathcal{H}^n(f(A))$, donc c'est bien l'aire de l'image qu'on calcule (avec multiplicité dans le cas général).

Noter que le résultat n'est intéressant que quand $m \geq n$; sinon les deux membres sont nuls.

La seconde partie du théorème (à savoir (3)) se déduit du reste par des arguments standards de théorie de la mesure: on note que (2) donne (3) quand g est une fonction caractéristique de borélien, puis (par combinaison linéaire) une fonction étagée borélienne, puis, par beppo-Lévi, une fonction borélienne positive, puis par soustraction, une fonction intégrable.

La formule (de la première partie) se localise très bien, au sens où s'il existe une famille dénombrable d'ensembles boréliens A_j disjoints tels que la formule est vraie pour chaque A_j, elle est aussitôt vraie pour $A = \cup_j A_j$: on note que $N_{f,A} = \sum_j N_{f,A_j}$, puis on applique Beppo-Lévi. Ca servira pour la démonstration.

Et bien sûr, si f n'est définie que sur un sous-ensemble E de $\mathbb{R}^n$, rien n'empêche de l'étendre à $\mathbb{R}^n$, puis d'appliquer le théorème à des ensembles A contenus dans E.

Le théorème s'étend assez facilement, grâce au principe de localisation, à des fonctions f définies sur un ouvert et qui sont seulement localement lipschitziennes. Mais n'abusez pas trop, si vous oubliez toute contrainte sur f, il peut se faire (escalier du diable quand $n = 1$) que $Df = 0$ presque-partout (donc le membre de droite est nul), mais l'image de l'ensemble de mesure nulle où f n'est pas différentiable soit de mesure positive.

Il y a aussi un théorème de l'aire pour une fonction lipschitzienne $f : E \to \mathbb{R}^m$, où E est un ensemble rectifiable de dimension n. L'énoncé et la démonstration demandent un peu plus de définitions (comment définir la différentielle au moins sur le plan tangent, puis le jacobien, presque-partout sur E), mais rien de très délicat si je me souviens bien.

Passons à la démonstration. Grâce au principe de localisation on va pouvoir se restreindre à des ensembles si petits que le résultat en sera plus simple.

D'abord on peut jeter de A n'importe quel ensemble Z de $\mathcal{H}^n$-mesure nulle, puisqu'alors Z ne contribue ni au membre de droite, ni à celui de gauche (puisque $N_{f,A}$ ne change que sur $f(Z)$, qui est de mesure nulle parce que $\mathcal{H}^n(Z) = 0$ et f est Lipschitzienne). Ceci permet de supposer que f est différentiable partout sur A.

En fait, allons directement un cran plus loin (même si on pourrait sans doute s'en tirer sans cela). On sait que pour tout $\varepsilon > 0$, f coïncide avec une fonction de classe C^1 sur $A \setminus Z_\varepsilon$, où $\mathcal{H}^n(Z_\varepsilon) \leq \varepsilon$. Si on démontre (2) sur $A \setminus Z_\varepsilon$, on pourra passer à la limite (encore Beppo-Levi, ou un découpage en tranches et la remarque sur la localité), et obtenir (2) pour A. Donc on peut supposer que f est de classe C^1. Juste pour le plaisir, je vais essayer dans la suite de ne pas utilier ceci (pour gagner l'argument d'extension de Whitney qui y est utilisé), mais la démonstration se simplifie un peu pour le lecteur qui est prêt à l'utiliser. Par contre, on va quand même utliser Lusin (qui dit que Df est égale à une fonction continue, sauf sur un ensemble de mesure arbitrairement petite, voir le chapitre 18).

Donc dans tous les cas (soit par C^1, soit par différentiabilité p.p. et continuité de Df sur des grands ensembles), on a jeté la partie de A où f n'est pas différentiable, et on suppose Df continue sur A.

Par Sard (voir un exercice fait plus haut), la partie Z_c de A où Df n'est pas de rang n ne contribue pas au membre de gauche, puisque $\mathcal{H}^d(f(Z_c)) = 0$; elle ne contribue pas non plus au membre de droite (puisque $J_f = 0$ sur Z_c), donc on peut jeter Z_c.

Comme pour tout $x \in A$, $Df(x)$ est inversible, il existe trivialement un entier $M \geq 1$ tel que $Df(x)$ et $Df(x)^{-1} : V \to \mathbb{R}^n$ (où $V = Df(x)(\mathbb{R}^n)$) sont de normes $\leq M$. Ou, en le disant de manière équivalente mais sans parler de V,

$$(4) \qquad M^{-1}|u| \leq |Df(x) \cdot u| \leq M|u| \text{ pour tout vecteur } u \in \mathbb{R}^n.$$

Donc, en découpant encore A en un nombre au plus dénombrable de régions, on peut supposer que le même M marche pour tout $x \in A$.

Donnons-nous aussi $\varepsilon > 0$. On va se contenter de démontrer les inégalités souhaitées à $C\varepsilon$ près, avec un C qui pourrait d'ailleurs dépendre de M mais ça n'est pas grave car maintenant M est fixé.

Pour tout $x \in A$, il existe r_x tel que

$$(5) \quad |f(y) - f(x) - Df(x) \cdot (y - x)| \leq M^{-1}\varepsilon|x - y| \text{ pour } y \in \mathbb{R}^n \cap B(x, r_x)$$

Quitte à encore découper A en morceaux (ce qui préserve aussi les inégalités à suivre, par le même argument de Beppo-Lévi), on peut supposer que l'on peut prendre pour r_x un même $r > 0$ fixé. Et, en découpant encore, que $A \subset Q$ pour un cube de diamètre r. Donc (5) vaudra pour $x, y \in A$. En découpant encore, on peut supposer que D_f varie si peu dans A que

$$(6) \qquad |Df(x) - Df(y)| \leq M^{-1}\varepsilon \text{ pour } x, y \in A$$

et

$$(7) \qquad (1 - \varepsilon)J_f(x) \leq J_f(y) \leq (1 + \varepsilon)J_f(x) \text{ pour } x, y \in A$$

A cause de (4) et (5), $|f(y) - f(x)| \geq |Df(x) \cdot (y - x)| - M^{-1}\varepsilon|x - y| \geq M^{-1}(1 - \varepsilon)|x - y| \geq |y - x|(2M)^{-1}$, donc f est injective, $N_{f,A}(y)$ vaut 1 si $y \in f(A)$ et 0 sinon, et donc la formule (1) se simplifie en $\mathcal{H}^n(f(A)) = \int_A J_f(x) d\mathcal{H}^n(x)$. A la place, on va montrer que

$$(8) \qquad (1 - C\varepsilon)J_f(x_0)\mathcal{H}^n(A) \leq \mathcal{H}^n(f(A)) \leq (1 + C\varepsilon)J_f(x_0)\mathcal{H}^n(A),$$

où x_0 est choisi au hasard dans A. Noter que $J_f(x_0)\mathcal{H}^n(A)$ est comparable à $\int_A J_f(x)d\mathcal{H}^n(x)$ à cause de (7). On en déduira en sommant sur nos tas de petits morceaux que (2) est vraie à un facteur multiplicatif $1 \pm C\varepsilon$ près, ce qui suffit puisque ε est arbitraire.

Fixons $x_0 \in A$. Notons S une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ telle que $Df(x_0) \circ S^{-1}$ soit une isométrie. Pour se rapprocher de notre définition du jacobien $J_f(x_0)$ on peut prendre une isométrie j de $V = Df(x_0)(\mathbb{R}^n)$ dans $\mathbb{R}^n$, et prendre $S = (j \circ Df(x_0))^{-1}$. Noter que $J_f(x_0) = |det(j \circ Df(x_0))|$, donc $|det(S)| = J_f(x_0)^{-1}$. A cause de ceci, du fait que $\mathcal{H}^n$ est proportionnel à dx dans $\mathbb{R}^n$, et au théorème de changement de variable dans le cas des applications linéaires bijectives de $\mathbb{R}^n$ (que donc vous êtes prêts à croire), on trouve que

$$\mathcal{H}^n(S(A)) = J_f(x_0)^{-1}\mathcal{H}^n(A). \tag{9}$$

Posons $g = f \circ S^{-1}$ (définie sur $S(A)$. Notons $z_0 = S^{-1}(x_0)$ Pour $z, z' \in S(A)$, en en notant $x = S^{-1}(z)$ et $x' = S^{-1}(z')$, $g(z') - g(z) = f(x') - f(x)$, qui est $\varepsilon M^{-1}|x' - x|$-proche de $Df(x) \cdot (x' - x)$ par (5) et donc $2\varepsilon M^{-1}|x' - x|$-proche de $Df(x_0) \cdot (x' - x) = Df(x_0) \circ S(z' - z)$. Et du coup, comme $Df(x_0) \circ S$ est une isométrie, la norme de $g(z') - g(z)$ est $2\varepsilon M^{-1}|x' - x|$-proche de celle de $z' - z$. Autrement dit,

(10)
$$\begin{aligned} ||g(z') - g(z)| - |z' - z|| &\leq ||Df(x_0) \circ S(z' - z)| - |z' - z|| + 2\varepsilon M^{-1}|x' - x| \\ &= 2\varepsilon M^{-1}|x' - x| \leq 2\varepsilon|z' - z| \end{aligned}$$

où pour la dernière inégalité on utilise le fait que $|z' - z| = |Df(x_0) \circ S(z' - z)| = |Df(x_0)(x' - x)| \geq M^{-1}|x' - x|$ puisque $Df(x_0) \circ S$ est une isométrie et par (4).

Bref, (10) dit que g est $1 \pm 2\varepsilon$-bilipschitzienne, ce qui donne

$$(1 - 2\varepsilon)^n\mathcal{H}^n(S(A)) \leq \mathcal{H}^n(g(S(A))) \leq (1 + 2\varepsilon)^n\mathcal{H}^n(S(A)) \tag{11}$$

Or $\mathcal{H}^n(S(A)) = J_f(x_0)^{-1}\mathcal{H}^n(A)$ par (9), et $\mathcal{H}^n(g(S(A))) = \mathcal{H}^n(f(A))$ par définition de g; on en déduit (8). $\square$

On essaie maintenant la formule de la coaire. Cette fois on se donne $f : \mathbb{R}^m \to \mathbb{R}^n$, avec $m > n$, et on estime la taille moyenne des surfaces de niveau en fonction du jacobien.

Théorème 2. *Soient m et n des entiers, avec $m \geq n$, et $f : \mathbb{R}^m \to \mathbb{R}^n$ une fonction Lipschitzienne. Alors pour tout borélien $A \subset \mathbb{R}^m$,*

$$\int_A J_f(x)dx = c_{m-n} \int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y))dy, \tag{11}$$

où c_{m-n} est la constante telle que la restriction de $\mathcal{H}^{m-n}$ à $\mathbb{R}^{m-n}$ soit c_{m-n} fois la mesure de Lebesgue. Et, pour g borélienne, positive sur $\mathbb{R}^m$ ou intégrable par rapport à la mesure de Lebesgue dx,

$$\int_{\mathbb{R}^m} g(x)J_f(x)dx = c_{m-n} \int_{\mathbb{R}^n} \left\{ \int_{f^{-1}(y)} g(u)d\mathcal{H}^{m-n}(u) \right\} dy. \tag{12}$$

Les commentaires d'abord. Sauf erreur de ma part, Federer a une convention différente pour la normalisation de $\mathcal{H}^d$, et c'est pour cela qu'il n'a pas besoin de c_{m-n}. Ici, on mélange un peu Lebesgue et $\mathcal{H}^d$.

Si $m = n$, utiliser la formule de l'aire pour le même résultat. La mesurabilité de $\mathcal{H}^{m-n}(A \cap f^{-1})(y)$ et (donc) de $\int_{f^{-1}(y)} g(u)d\mathcal{H}^{m-n}(u)$ viennent en prime. La seconde partie découle assez facilement de la première.

Donnons à nouveau une définition bizarre de $J_f(x)$, quand $Df(x)$ existe. Cette fois, $Df(x)$ va de $\mathbb{R}^m$ dans un $\mathbb{R}^n$ plus petit. Pour éviter les formes multilinéaires, disons que $J_f(x) = 0$ si le rang de $Df(x)$ est $< n$, et que sinon l'adjoint ${}^tD$ de $Df(x)$ est une application linéaire bijective de $\mathbb{R}^n$ dans un sous-espace V de dimension n (en fait, l'orthogonal du noyau), qu'on identifie à $\mathbb{R}^n$ par une isométrie φ, et que l'on prend pour J_f la valeur absolue du déterminant jacobien de $\varphi \circ^t D$. Une autre manière de dire est que le jacobien de Df est défini comme étant égal à celui de sa transposée.

A nouveau, rien n'empêche de se restreindre à des ensembles plus petits (si f n'était pas initialement définie partout, ou au contraire si elle est seulement localement lipschitzienne).

Cette fois, je n'ose pas essayer de le démontrer, mais je vais essayer de vous convaincre que le théorème a un sens. Voir les théorèmes 3.2.11 et 3.2.12 de [Fe], et aussi quelques pages plus loin (3.2.22) pour une extension où f est définie sur un ensemble rectifiable.

On commence par le cas le plus facile: celui où f est la projection π de $\mathbb{R}^m$ sur $\mathbb{R}^n \subset \mathbb{R}^m$. Notons (x, y) les points de $\mathbb{R}^m$, avec $x \in P = \mathbb{R}^{m-n}$ et $y \in \mathbb{R}^n$, et posons donc $\pi(x, y) = y$. Il est facile de voir que dans ce cas $J_f = 1$.

Pour $y \in \mathbb{R}^n$, $\pi^{-1}(y)$ est le $(m-n)$-plan $y+P$; On se donne $A \subset \mathbb{R}^m$; le membre de gauche de (11) est donc $|A|$. Le membre de droite est $c_{m-n} \int_{y \in \mathbb{R}^n} \mathcal{H}^{m-n}(A \cap y + P) dy = \int_{y \in \mathbb{R}^n} \big\{ \int_{x \in P} \mathbf{1}_A(x,y) dx \big\} dy$. Dans ce cas, donc, il s'agit de la formule de Fubini.

Noter aussi que si on multiplie f par λ, le Jacobien est multiplié par λ^n (et pas λ^m), donc aussi le membre de gauche de (11); les ensembles de niveaux restent les mêmes, sauf qu'ils sont maintenant paramétrés par λy, donc le membre de droite aussi est multiplié par λ^n (bref, il y a plus d'ensembles de niveaux).

Si l'on sait (11) pour f et qu'on essaie pour $f \circ \varphi$, où φ est une isométrie de $\mathbb{R}^m$, Df est remplacé par $Df \circ \varphi$, son transposé par ${}^t\varphi \circ^t Df$, et le jacobien ne bouge pas puisque ${}^t\varphi$ est une isométrie de l'image V de ${}^tDf$ sur l'image de ${}^t\varphi \circ^t Df$. Le membre de gauche ne bouge donc pas. Le membre de droite non plus, puisque $f^{-1}(y)$ est simplement remplacé par son image par l'isométrie φ^{-1}.

De même, si l'on remplace f par $\psi \circ f$, où ψ est une application linéaire bijective de $\mathbb{R}^n$, le jacobien J_f est multiplié par J_ψ (on a pré-composé ${}^tDf$ par ${}^t\psi$), et le membre de droite aussi est multiplié par J_ψ, puisqu'on garde les mêmes ensembles de niveau $(\psi \circ f)^{-1}(y) = f^{-1}(\psi^{-1}(y))$, et par un changement de variable linéaire sur $\mathbb{R}^n$.

On en sait assez pour le cas où f est linéaire (ou affine) de rang n: en précomposant par φ, on peut supposer que $V = {}^tDf(\mathbb{R}^n)$ est l'orthogonal de $P = \mathbb{R}^{m-n}$ de tout-à-l'heure. Noter que Df s'annule sur P, puisque pour $p \in P$, $\langle Df(p), y\rangle = \langle p, ^t D(y)\rangle = 0$ pour $y \in \mathbb{R}^n$. Par contre, la restriction de Df à V est une injection, parce que si $Df(x) = 0$, alors $\langle x, ^t D(y)\rangle = \langle Df(x), y\rangle = 0$ pour $y \in \mathbb{R}^n$, donc $x \in V^\perp$. Donc cette restriction est une bijection, et si on compose $f \circ \varphi$ avec l'inverse ψ de cette restriction, on trouve que $\psi \circ f \circ \varphi$ est la projection de tout-à-l'heure, pour laquelle on a vu le théorème.

Le cas de f affine de rang $< n$ est facile (le membre de droite est nul parce que $f^{-1}(y)$ est vide pour presque tout y, mais ne donne quand même pas une idée exacte de la contribution des points où $Jf(x) = 0$ dans le cas général.

Le cas où f est de classe C^1 (ou d'ailleurs Lipschitz, je crois), mais avec une dérivée qui reste de rang n est assez proche du cas linéaire pour pouvoir conclure pas trop difficilement.

Pour la contribution des points où f n'est pas différentiable, il s'agit simplement de montrer que la contribution au second membre est nulle. C'est ce que vous avez fait en TD, sauf erreur de ma part, en prouvant que

$\int_{\mathbb{R}^n} \mathcal{H}^{m-n}(A \cap f^{-1}(y)) d\mathcal{H}^n(y) \leq C||f||_{Lip}^m \mathcal{H}^m(A)$ pour f lipschitzienne (et même avec des dimensions non entières).

Et c'est, de manière surprenante, la contribution vraiment pénible est celle des points où f est différentiable, mais J_f est nulle. Il s'agit de voir que sa contribution à droite est nulle. C'est un peu plus délicat que pour le théorème de Sard, mais sauf erreur c'est le même esprit.

17. FONCTIONS A VARIATION BORNÉE

On va sans doute aller un peu vite. On renvoie au livre de Giusti pour des détails et des applications. Voir aussi celui d'Ambrosio, Fusco, et Pallara.

Deux justifications (pas incompatibles) pour l'intérêt de cet espace. Pour les applications en traitement de l'image, on veut souvent représenter une image comme somme de fonctions plutôt lisses, mais multipliées par des fonctions caractéristiques d'ensembles à bord assez régulier. La dérivée de ces fonctions comportera donc des morceaux qui sont des mesures, comme pour les fonctions de BV.

Pour définir des surfaces minimales en codimension 1, un bon moyen est de les voir comme frontières d'ensembles dans $\mathbb{R}^n$, et alors BV (et plus précisément les ensembles ce Caccioppoli décrits ci-dessous) a de bonnes propriétés de compacité, qui permettront de trouver des surfaces minimales. Dit en d'autres termes, BV est le bon cadre parce que si on se donne une suite minimisante pour certains problèmes de surfaces minimales, c'est dans BV que les sous-suites convergeront et que se trouveront les solutions. Ce qui ne nous empêchera pas, ensuite, d'essayer de démontrer que les solutions trouvées dans BV sont en fait plus régulières.

Soit f intégrable sur un domaine $\Omega \subset \mathbb{R}^n$ (c'est ce que demande Giusti; moi j'aurais volontiers seulement demandé localement intégrable). On dira que f est à variation bornée sur Ω ($f \in BV(\Omega)$) si sa dérivée au sens des distributions (sur Ω) est une mesure finie.

Autrement dit, $f \in BV(\Omega)$ s'il existe des mesures (boréliennes) finies (réelles ou même complexes pour des fonctions à valeurs complexes) $\mu_1, \cdots, \mu_n$ sur Ω telles que $\int_\Omega f \frac{\partial g}{\partial x_i} = -\int_\Omega g d\mu_i$ pour $g \in C_c^\infty(\Omega)$.

Noter que par Riesz, une mesure boréliene finie sur Ω, c'est la même chose qu'une forme linéaire continue sur $C_c(\Omega)$ (et on peut remplacer par C_c^∞ par densité, à condition bien sûr de garder la norme sup). Donc $f \in BV(\Omega)$ si pour $1 \leq i \leq n$, $|\int_\Omega f \frac{\partial g}{\partial x_i}| \leq C||g||_\infty$ pour $g \in C_c^\infty(\Omega)$. Il est d'usage de regrouper ces n conditions en une seule, en prenant g à valeurs dans $\mathbb{R}^n$, et en disant que $f \in BV(\Omega)$ s'il existe C telle que

$$(1) \qquad \Big| \int_\Omega f \mathrm{div}(g) \Big| \leq C ||g||_\infty \;\; \text{pour } g \in C_c^\infty(\Omega; \mathbb{R}^n),$$

et où $\mathrm{div}(g) = \sum_i \frac{\partial g}{\partial x_i}$ est la divergence de g. On note $\int_\Omega |Df|$ la plus petite constante C qui marche. Autrement dit,

$$(2) \qquad \int_\Omega |Df| = \sup \Big\{ \Big| \int_\Omega f \,\mathrm{div}(g) \Big| ;\, g \in C_c^\infty(\Omega; \mathbb{R}^n) \text{ avec } ||g||_\infty \leq 1 \Big\}$$

(et on peut aussi oter la valeur absolue autour de $\int_\Omega f \mathrm{div}(g)$).

Ainsi $\int_\Omega |Df|$ est la norme de la forme linéaire qui à g dans une classe dense de $C_c(\Omega; \mathbb{R}^n)$ associe $\langle \sum_i \frac{\partial f}{\partial x_i}, g_i \rangle$, donc c'est la masse de la variation totale de la mesure à valeurs vectorielles $Df = (\frac{\partial f}{\partial x_i})_i$. On rappelle que pour une mesure finie ν, il y a une mesure positive finie $|\nu|$, la variation totale de ν, telle que par exemple $d\nu = h|d\nu|$, avec une fonction mesurable h telle que $|h(x)| = 1$ pour $|\nu|$-presque tout x.

Quelques remarques. D'abord, si $\Omega' \subset \Omega$ et $f \in BV(\Omega)$, alors $f \in BV(\Omega')$ avec une norme plus petite. Aussi, si $f \in W^{1,1}(\Omega)$, alors $f \in BV(\Omega)$, avec la même norme. Autrement dit, si $Df \in L^1$, sa variation totale est bien $|Df(x)|dx$ et la masse de sa variation totale est $\int_\Omega |Df(x)|dx$.

La réciproque est fausse; la fonction caractéristique du carré unité dans $\mathbb{R}^2$ est dans BV, par exemple $\frac{\partial f}{\partial x_1}$ est e_1 fois la différence entre les mesures de longueurs sur les cotés verticaux.

Pour des fonctions définies sur un intervalle, $f \in BV$ si $f \in L^1$ et f est obtenue par intégration d'une mesure finie. Par exemple, on peut prendre la fonction de Lebesgue (escalier du diable): on part d'une mesure μ positive finie sur un ensemble de Cantor $K \subset [0, 1]$, et on pose $f(x) = \int_{[0,x]} d\mu$. Il est assez facile de vérifier que la dérivée de f au sens des distributions est bien μ (alors que son gradient est nul presque partout si K est de mesure nulle). Donc ne pas confondre dérivée au sens des distributions et dérivée presque partout.

En fait, sur un intervalle $[a, b]$, $f \in BV$ si et seulement si f est l'intégrale d'une mesure finie (la réciproque est vraie aussi). La notion de BV est la même que celle que vous connaissez peut-être déjà, où $f \in BV$ si sa variation totale (la borne supérieure des sommes $\sum_{j=1}^k |f(x_j) - f(x_{j-1}|$, où l'on considère les subdivisions $a = x_0 < x_1 \ldots < x_k = b$ de l'intervalle $[a, b]$) est finie. Les fonctions de BV sur $[a,b]$ sont aussi les différences de deux fonctions croissantes, ce qui correspond au fait que sur un intervalle, les mesures (signées) finies sont les différences de deux mesures de Stiljes.

Muni de la norme $||f||_{BV(\Omega)} = ||f||_{L^1(\Omega)} + \int_\Omega |Df|$, $BV(\Omega)$ est un espace de Banach, comme on le on vérifie sans trop de mal (surtout avec le théorème de semi-continuité ci-dessous).

Un **ensemble de Caccioppoli** (ou ensemble de périmètre fini) est un ensemble (mesurable) de $\mathbb{R}^n$ dont la fonction caractéristique est dans BV. Un peu plus généralement, E est de Caccioppoli dans Ω si $\mathbf{1}_E \in BV(\Omega)$. Et le **périmètre de E dans Ω** est la variation totale $\mathcal{P}(E,\Omega) = \int_\Omega |D\mathbf{1}_E|$. Noter que ces notions ne changent pas quand on remplace E par un ensemble E' équivalent (au sens où la mesure de la différence symétrique $E\Delta E'$ est de mesure nulle, ou encore où les fonctions caractéristiques sont égales presque partout).

Lorsque E est un domaine borné de classe C^2, disons, son périmètre est fini. En effet, pour $g \in C_c^\infty(\Omega;\mathbb{R}^n)$, Gauss-Green dit que

$$\int \mathbf{1}_E \mathrm{div}(g) = \int_E \mathrm{div}(g) = \int_{\partial E} g \cdot \nu d\sigma,$$

où ν est la normale extérieure, et la mesure $d\sigma$ est aussi la restriction de $\mathcal{H}^{n-1}$ à ∂E. Donc non seulement E est de Caccioppoli, mais en plus son périmètre est inférieur à $\mathcal{H}^{n-1}(\partial E)$, et même égal puisque $D\mathbf{1}_E$ est le produit de $d\sigma$ par $-\nu$ qui est unitaire.

Par contre, la réciproque ne marche pas du tout, ne serait-ce que parce que $D\mathbf{1}_E$ ne change pas quand on remplace E par un ensemble équivalent, alors que ∂E peut changer dramatiquement. Penser à $E = \mathbb{Q}^2$ dans $\mathbb{R}^2$, ou à un carré moins un segment.

L'intérieur d'un flocon de neige n'est pas de Caccioppoli (mais je ne sais pas si c'est si trivial à vérifier); l'ensemble de Cantor à 4 coins si, mais pour des raisons décevantes.

Pour un ensemble de Caccioppoli, on vérifie sans mal que $D\mathbf{1}_E$ est à support dans ∂E, car ailleurs, $\mathbf{1}_E$ est localement constante, donc sa dérivée est nulle (pour vérifier directement, dire que son effet sur des fonctions à support contenu dans l'intérieur ou l'extérieur est nul, en intégrant par parties). On peut encore écrire

$$\int \mathbf{1}_E \mathrm{div}(g) = \int_E \mathrm{div}(g) = \int_{\partial E} g \cdot D\mathbf{1}_E,$$

par définition de la dérivée et puisque son support est contenu dans ∂E. Ca donne une formule de Gauss-Green pour les ensembles de Caccioppoli (qui d'ailleurs sera plus facile à utiliser si on connaît mieux $D\mathbf{1}_E$).

Théorème (semicontinuité de la variation totale). *Soient $\Omega \subset \mathbb{R}^n$ et $\{f_k\}$ une suite dans $BV(\Omega)$, qui converge vers f dans $L^1(\Omega)$. On suppose que la suite est bornée dans $BV(\Omega)$, c.-à.-d., que $\int_\Omega |Df_k| \leq M$. Alors $f \in BV(\Omega)$, et $\int_\Omega |Df| \leq \liminf_{k\to+\infty} \int_\Omega |Df_k|$.*

Démonstration: il s'agit donc de voir que, si L est la liminf, $\int_\Omega f\mathrm{div}(g) \leq L$ pour $g \in C_c^\infty(\Omega;\mathbb{R}^n)$ de norme $||g||_\infty \leq 1$. C'est facile, puisque $\int_\Omega f\mathrm{div}(g) = \lim_{k\to+\infty} \int_\Omega f_k \mathrm{div} g \leq L$. $\square$

Noter que cet énoncé se localise tout seul (il peut être utile de l'appliquer à un sous-ensemble) Il vaut aussi pour une suite bornée de fonctions dans $W^{1,1}$ qui converge dans L^1. La limite est donc dans BV.

Voyons une sorte de réciproque. On va travailler dans $\mathbb{R}^n$ tout entier, voir Giusti pour la localisation.

Théorème (Approximation par des fonctions C^∞). *On se donne $f \in BV(\mathbb{R}^n)$. Il existe une suite $\{f_k\}$ de fonctions C^∞, telles que*

$$\lim_{k\to+\infty} f_k = f \quad \textit{dans } L^1(\mathbb{R}^n), \tag{3}$$

$$\int_\Omega |Df| \leq \liminf_{k\to+\infty} \int_\Omega |Df_k| \quad \textit{pour tout } \Omega \subset \mathbb{R}^n \textit{ ouvert}, \tag{4}$$

$$\limsup_{k\to+\infty} \int_F |Df_k| \leq \int_F |Df| \quad \textit{pour tout } F \subset \mathbb{R}^n \textit{ compact}, \tag{5}$$

où l'on a noté $\int_F |Df|$ l'intégrale sur F de la mesure "variation totale de Df", autrement dit $\int_F |Df| = \inf \left\{ \int_\Omega |Df|\,; \Omega \text{ est un ouvert contenant } F \right\}$ (par régularité de la dite mesure).

On aurait pu espérer que (5) vaudrait aussi pour F ouvert, mais c'est un peu naïf. Si f est la fonction caractéristique d'une boule ouverte B, et F est le complémentaire de la sphère, il y a un gros problème puisqu'il faudrait que $\int_{\mathbb{R}^n} |Df_k| = \int_F |Df_k|$ tende vers 0. Essayer $f_k = \mathbf{1}_{B(0,1+2^{-k})}$.

Par contre, il y a beaucoup d'ouverts (par exemple, $\Omega = B(0,R)$ pour presque tout R) tels que $\int_{\partial\Omega} |Df| = 0$, et pour de tels ouverts, on a bien

$$\lim_{k\to+\infty} \int_\Omega |Df_k| = \int_\Omega |Df| \tag{6}$$

Pour la démonstration, on se donne une fonction φ positive, C^∞, radiale, d'intégrale 1, puis on pose $\varphi_k(x) = 2^{kn}\varphi(2^k x)$ pour $x \in \mathbb{R}^n$, et enfin on prend $f_k = f * \varphi_k$. Le fait que f_k tend vers f dans L^1 est une histoire standard sur les approximations de l'identité, puisque $f \in L^1$. Pour (4), c'est le théorème précédent, donc il reste (5). Il suffit de voir que

$$\limsup_{k\to+\infty} \int_F |Df_k| \leq \int_\Omega |Df| \tag{7}$$

pour tout ouvert Ω borné contenant F, puisque $\int_F |Df|$ est une limite de $\int_\Omega |Df|$ pour de tels Ω. On se donne ensuite Ω', avec $\Omega' = \{x \in \mathbb{R}^n\,;\, \mathrm{dist}(x,F) < \frac{1}{2}\,\mathrm{dist}(F, \mathbb{R}^n \setminus \Omega)\}$, qui est relativement compact dans Ω. Il suffira maintenant de vérifier que pour k assez grand, $\int_{\Omega'} |Df_k| \leq \int_\Omega |Df|$ (puisque $\int_F |Df_k| \leq \int_{\Omega'} |Df_k|$). Ou encore, par définition de $\int_{\Omega'} |Df_k|$, que $|\int_{\Omega'} f_k \,\mathrm{div}(g)| \leq \int_\Omega |Df|$ pour $g \in C_c^\infty(\Omega')$, de norme $||g||_\infty \leq 1$.

Mais $\int_\Omega f_k \,\mathrm{div}(g) = \int_\Omega (f * \varphi_k)\,\mathrm{div}(g) = \int_\Omega f\,[\mathrm{div}(g) * \varphi_k] = \int_\Omega f \;\mathrm{div}(g * \varphi_k)$ par calculs en principe faciles. Or $||g * \varphi_k||_\infty \leq ||g||_\infty \leq 1$ parce que $\int |\varphi_k| = \int \varphi_k = 1$, donc $|\int_\Omega f \;\mathrm{div}(g * \varphi_k)| \leq \int_\Omega |Df|$, ce qui permet de conclure. $\square$

Une remarque parfois utile. Quand, même en dehors du théorème, on a une suite $\{f_k\}$ telle que (3) ait lieu, ainsi que (6) pour un certain Ω, on peut en déduire des estimations de semicontinuité supérieure. En effet, si F est un fermé dans Ω,

(8)

$$\begin{aligned}\limsup_{k\to+\infty} \int_F |Df_k| &= \lim_{k\to+\infty} \int_\Omega |Df_k| - \liminf_{k\to+\infty} \int_{\Omega\setminus F} |Df_k| \\ &= \int_\Omega |Df| - \liminf_{k\to+\infty} \int_{\Omega\setminus F} |Df_k| \geq \int_\Omega |Df| - \int_{\Omega\setminus F} |Df| = \int_F |Df|.\end{aligned}$$

Comment utiliser le théorème? Voici juste un exemple. Il est classique (mais malheureusement pas fait dans ces notes) qu'il existe une constante C_n

telle que, pour toute fonction C^1 à support compact sur $\mathbb{R}^n$, on ait l'inégalité de Sobolev

$$(9) \qquad ||f||_{\frac{n}{n-1}} \leq C_n \int_{\mathbb{R}^n} |\nabla f|.$$

En fait, on en sait beaucoup sur (9), jusqu'à connaître la meilleure constante C_n et les fonctions optimales.

Comme corollaire de (9) et du théorème d'approximation, on obtient facilement que

$$(10) \qquad ||f||_{\frac{n}{n-1}} \leq C_n \int_{\mathbb{R}^n} |Df| \text{ pour } f \in BV(\mathbb{R}^n).$$

En effet, on prend une suite $\{f_k\}$ comme au théorème d'approximation. Pour être honnête, ces fonctions ne sont pas à support compact, donc on les remplace par des troncatures à l'infini qui ont les mêmes propriétés; voir Giusti pour les détails.

On applique (9) à f_k, et on trouve que $||f_k||_{\frac{n}{n-1}} \leq C_n \int_{\mathbb{R}^n} |\nabla f_k| \leq C_n \int_{\mathbb{R}^n} |Df_k|$. Par ailleurs, $Df_k = Df * \varphi_k$, et comme dans la démonstration du théorème d'approximation, $\int_{\mathbb{R}^n} |Df_k| \leq \int_{\mathbb{R}^n} |Df|$ car quand on calcule la norme (agissant sur les fonctions test, muni de la norme sup), cela revient à composer par l'application $g \to g * \varphi_k$, qui est contractante.

En particulier $\{f_k\}$ est une suite bornée dans $L^{\frac{n}{n-1}}$.

Puisque $\{f_k\}$ converge vers f dans L^1, on peut extraire une sous-suite qui converge presque-partout. Alors $\displaystyle\int |f|^{\frac{n}{n-1}} \leq \liminf_{k\to+\infty} \int |f_k|^{\frac{n}{n-1}} \leq \left[C_n \int_{\mathbb{R}^n} |Df|\right]^{\frac{n-1}{n}}$ par Fatou et l'estimation ci-dessus, et (10) s'ensuit.

Encore une conséquence: l'inégalité isopérimétrique qui dit que pour $E \subset \mathbb{R}^n$ de périmètre fini et borné (ou de mesure finie),

$$(11) \qquad |E|^{\frac{n-1}{n}} \leq C_n P(E)$$

(où $P(E)$ est le périmètre de E). Appliquer Sobolev à la fonction caractéristique de E.

Dans ce cas aussi la constante optimale est connue, et les ensembles optimaux sont les boules.

De la même manière, on obtient aisément les inégalités de Poincaré pour $f \in BV(\mathbb{R}^n)$ (ou dans un domains simple) avec des exposants $1 \leq p < \frac{n}{n-1}$, à partir des mêmes inégalités valables pour $f \in W^{1,1}$. [Exercice.]

Avant de passer à une application, énonçons un petit résultat de compacité.

Théorème. *Soient Q un cube ou une boule (ouvert) de $\mathbb{R}^n$, $M > 0$, et notons A l'ensemble des fonctions $f \in BV(Q)$ telles que $\int_Q |Df| \leq M$ et $|\int_Q f| \leq M$. Alors A est une partie compacte de $L^1(Q)$.*

Notons que la restriction sur l'intégrale de f (ou quelque chose de semblable) est nécessaire, à cause des fonctions constantes et puisque $\mathbb{R}$ n'est pas compact.

Rappelons que notre ensemble est une partie complète de L^1, parceque L^1 est complet et A est fermé: la condition d'intégrale passe à la limite, et pour la condition sur la variation, on utilise le théorème de semi-continuité inférieure. Il s'agit donc de montrer que A (muni de la distance $||f - g||_1$) est précompact. On se donne donc $\varepsilon > 0$ et on doit vérifier que A peut être recouvert par un nombre fini de boules de rayon ε.

On va se contenter du cas où Q est un cube; le cas d'une boule s'y ramène par changement de variable bilipschitzien, ou en modifiant un peu la construction avec des ensembles moins propres que des petits cubes. On se donne $\eta > 0$ très petit (à calculer à la fin en fonction de ε, et on recouvre Q par des cubes R_j d'intérieurs disjoints, de diamètre $\leq \eta$. Il en faut un nombre fini.

Noter qu'à cause de l'inégalité de Poincaré, $\int_Q |f| \leq |\int_Q f| + \int_Q |f - m_Q(f)| \leq M + C_Q \int_Q |Df| \leq C'_Q M$ pour $f \in A$. Alors

$$\frac{1}{|R_j|}\int_{R_j} |f| \leq \frac{1}{|R_j|}\int_Q |f| \leq \frac{C'_Q M}{|R_j|} \quad \text{pour } f \in A. \tag{12 = 8}$$

On a besoin d'un ensemble fini. Prenons l'ensemble des fonctions g de la forme

$$g = \sum_j \alpha_j \mathbf{1}_{R_j}, \tag{9}$$

où chaque α est un multiple entier de η tel que $|\alpha| \leq \frac{C'_Q M}{|R_j|}$. Ca fait beaucoup de fonctions g, mais un nombre fini quand même. On va d'ailleurs voir que le fait que la norme de g dans BV puisse être bien plus grande que M n'a aucune importance.

Soit donc $f \in A$ donnée. On choisit g comme en (9), de manière que $\left|\alpha_j - m_j\right| \leq \eta$, en posant $m_j = \frac{1}{|R_j|}\int_{R_j} f$. C'est possible à cause de (8). Et

$$\begin{aligned} ||f-g||_1 &= \sum_j \int_{R_j} |f-g| \leq \sum_j \int_{R_j} |f-m_j| + \sum_j \int_{R_j} |m_j - \alpha_j| \\ &\leq C \sum_j \int_{R_j} \operatorname{diam}(R_j)|Df| + \eta \sum_j |R_j| \\ &\leq C\eta \sum_j \int_{R_j} |Df| + \eta|Q| \leq C\eta M + \eta|Q| \end{aligned} \tag{10}$$

par Poincaré sur chaque R_j; l'homogénéité joue pour nous, et c'est normal puisque f est un peu régulière. Bref, quand on prend η assez petit, le membre de droite est plus petit que ε; on en déduit le théorème. □

Remarque: il y a plein de variantes de ce genre d'argument; je ne suis pas sûr d'avoir choisi la plus simple. L'argument marche encore pour un domaine connexe assez général, en travaillant un peu plus la géométrie. Mais il ne faut pas permettre un domaine qui soit trop pincé par endroits, pour pouvoir contrôler, y compris localement, la norme L^1 à partir de $\int |Df|$ et de l'intégrale de f.

Variante du théorème: prenons Ω un ouvert borné, et soit A l'ensemble des fonctions caractéristiques d'ensembles F dont le périmètre dans Ω est au plus M. A nouveau A est compact dans L^1.

On peut même prendre les fonctions $f \in BV$ telles que $|f(x)| \leq M$ presque-partout sur Ω, et $\int_\Omega |Df| \leq M$, avec la même démonstration.

Démonstration. D'abord, A est encore complet: si f est la limite dans L^1 d'une suite $\{f_k\}$ dans A, on peut extraire une sous-suite qui converge presque partout vers f, et donc $f(x) \in \{0,1\}$ presque-partout. Le fait que $f \in BV(\Omega)$ est encore dû au théorème de semicontinuité.

Ensuite on fait comme plus haut. Soit $\varepsilon > 0$. On se donne $\eta > 0$ (à choisir bientôt), et on essaie de recouvrir Ω par des petits cubes de diamètre au plus η presque disjoints. On en met autant qu'on peut, mais il reste une partie V de Ω. Malgré tout, $|\Omega \setminus V|$ tend vers 0 quand η tend vers 0 (par convergence dominée, pour être un peu paresseux). A part ça, on utilise la même formule (9), mais on peut se contenter des cubes choisis, et (pour ce qui est des fonctions caractéristiques) de coefficients $\alpha_j \in [0,1]$. Au lieu de

(10), on a

$$(11)\qquad ||f-g||_1 \le \sum_j \int_{R_j} |f-g| + |\Omega\setminus V| \le C\eta M + \eta|Q| + |\Omega\setminus V|$$

par le même calcul que plus haut, et on conclut pareil.

Passons à un exemple d'application. On pense à un problème du type suivant. On se donne un compact $K\subset\mathbb{R}^n$, et un ouvert borné (c'est plus simple) $V\supset K$. On cherche un ensemble F tel que $K\subset F\subset V$, et qui minimise le périmètre. Bien sûr, on pourrait préférer minimiser $\mathcal{H}^{n-1}(\partial F)$, mais c'est un peu plus compliqué, et il n'est pas certain que ce soit plus naturel de toute façon. Voir plus bas.

Corollaire. *On se donne un ensemble T borné de périmètre fini dans $\mathbb{R}^n$, et aussi un domaine $\Omega\subset\mathbb{R}^n$ borné. On note $\mathcal{F}$ la classe des ensembles $F\subset\mathbb{R}^n$, de périmètre fini, et tels que $E=T$ hors de Ω. Alors il existe $E\in\mathcal{F}$ qui minimise le périmètre $P(E)$ dans la classe $\mathcal{F}$.*

Dans l'exemple précédent, on choisirait Z de périmètre fini, tel que $K\subset Z\subset V$, et on choisirait $\Omega=V\setminus K$.

Retour à notre énoncé. Il s'agit d'énoncer un résultat avec des contraintes de type Dirichlet au bord de Ω, sans avoir à parler de traces ni avoir à supposer que Ω est régulier.

Soit R assez grand pour que $T\subset B(0,R)$ et $\Omega\subset B(0,R)$ aussi. Notons M l'infimum du périmètre dans la classe $\mathcal{F}$, et soit $\{F_k\}$ une suite minimisante.

Ainsi, le périmètre de F_k reste borné, donc par compacité il existe une sous-suite (on la notera encore $\{F_k\}$ dont les fonctions caractérstiques convergent dans L^1 vers une limite qui est dans l'ensemble A du second résultat de compacité, c.-à.-d. la fonction caractéristique de E, de périmètre fini.

Noter que $E\in\mathcal{F}$, par exemple parceque l'on peut extraire une sous-suite qui converge vers $\mathbf{1}_E$ presque-partout, et puisque $\mathbf{1}_{F_k}(x)=1$ pour $x\in T\setminus\Omega$ pour tout k.

Finalement, $P(E)\le\liminf_{k\to+\infty}P(F_k)\le M$ par le théorème de semi-continuité, ce qui prouve le corollaire. □

Remarque. C'est un peu trop facile. On a choisi le problème de surface minimale (ici, la surface de la frontière) à peu près le plus facile à résoudre, en particulier parce que nos surfaces sont des frontières (en un sens correct) d'ensembles. Ceci donne de la compacité souvent bien utile.

On devrait continuer notre étude en voyant quel genre de surface minimale on obtient. Il se trouve que c'est possible, et qu'on trouve des ensembles pour lesquels la mesure dérivée $D\mathbf{1}_E$ est effectivement bien portée par ∂B (à l'intérieur de Ω seulement; ailleurs on ne contrôle pas si bien les choses que ça). De sorte qu'en fait E minimise aussi $\mathcal{H}^{n-1}(\partial E)$ localement dans Ω.

On va maintenant s'intéresser à la régularité (faible) des ensembles de périmètre fini dans $\mathbb{R}^n$. Soit donc $E \subset \mathbb{R}^n$ un ensemble de périmètre fini, et notons pour simplifier $f = \mathbf{1}_E$ sa fonction caractéristique. Puis soit $\partial^* E$ l'ensemble borélien des points $x \in \mathbb{R}^n$ tels que

$$\int_{B(x,r)} |Df| > 0 \text{ pour tout } r > 0 \tag{12}$$

et

$$\nu(x) = \lim_{r \to 0} \frac{\int_{B(x,r)} Df}{\int_{B(x,r)} |Df|} \tag{13}$$

existe, avec de plus

$$|\nu(x)| = 1. \tag{14}$$

On appelle $\partial^* E$ la frontière réduite de E. Les résultats suivants sont dûs en grande partie à E. De Giorgi.

Théorème 1. *Gardons $E \subset \mathbb{R}^n$ de périmètre fini, $f = \mathbf{1}_E$, et $\partial^* E$ comme ci-dessus. Alors*

$$|Df|\text{-}\textit{presque tout } x \in \mathbb{R}^n \textit{ est dans } \partial^* E, \tag{15}$$

$$\partial^* E \subset \partial E, \tag{16}$$

$$|Df| = \mathcal{H}^{n-1}_{|\partial^* E}\,, \tag{17}$$

$$\partial^* E \textit{ est rectifiable.} \tag{18}$$

Théorème 2. *Supposons de plus que* $0 \in \partial^* E$ *et que* $\nu(x) = e_1$. *Posons, pour* $r > 0$ *et* $\varepsilon > 0$, $T^+_{r,\varepsilon} = \{x \in B(0,r)\,;\, x_1 \geq \varepsilon r\}$, $T^-_{r,\varepsilon} = \{x \in B(0,r)\,;\, x_1 \leq -\varepsilon r\}$, *et* $S_{r,\varepsilon} = B(0,r) \setminus (T^+_{r,\varepsilon} \cup T^-_{r,\varepsilon}) = \{x \in B(0,r)\,;\, |x_1| < \varepsilon r\}$. *Alors, pour tout* $\varepsilon > 0$,

$$\lim_{r \to 0} r^{-n} |E \cap B(0,r) \cap T^+_{r,\varepsilon}| = 0, \tag{19}$$

$$\lim_{r \to 0} r^{-n} |B(0,r) \cap T^-_{r,\varepsilon} \setminus E| = 0, \tag{20}$$

et pour la mesure $|Df|$,

$$\lim_{r \to 0} r^{-n+1} \int_{B(0,r) \cap T^{\pm}_{r,\varepsilon}} |Df| = 0, \tag{21}$$

$$\lim_{r \to 0} r^{-n+1} \int_{B(0,r) \cap S_{r,\varepsilon}} |Df| = \omega_{n-1}, \tag{22}$$

où ω_{n-1} *est la mesure de Hausdorff du disque unité de* $\mathbb{R}^{n-1}$.

On ne va pas tout démontrer, juste faire les remarques faciles. Noter qu'il y a un peu de redondance dans ces énoncés.

Noter que nos définitions ne changent pas quand on remplace E par un ensemble équivalent. C'est bien.

Dans le cas d'un domaine régulier, on retrouve $\partial^* E = \partial E$ et $Df = \nu |Df| = \nu \mathcal{H}^{n-1}_{|\partial^* E}$, où ν est la normale extérieure. Les résultats disent que c'est un peu pareil dans le cas général.

Si E est un cube, les coins et les arêtes ne sont pas dans $\partial^* E$ (la condition dans (13) y est violée et les conditions du théorème 2 sont fausses). Pareil pour un cusp.

Rien n'empêche d'écrire $Df = \nu |Df|$ par Radon-Nikodym, avec $|\nu(x)| = 1$ pour $|Df|$-presque tout x. Si x est dans le support de $|Df|$, il vérifie (12). Le fait qu'on ait (13) pour $|Df|$-presque tout x est le théorème de différentiabilité de Lebesgue.

Ainsi, $|Df|$-presque tout x est dans $\partial^* E$, comme annoncé en (15). Le fait que $\partial^* E \subset \partial E$ est facile, soit directement, soit parce qu'on a déjà vu que support de $|Df|$ est contenu dans ∂E. Notons d'ailleurs qu'il est assez facile de modifier E sur un ensemble de mesure nulle, de façon que E contienne tous les points de $x \in \mathbb{R}^n$ tels que $|B(x,r) \setminus E| = 0$ pour un $r > 0$, et ne contienne aucun point x tel que $|B(x,r) \cap E| = 0$ pour un $r > 0$. Alors $\partial E \subset \overline{\partial^* E}$.

Evidemment il est facile de produire des exemples où $\partial^* E$ est bien plus petit que ∂E. On peut aussi de débrouiller pour que ∂E ait une partie non rectifiable, même dans le cas plus propre ci-dessus où $\partial E \subset \overline{\partial^* E}$. En fait, construire $\partial^* E$, par exemple avec des petits carrés, et se débrouiller pour que $\overline{\partial^* E}$ ait un morceau non rectifiable, par exemple un flocon de neige ou morceau de Cantor.

Après, on se met dans les conditions du théorème 2, et on essaie d'obtenir des informations sur E dans les petites boules. Noter que (21) ressemble à de la rectifiabilité, si l'on a (17).

Donc, on ne démontre pas le reste (ça prend un certain temps dans Giusti, même si je le soupçonne de faire des petits détours). Quand même, si on oublie qu'on doit avoir des estimations un peu précises, c'est plus facile à faire. Supposons par exemple que pour une boule $B(x,r)$, on ait l'égalité exacte $\int_{B(x,r)} Df = \nu \int_{B(x,r)} |Df|$ pour le vecteur unitaire ν. Par exemple, $\nu = e_1$. Ca veut dire que les dérivées partielles $\frac{\partial f}{\partial x_i}$, $i \neq 1$, sont nulles, donc par exemple que si φ est une petite fonction bosse à support dans $B(0, r/2)$, $\int_E f(x)dx = \int_E f(x+h)dx$ pour tout petit vecteur vertical h. En fait, dans $B(0, r/2)$, f ne dépend que de la variable x_1. En plus, c'est une fonction caractéristique, et ceci laisse suffisament peu de choix: f ressemble à une fonction caractéristique de demi-espace. Si en plus on sait qu'on a le même type de description pour des boules plus petites et que $\int_{B(x,r)} |Df| \neq 0$, le plan vertical à la frontière passe près de l'origine.

Sauf erreur de ma part, quand on a ceci, il n'est pas trop dur de revenir en arrière et de vérifier deux choses assez logiques, mais que je n'ai pas justifiées jusqu'ici. D'abord, pour un ensemble E de périmètre fini, on a que $\mathcal{H}^d(\partial E) \geq \mathcal{H}^d(\partial^* E) \geq Per(E)$, où la première partie vient de (16), et la seconde de (22) (et d'un petit argument de densité).

De plus, si E est tel que $\mathcal{H}^d(\partial E) < +\infty$, alors E est de périmètre fini. Il me semble que pour ceci, on doit pouvoir approximer E par des domaines dont la frontère est contenue dans une union de sphères (en recouvrant ∂E), puis passer à la limite en disant que ces approximations ont des périmètres uniformément bornés. J'ai pas vérifié, donc prenez avec un grain de sel. Ce second résultat permet ensuite de dire que $Per(E) \leq \mathcal{H}^d(\partial E)$ en utilisant le précédent.

Disons maintenant quelques mots sur SBV (special bounded variation) et les résultats d'existence pour la fonctionnelle de Mumford-Shah. Toujours avec un minimum de démonstration. Cette fois on renvoie au livre d'Ambrosio, Fusco et Pallara.

D'abord, il y a une description de Df quand $f \in BV$. Une première chose est le découpage de $|Df|$ en une partie absolument continue par rapport à la mesure de Lebesgue, puis une partie singulière. Si l'on garde la notation $Df = \nu|Df|$, cela donne une décomposition $Df = \nabla f\, dx + D_s f$, où la partie absolument continue $\nabla f\, dx$ se calcule effectivement comme différentielle presque-partout (en gros, par différentiabilité de Lebesgue). Dans la partie singulière, on arrive encore à couper en deux. La première partie est portée par un ensemble rectifiable K^*, et correspond à un saut de f. C'est ce que l'on a pour la fonction caractéristique de E de périmètre fini, où l'ensemble K^* est la frontière réduite, le saut est 1, et a lieu dans la direction orthogonale ν. Dans le cas de BV, on autorise aussi un saut $j(x)$ non nécessairement égal à 1, et qu'on peut calculer comme

$$j(x) = \lim_{r \to 0} [m_+(r) - m_-(r)],$$

où

$$m_\pm(r) = \frac{1}{|B_\pm(r)|} \int_{B_\pm(r)} f \quad \text{et} \quad B_{\pm(r)} = \big\{ y \in B(x,r)\,;\, \pm\langle y - x, \nu(x)\rangle \geq 0 \big\}.$$

Cette première partie de la partie singulière de Df est donc du type suivant:

$$D_s(x) = \mathbf{1}_{K^*}(x)\, j(x)\, \nu(x)\, d\mathcal{H}^{n-1}(x).$$

Et il reste un dernier morceau de mesure singulière, dont on ne dit pas grand-chose, et qu'on appellera partie cantorienne à cause de l'exemple de la fonction de Lebesgue (escalier du diable).

Un espace amusant est l'espace SBV (avec S pour spécial) des fonctions de BV pour lesquelles la partie cantorienne est nulle. Sauf erreur, cet espace a été introduit et étudié par Ambrosio. Il est aimablement stable par limites, ce qui lui donne d'agréables propriétés de compacité.

Suite possible: défnition de la fonctionnelle de Mumford-Shah dans le cas standard, puis version faible, puis description vague de l'approche par SBV pour l'existence de segmentations minimisantes. Sans doute n'aurai-je pas le temps d'écrire des notes là-dessus!!

18. PETITS COMPLEMENTS: LUSIN ET EGOROFF

Ce chapitre devrait normalement se situer plus tôt. Il s'agit de dire deux mots du théorème de Lusin et de son application à l'existence de grands ensembles sur lesquels une fonction Lipschitzienne donnée coïncide avec des fonctions de classe C^1.

Théorème (Lusin). *On se donne un espace métrique localement compact E (par exemple, $E = \mathbb{R}^n$), une mesure finie μ sur E, muni des boréliens, et une fonction f définie sur E, borélienne, et à valeurs dans $\mathbb{R}^m$, ou $\mathbb{C}^m$, ou $\overline{\mathbb{R}}$. Mais si elle est à valeurs dans $\overline{\mathbb{R}}$, on la suppose finie presque-partout. Alors pour tout $\varepsilon > 0$, il existe g continue (et finie partout), telle que $\mu(\{x \in E\,;\, f(x) \neq g(x)\}) \leq \varepsilon$.*

On commence par quelques réductions faciles. On peut supposer que f est finie partout (la remplacer par 0 sur un ensemble de mesure nulle ne change pas la conclusion), puis à valeurs réelles (appliquer l'argument à chaque coordonnée et avec ε/n, puis prendre l'union des mauvais ensembles), puis même à valeurs dans $]-M, M[$ (choisir M assez grand pour que $\mu(\{x \in E\,;\, |f(x)| \geq M\}) \leq \varepsilon/2$), et enfin à valeurs dans $[0, 1[$ (considérer $[M + f(x)]/2M$.)

Par un argument standard de théorie de la mesure, on peut écrire $f = \sum_{k\geq 1} 2^{-k}\, \mathbf{1}_{E_k}$, avec des ensembles E_k boréliens. Les sommes partielles sont les approximations standard de f par des fonctions à valeurs dans $2^{-k}\mathbb{N}$, et une manière d'écrire la somme directement est de dire que pour tout x, $f(x) = \sum_{k\geq 1} 2^{-k}\, \mathbf{1}_{E_k}$ est l'écriture standard de $f(x)$ en base 2. Et sauf erreur de calcul, $\mathbf{1}_{E_k}(x) = e(2^k f(x)) - 2e(2^{k-1} f(x))$, où $e(y)$ est la partie entière de y; d'où la mesurabilité.

Maintenant, la mesure μ étant régulière, on sait que pour tout k, on peut trouver un compact K_k et un ouvert V_k tels que $K_k \subset E_k \subset V_k$ et $\mu(V_k \backslash K_k) < \varepsilon_k$, où l'on peut choisir $\varepsilon_k > 0$ très petit, par exemple, $\varepsilon_k = 2^{-k}\varepsilon$.

Par un théorème de topologie (Urysson, je crois), il existe une fonction continue g_k telle que $g_k(x) = 1$ sur K_k, $g_k(x) = 0$ hors de V_k, et $0 \leq g_k(x) \leq 1$ sur $V_k \setminus K_k$. D'ailleurs, ici l'espace est métrique et il est facile de donner une formule pour $g_k(x)$ en fonction de la distance de x au complémentaire de V_k. La série $\sum_k 2^{-k} g_k$ converge normalement, et sa somme g est continue. Il reste à voir que $g(x) = f(x)$ hors d'un ensemble Z tel que $\mu(Z) \leq \varepsilon$.

Mais on peut prendre $Z = \cup_k [V_k \setminus K_k]$, et utiliser le fait que $\mu(V_k \setminus K_k) < \varepsilon_k$. $\square$

Le théorème a un corollaire facile: la fonction f est une limite simple presque-partout de fonctions continues. Mais à mon sens c'est moins impressionnant.

Avant le second corollaire, qui concerne les fonctions lipschitziennes, énoncons le théorème d'Egoroff.

Théorème d'Egoroff. *Soit* $(E, \mathcal{A}, \mu)$ *un espace de probabilité, et soit* $\{\varphi_n\}$ *une suite de fonctions mesurables, disons de* E *dans* $\mathbb{R}$*, telle que* $\lim_{n\to+\infty} \varphi_n(x) = 0$ *pour* μ*-preque tout* $x \in E$*. Alors, pour tout* $\varepsilon > 0$*, il existe* $F \subset E$ *mesurable tel que* $\mu(E \setminus F) \leq \varepsilon$ *et*

$$\lim_{n\to+\infty} \varphi_n(x) = 0 \text{ uniformément sur } F. \tag{1}$$

Donc on doit jeter un petit ensemble pour rendre la limite uniforme. C'est sans doute à cause de cette ressemblance avec Lusin ci-dessus qu'on confond souvent.

J'ai décidé de prendre une limite nulle pour simplifier, mais bien sûr si la limite avait été une fonction φ, rien de plus simple que d'appliquer l'énoncé ci-dessus aux fonctions $\varphi_n - \varphi$.

Quitte à remplacer les φ_n par 0 sur un ensemble de mesure nulle, et ajouter cet ensemble à $E \setminus F$, on peut supposer que $\lim_{n\to+\infty} \varphi_n(x) = 0$ partout. Posons $\alpha_k = 2^{-k}$. Fixons k et notons

$$A_{N,k} = \big\{x \in E\,;\, |\varphi_n(x)| \geq \alpha_k \text{ pour au moins un } n \geq N\big\}$$

pour $N \geq 1$. Puisque $\lim_{n\to+\infty} \varphi_n(x) = 0$ partout, l'intersection (visiblement décroissante) des $A_{N,k}$ est vide. Donc il existe $N(k)$ tel que $\mu(A_{N(k),k}) \leq \varepsilon 2^{-k-1}$.

Ensuite posons $Z = \cup_k A_{N(k),k}$, et $F = E \setminus Z$. Alors $\mu(E \setminus F) = \mu(Z) \leq \sum \mu(A_{N(k),k}) \leq \varepsilon$, et il ne reste plus qu'à vérifier (1).

On doit montrer que pour tout $\alpha > 0$, il existe N tel que $|\varphi_n(x)| \leq \alpha$ pour $x \in F$ et $n \geq N$. On choisit k tel que $\alpha_k < \alpha$ et $N = N(k)$. Alors si $x \in F$, x n'est pas dans $A_{N(k),k} \subset Z$, donc $\varphi_n(x) \leq \alpha_k < \alpha$ pour tout $n \geq N$, come souhaité. $\square$

Passons au corollaire sur les fonctions lipschitziennes.

Corollaire. *Soient $K \subset \mathbb{R}^n$ un compact et $f : K \to \mathbb{R}^m$ une fonction Lipschitzienne. Alors pour tout $\varepsilon > 0$ il existe $g : \mathbb{R} \to \mathbb{R}$, de classe C^1, telle que $\{x \in K\,;\, f(x) \neq g(x)\}$ soit de mesure de Lebesgue au plus ε.*

Evidemment, par le théorème de Whitney on peut supposer que f est définie sur $\mathbb{R}^n$ tout entier.

Ensuite, on peut se contenter de démontrer le théorème pour $f : K \to \mathbb{R}$, puisqu'ensuite il suffit d'appliquer le résultat coordonnée par coordonnée, et avec ε/m, puis de prendre l'union des ensembles exceptionnels.

C'est bien possible qu'on puisse remplacer le compact par $\mathbb{R}^n$ tout entier. Je n'ai pas vraiment essayé ni vérifié dans les livres.

Pour la démonstration, on applique le théorème de Lusin à la différentielle de f, que l'on note F (et qu'on confondra allègrement avec un gradient), et qui existe presque-partout par le théorème de Rademacher. C'est une fonction borélienne (si on écrit ce qu'est une différentielle), mais de toute façon elle est égale presque-partout à la dérivée au sens des distributions (on l'a démontré au chapitre sur la différentiation des fonctions de $W^{1,p}$), donc on pourrait la modifier sur un ensemble de mesure nulle pour la rendre borélienne.

On trouve donc une fonction continue G (à valeurs dans les applications linéaires de $\mathbb{R}^n$ dans $\mathbb{R}$, qu'on assimile à $\mathbb{R}^n$, mais Lusin s'applique quand même), qui coïncide avec $F = Df$ sur un ensemble E_1 tel que $|K \setminus E_1| \leq \varepsilon/3$. (On met le fait que $Df(x)$ existe pour $x \in E_1$ dans la définition de E_1, comme cela on n'aura pas d'ennuis).

En plus, on va avoir besoin d'une précaution supplémentaire un peu désagréable, qui va nous amener à appliquer également le théorème d'Egoroff. On sait que pour $x \in E_1$, $Df(x)$ existe, ce qui signife que si l'on pose

$$(2) \quad \varphi_n(x) = \sup\{|y-x|^{-1}|f(y) - f(x) - Df(x).(y-x)|\,;\, y \in B(x, 2^{-n})\}$$

alors φ_n tend vers 0 quand n tend vers 0. Par Egoroff, il existe un nouveau borélien E_2 contenu dans E_1, tel que $|E_1 \setminus E_2| \leq \varepsilon/3$, et pour lequel $\lim_{n\to+\infty} \varphi_n(x) = 0$ uniformément sur E_2, comme en (1). On peut même le prendre compact (parce que la mesure est régulière, donc ceci nous coûte juste un peu de mesure). Autrement dit, pour tout $\alpha > 0$, il existe $N \geq 0$ tel que $\varphi_n(x) \leq \alpha$ pour $x \in E_2$ et $n \geq N$. Ou encore, tel que

$$(3) \quad |y-x|^{-1}|f(y)-f(x)-Df(x).(y-x)| \leq \alpha \text{ pour } x \in E_2 \text{ et } y \in B(x, 2^{-N}).$$

Traduisons encore ceci. Posons

$$\varepsilon_1(r) = \sup_{x \in E_2} \sup_{y \in B(x,r)} |y-x|^{-1} |f(y) - f(x) - Df(x).(y-x)|$$

pour $r > 0$. C'est une fonction croissante de r, et (3) nous dit que $\lim_{r \to 0^+} \varepsilon_1(r) = 0$. Et en traduisant brutalement la définition de ε_1,

$$|f(y) - f(x) - Df(x).(y-x)| \leq |y-x| \varepsilon_1(|x-y|) \tag{4}$$

pour $x \in E_2$ et $y \in \mathbb{R}^n$.

Maintenant, comme E est compact et $Df = F = G$ y est continue, elle y est aussi uniformément continue, de sorte qu'il existe un module de continuité, à savoir une fonction ε_2 croissante, tendant aussi vers 0, telle que

$$|Df(x) - Df(y)| \leq \varepsilon_2(|x-y|) \text{ pour } x, y \in E_2. \tag{5}$$

On peut d'ailleurs la changer, si besoin est, pour qu'elle vérifie la condition (26.5) dans le second théorème de Whitney (toutes les références qui suivent concernent le paragraphe 12.d): la remplacer par une fonction comme en (29.5), qui tend encore vers 0.

Choisissons pour finir un ensemble compact $E \subset E_2$, tel que $|E_2 \setminus E| \leq \varepsilon/3$ et donc $|K \setminus E| \leq \varepsilon$. On applique le second théorème d'extension de Withney à la restriction de f à E. Le jet de Whitney d'ordre 1 est donné par la fonction G (qui coincide avec $F = Df$ sur E), et les majorations dans les hypothèses sont satisfaites avec $\varepsilon(r) = \varepsilon_1(r) + \varepsilon_2(r)$, grâce à (4) et (5).

On trouve une fonction g de classe C^1, qui étend la restriction de f à E, comme souhaité pour le corollaire. On a même gagné un module de continuité pour Dg, mais d'un autre côté ce module, qui vient de Lusin, est sans doute assez mauvais, surtout quand ε est petit. $\square$

19. PROPRIETES QUANTIFIEES DE REGULARITE

On va commencer par un petit retour à la régularité des fonctions lipschitziennes. On cherche des propriétés mieux quantifiées que la différentiabilité presque-partout.

On va donner un énoncé dans ce sens (parmi beaucoup), puis on en verra un bout de démonstration et des petites variantes.

On se donne une fonction f, définie sur $\mathbb{R}^n$ (mais ce qu'on va faire pourrait se localiser). Disons qu'on s'intéresse à l'approximation de f par des fonctions affines. Noter que la différentiabilité presque-partout est un peu de ce type, mais on va essayer de mesurer plus précisément.

On se donne un exposant annexe $q \in [1, +\infty]$, et on mesure la proximité de f aux fonctions affines en posant

$$\gamma_q(x,r) = \frac{1}{r} \inf_{a \in \mathcal{A}} \left\{ \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - a(y)|^q dy \right\}^{1/q} \tag{1}$$

pour $x \in \mathbb{R}^n$ et $r > 0$, où l'on a noté $\mathcal{A}$ la classe des fonctions affines. Pour $q = +\infty$, on fait la modification standard, ce qui donne

$$\gamma_\infty(x,r) = \inf_{a \in \mathcal{A}} \sup_{y \in B(x,r)} \left\{ \frac{1}{r} |f(y) - a(y)| \right\}. \tag{2}$$

Noter que $\gamma_q(x,r)$ est une fonction croissante de q, par Hölder, et que $\gamma_q(x,r) \leq ||f||_{lip}$ à peu près trivialement (en essayant $a(y) = f(x)$).

Mais la bonne surprise est que, bien qu'on ne puisse pas s'attendre à ce que $\gamma_q(x,r)$ soit plus petit que C a priori, il se trouve que $\gamma_q(x,r)$ est souvent bien plus petit, au point que des intégrales du type $\int_x \int_r \gamma_q(x,r)^2 \frac{dx\,dr}{r}$ convergent, comme dans les énoncés suivant. On donne d'abord l'énoncé pour l'espace de Sobolev $\mathcal{H}^1$, puis pour les fonctions Lipschitziennes.

Théorème 1 [Dorronsoro]. *Pour $1 \leq q < \frac{2n}{n-2}$ si $n \geq 2$, et pour tout $q \in [1, +\infty]$ si $n = 2$, il existe $C_q > 0$ telle que pour $f \in W^{1,2}(\mathbb{R}^n)$,*

$$\int_0^{+\infty} \int_{\mathbb{R}^n} \gamma_q(x,r)^2 \, \frac{dx\,dr}{r} \leq C_q ||\nabla f||_2^2. \tag{3}$$

Corollaire 1. *Pour $1 \leq q < \frac{2n}{n-2}$ si $n \geq 2$, et pour tout $q \in [1, +\infty]$ si $n = 2$, il existe $C_q > 0$ telle que pour $f \in L^1_{loc}(\mathbb{R}^n)$ telle que $\nabla f \in BMO$, et tout choix de $x \in \mathbb{R}^n$ et $r > 0$,*

$$\int_{t=0}^{r} \int_{y \in B(x,r)} \gamma_q(y,t)^2 \, \frac{dy\,dt}{t} \leq C_q ||\nabla f||_{BMO}^2 \, r^n. \tag{4}$$

Quelques commentaires avant de parler des démonstrations.

Dans le second énoncé, on peur donc en particulier prendre f lipschitzienne, puisque dans ce cas l'on a vu que $f \in W^1_{loc}(\mathbb{R}^n)$, avec un gradient ∇f borné, et même $||\nabla f||_{BMO} \leq 2||f||_{lip}$.

Ces résultats sont en fait des résultats d'orthogonalité, ce qui explique que le centre de la démonstration passe par une transformation de Fourier (et je crois qu'il serait bien difficile de faire autrement).

La contrainte sur q est malheureusement nécessaire: il y a des contre-exemples à pointes dont l'existence est liée au fait que les inégalités de Sobolev ne sont pas vraies pour tous les exposants. C'est dommage, on aurait préféré prendre $q = +\infty$ même pour $n \geq 2$. Mais on peut montrer que, dans le cas du corollaire ou la fonction f est lipschitzienne, $\gamma_\infty(x, r) \leq C\gamma_q(x, 2r)^{\frac{q}{d+q}}$ (avec C qui dépend de la constante de Lipschitz de f, et pour une démonstration plus facile, en ne mettant dans la définition de γ_q que des fonctions affines qui ont une norme de Lipschitz inférieure à $C||\nabla f||_\infty$). Donc on aura quand même un certain contrôle sur les $\gamma_\infty(x, r)$. Voir la preuve de (6) dans le prochain chapitre.

Ces résultats ne donnent pas directement la différentiabillité presque partout, parce que la fonction affine $a_{x,r}$ qui approxime bien f dans $B(x, r)$ peut dépendre de r, mais par contre ils sont agréablement quantitatifs.

La condition du corollaire dit que la mesure $\gamma_q(y, t)^2 \,\frac{dy\,dt}{t}$ est une mesure de Carleson sur le demi-espace supérieur. De telles conditions apparaîssent naturellement, pour cause de bonne homogénéité, et sont souvent utiles, typiquement en connexion avec l'espace BMO. D'ailleurs, sauf erreur de ma part, la condition (4) (avec un second membre Cr^d) est une caractérisation de la condition $\nabla f \in BMO$.

On va pouvoir démontrer ces résultats quand $q = 2$. Pour $q > 2$, la démonstration se complique, il faut faire appel à des décompositions de Littlewood-Paley, et peut-être même à des opérateurs d'intégrales singulières à valeurs vectorielles, mais bien sûr ça n'est sensé être "que de la technique". Bien content quand même de ne pas avoir à le faire ici. Par ailleurs, le résultat de Dorronsoro est encore plus général, et sert de caractérisation de divers espaces fonctionnels.

On va commencer par démontrer le corollaire à partir du théorème, par un argument de localisation standard (à x et r fixés, c'est logique que les choses ne dépendent que des valeurs de f dans $B(x, 2r)$).

Donc on se donne f lipschitzienne, ou même à dérivée dans BMO, et aussi $x \in \mathbb{R}^n$ et $r > 0$, et on va appliquer le théorème 1 à une fonction auxiliaire.

D'abord, on observe que l'on peut ajouter une fonction affine à f, sans rien changer aux γ_q ni au membre de gauche de (4). On peut donc supposer que $m_{B(x,4r)}f = 0$ et même, si $\nabla f \in BMO$, que $m_{B(x,4r)}\nabla f = 0$.

On choisit une gentille fonction de coupure φ, telle que $\varphi(y) = 1$ sur $B(x, 2r)$, $\varphi(y) = 0$ hors de $B(x, 4r)$, et $0 \leq \varphi(y) \leq 1$ et $|\nabla\varphi(y)| \leq 2r^{-1}$ partout, puis on pose $g(y) = f(y)\varphi(y)$. Commençons par observer que les deux fonctions $\gamma_q(y, t)$, pour f et pour g, coïncident sur la région où $y \in B(x, r)$ et $0 < t < r$, de sorte qu'il suffira de montrer que l'analogue de (4) pour g est vrai. On élargit carrément le domaine d'intégration à $\mathbb{R}^n\times]0, +\infty[$, et on utilise le théorème; on voit qu'il suffira de vérifier que $g \in W^{1,2}(\mathbb{R}^n)$, avec

$$(4) \qquad ||\nabla g||_2^2 \leq C||\nabla f||_\infty^2\, r^n \quad \text{ou} \quad ||\nabla g||_2^2 \leq C||\nabla f||_{BMO}^2\, r^n,$$

suivant le cas considéré. Le cas où f est lipschitzienne est plus facile, donc parlons seulement du cas où $\nabla f \in BMO$. Noter que $f \in W^{1,2}_{loc}$, avec $\int_{B(x,4r)} |\nabla f|^2 \leq C||\nabla f||_{BMO}^2\, r^n$, parce que $\nabla f \in BMO$ et sa moyenne sur $B(x, 4r)$ est nulle. Après, $\nabla(\varphi f) = \varphi\nabla f + f\nabla\varphi$ (faire un peu de théorie élémentaire des distributions si f n'est pas supposée régulière, en observant qu'on peut prendre φ de classe C^∞). Le premier terme donne une contribution à $\int |\nabla g|^2$ inférieure à $\int_{B(x,4r)} |\nabla f|^2 \leq C||\nabla f||_{BMO}^2\, r^n$, comme nous venons de dire. Le second terme donne au plus

$$r^{-2}\int_{B(x,4r)} |f|^2 = r^{-2}\int_{B(x,4r)} |f - m_{B(x,4r)}f|^2 \leq C\int_{B(x,4r)} |\nabla f|^2 \leq C||\nabla f||_{BMO}^2\, r^n$$

par Poincaré. Donc on a (4) et le corollaire se déduit du théorème.

Maintenant on va s'occuper du théorème, mais seulement dans le cas où $q = 2$. On commence par un résultat de la même veine, mais plus facile, et qui donnera une idée de la démonstration. On se place en dimension 1 pour simplifier, mais la démonstration du théorème 1 contiendra la généralisation à $\mathbb{R}^n$.

Théorème 2. *Il existe $c > 0$ telle que pour $f \in W^{1,2}(\mathbb{R})$,*

$$(5) \qquad \int_{x\in\mathbb{R}}\int_{t>0} \frac{|f(x+t) + f(x-t) - 2f(x)|^2\, dxdt}{t^3} = c\int_{\mathbb{R}} |\nabla f|^2.$$

A nouveau, noter qu'on s'attend simplement à ce que, pour f Lipschitzienne, le rapport $t^{-2}|f(x+t) + f(x-t) - 2f(x)|^2$ soit borné, alors que (5) dit qu'il est bien plus petit en moyenne, puisqu'il est intégrable contre la mesure localement infinie $\frac{dxdt}{t}$.

Il y a d'autres critères du même genre pour mesurer la régularité (plutôt faible) des fonctions (voir le livre de Stein (Singular Integrals and Differentiability Properties of Functions). L'idée de mesurer ceci en utilisant une fonction définie sur le demi-espace supérieur $\mathbb{R}^n\times]0,+\infty[$ (à voir aussi comme l'ensemble des boules), et aussi classique et fructueuse.

Pour des régularités plus faibles, on pourraît se contenter d'utiliser les différences simples $|f(x+t)-f(x)|$, mais pas ici.

Pour la démonstration, on commence par fixer t, et à calculer

$$I(t)=\int_{\mathbb{R}}|f(x+t)+f(x-t)-2f(x)|^2dx$$

par Plancherel. On trouve que

$$(6)\quad I(t)=\int_{\mathbb{R}}\left|\widehat{f}(\xi)[2-e^{2i\pi t\xi}-e^{-2i\pi t\xi}]\right|^2d\xi=4\int_{\mathbb{R}}|\widehat{f}(\xi)|^2|\cos(2\pi t\xi)-1|^2d\xi$$

Ensuite, on note J le membre de gauche de (5), et on intègre (6) contre $t^{-3}dt$ pour obtenir

$$(7)\qquad J=\int I(t)t^{-3}dt=\int_\xi\int_t|\xi\widehat{f}(\xi)|^2\ \frac{|\cos(2\pi t\xi)-1|^2}{|t\xi|^2}\ \frac{dt\,d\xi}{t}$$

L'intégrale intérieure en dt/t converge, et en plus ne dépend pas de ξ (car dt/t est invariante par dilatations), donc il reste $J=c\int_\xi|\xi\widehat{f}(\xi)|^2d\xi=c\int_{\mathbb{R}}|\nabla f|^2$. $\square$

On retourne au théorème 1. Pour $f\in W^{1,2}(\mathbb{R}^n)$ et $(x,t)\in\mathbb{R}^n\times]0,+\infty[$, on doit trouver une fonction affine $a_{x,t}$. La chose la plus raisonnable serait de prendre la fonction affine dont la moyenne sur $B(x,t)$ est la moyenne de f sur $B(x,t)$, ou, mieux même, une version adoucie du genre $\int_{B(0,1)}\varphi(u)f(x+tu)du$, avec φ régulière et à support dans $B(0,1)$, et pareillement le gradient (constant) de $a_{x,t}$ est une moyenne adoucie de ∇f sur $B(x,t)$. C'est ce qu'il faudrait faire pour le résultat général, mais avec $q=2$ on va pouvoir s'en sortir avec un choix de $a_{x,t}$ moins régulier.

Notons e_j le $j^{\text{iéme}}$ vecteur de base, et pour tout $u=(u_1,\cdots,u_n)\in\mathbb{R}^n$, posons

$$a_{x,t}(x+tu)=f(x)+\sum_{j=1}^n\frac{u_j}{2}[f(x+te_j)-f(x-te_j)].$$

C'est une fonction affine, et on va voir si elle marche. Noter que le gradient coïncide avec celui de f si f est affine. Notons aussi, pour simplifier un argument du chapitre suivant, que la fonction a est lipschitzienne avec une constante $\leq \sqrt{n}||f||_{lip}$ si f est Lipschitzienne, puisque pour chacune des n coordonnées de ∇a, on a $|\dfrac{\partial a}{\partial u_j}| \leq \dfrac{1}{2t}|f(x+te_j) - f(x-te_j)| \leq ||f||_{lip}$.

On fixe $t > 0$ et $u \in B(0,1)$, et on calcule

$$I(u,t) = \int_x |f(x+tu) - a_{x,t}(x+tu)|^2 dx.$$

Par Plancherel, $I(u,t) = \int_\xi |\widehat{f}(\xi)|^2 |b(\xi)|^2 d\xi$, avec

$$b(\xi) = e^{2i\pi tu\cdot\xi} - 1 - \sum_{j=1}^{n} \frac{u_j}{2}[e^{2i\pi t\xi_j} - e^{-2i\pi t\xi_j}]$$

Pour $|t\xi| \geq 1$, on se souviendra juste que $b(\xi) \leq 3$. Sinon, on utilise un développement d'ordre 2 des exponentielles. La valeur en 0 est nulle (pas de contribution de la dernière somme). Pour le terme d'ordre 1, il sort

$$2i\pi tu \cdot \xi - \sum_{j=1}^{n} \frac{u_j}{2}\,[2i\pi t\xi_j + 2i\pi t\xi_j] = 0.$$

Bref, $b(\xi) \leq C|t\xi|^2$, avec une constante universelle C. Et du coup

$$(8) \qquad t^{-2}I(u,t) \leq \int_\xi |\xi|^2|\widehat{f}(\xi)|^2 \frac{|b(\xi)|^2}{t^2|\xi|^2} d\xi \leq C\int_\xi |\xi|^2|\widehat{f}(\xi)|^2 \frac{t^2|\xi|^2}{1+t^4|\xi|^4}\, d\xi$$

(en distinguant deux cas), puis (pour u fixé)

(9)

$$\int_t t^{-2}I(u,t)\,\frac{dt}{t} \leq C\int_\xi |\xi|^2|\widehat{f}(\xi)|^2 \int_t \frac{t^2|\xi|^2}{1+t^4|\xi|^4}\,\frac{dt}{t}\, d\xi \leq C\int_\xi |\xi|^2|\widehat{f}(\xi)|^2\, d\xi \leq C||\nabla f||_2^2$$

puisque l'intégrale intérieure converge et ne dépend pas de ξ. Ceci dit, noter que les constantes C sont uniformément bornées pour $u \in B(0,1)$. Maintenant, si J est le membre de gauche de (3) (avec $q = 2$), il vient

$$\begin{aligned}
J &= \int_{x\in\mathbb{R}^n}\int_{t>0}\Big\{\frac{1}{|B(x,t)|}\int_{B(x,t)} t^{-2}|f(z) - a_{x,t}(z)|^2 dz\Big\}\frac{dxdt}{t} \\
&= c\int_{x\in\mathbb{R}^n}\int_{t>0}\int_{u\in B(0,1)} t^{-2}|f(x+tu) - a_{x,t}(x+tu)|^2\ \frac{dudxdt}{t} \\
&= c\int_{u\in B(0,1)}\Big\{\int_{t>0}\int_{x\in\mathbb{R}^n} t^{-2}|f(x+tu) - a_{x,t}(x+tu)|^2\ \frac{dxdt}{t}\Big\}\, du
\end{aligned}$$

avec $c = |B(0,1)|^{-1}$ et par Fubini. Pour tout u, l'intégrale intérieure est au plus $C||\nabla f||_2^2$ par (9), donc $I \leq C||\nabla f||_2^2$, comme souhaité pour le cas théorème 1 quand $q = 1$. □

On pourrait bien utiliser plus loin la petite remarque suivante. Si f est Lipschitzienne, la fonction affine qui a été utilisée dans la preuve est elle aussi lipschitzienne, avec $||a||_{lip} \leq n||f||_{lip}$. Pour f est Lipschitzienne, on aurait pu définir les nombres $\widetilde{\gamma}_q$ comme les γ_q ci-dessus, mais en spécifiant que dans (1) et (2), on ne regarde que les fonctions affines telles que $||a||_{lip} \leq C(n,q)||f||_{lip}$ (avec une constante $C(n,q)$, qu'on pourrait choisir à l'avance. Par exemple, $C(n,2) = n$.

On a prouvé que les $\widetilde{\gamma}_2$ vérifient aussi l'estimation de Dorronsoro, et en fait $\widetilde{\gamma}_q = \gamma_q$ de toute façon, en tout cas si $C(n,q)$ est choisie assez grand.

Ceci est en fait vrai pour tout q comme dans le théorème 1, pour deux raisons simples La première est que, comme dans notre cas de γ_2, quand on démontre le théorème 1, on choisit en fait des fonctions a qui par construction sont lipschitziennes telles que $||a||_{lip} \leq C(n,q)||f||_{lip}$.

La seconde est que de toute façon, si a minimise ou minimise pratiquement dans la définition (1) ou (2) de $\gamma_q(x,r)$, on peut vérifier qu'elle est forcément lipschitzienne, avec $||a||_{lip} \leq C(n,q)||f||_{lip}$. C'est pas bien difficile: supposons que $||f||_{lip} = M$ mais $||a||_{lip} = AM$ avec un grand A. En particulier, il y a une direction $e \in \partial B(0,1)$ où la dérivée de a est de taille AM. Disons, quitte à remplacer e par $-e$, que la dérivée dans la direction de e est $\geq AM$. Et supposons encore que $a(x) \geq f(x)$ (sinon, on partirait à la place dans la direction $-e$). Alors, dans la boule $B(x + 2er/3, r/3)$, on a encore $a(y) \geq a(x) + rAM/3 \geq f(x) + rAM/3 \geq f(y) + rAM/3 - Ar \geq\geq f(y) + rAM/4$. Et du coup (je prends la cas où $q < +\infty$; le cas où $q = +\infty$ est plus simple)

$$\begin{aligned}
2\gamma_q(x,r) &\geq \frac{1}{r}\Big\{\frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y) - a(y)|^q dy\Big\}^{1/q} \\
&\geq \frac{1}{r}\Big\{\frac{1}{|B(x,r)|}\int_{B(x+2er/3,r/3)} |f - a|^q\Big\}^{1/q} \\
&\geq \frac{AM}{4}\left\{\frac{|B(x+2er/3,r/3)|}{|B(x,r)|}\right\}^{1/q} = \frac{AM}{4\cdot 3^{n/q}}
\end{aligned}$$

C'est bien une contradiction si A est choisi assez grand, puisqu'on sait (en essayant brutalement $a = f(x)$) que $\gamma_q(x,r) \leq ||f||_{lip} = M$.

Continuons avec le cas des fonctions lipschitziennes, en vérifiant que pour tout $M > 0$, il existe $C(n, M)$ telle que si f est lipschitzienne, avec $||f||_{lip} \leq M$, alors

$$\widetilde{\gamma}_\infty(x, r) \leq C(n, M)\, \widetilde{\gamma}_2(x, 2r)^{\frac{2}{n+2}} = C(n, M)\, \gamma_2(x, 2r)^{\frac{2}{n+2}} \tag{10}$$

pour $x \in \mathbb{R}^n$ et $r > 0$. Donc $\widetilde{\gamma}_2$ est défini comme ci-dessus, avec une constante $C(n, 2)$ assez grande pour que la seconde inégalité de (10) soit vraie. Et pour $\widetilde{\gamma}_\infty(x, r)$, le plus simple est de prendre la même constante $C(n, 2)$. J'admets, c'est un peu compliqué parce que la classe des fonctions affines autoriseées dépend de f, mai c'est logique aussi.

On pourrait maisément calculer comment $C(n, M)$ dépend de M. La même démonstration donne aussi que $\widetilde{\gamma}_\infty(x, r) \leq C(n, M, q)\widetilde{\gamma}_q(x, 2r)^{\frac{q}{q+2}} = C(n, M, q)\gamma_q(x, 2r)^{\frac{q}{q+2}}$ (où $\widetilde{\gamma}_\infty$ et $\widetilde{\gamma}_q$ sont définis avec la constante $C(n, q)$ plus haut).

Notons déjà que comme on peut essayer $a(y) = f(x)$ dans (2), on trouve que $\widetilde{\gamma}_\infty(x, r) \leq CM$ de toute façon (rien qu'en sachant que f est Lipschitzienne), donc on peut supposer que $\widetilde{\gamma}_2(x, 2r) \leq c$, avec c aussi petit qu'on veut (dépendant aussi de M), car sinon (10) est trivial.

Choisissons a affine telle que $||a||_{lip} \leq C(n, 2)||f||_{lip}$ et

$$\frac{1}{2r}\Big\{\frac{1}{|B(x, 2r)|}\int_{B(x,2r)} |f(y) - a(y)|^2 dy\Big\}^{1/2} \leq 2\widetilde{\gamma}_2(x, r) \tag{11}$$

(comparer avec la définition (1)), posons $\lambda = C(n, M)\, \widetilde{\gamma}_2(x, 2r)^{\frac{2}{n+2}}$, et vérifions que

$$|f(z) - a(z)| \leq \lambda r \text{ pour } z \in B(x, r). \tag{12}$$

On en déduira bien que $\widetilde{\gamma}_\infty(x, r) \leq \lambda$, et (10) suivra.

Supposons que $|f(z) - a(z)| > \lambda r$ pour un $z \in B(x, r)$. Posons $\rho = [2(1 + C(n, 2))M]^{-1}\lambda r$, puis $D = B(z, \rho)$. Puisque $\widetilde{\gamma}_2(x, 2r) \leq c$ avec c aussi petit qu'on veut, on a que $\rho < r$, donc $D \subset B(x, 2r)$.

Par ailleurs, $f - a$ est $(1 + C(n, 2))M$-lipschitzienne (par définition de f et de a), donc $|f(w) - a(w)| \geq |f(z) - a(z)| - (M + 1)\rho > \lambda r - \lambda r/2 > \lambda r/2$ pour $w \in D$, ce qui donne en intégrant

$$\begin{aligned}\int_{B(x,2r)} |f(y) - a(y)|^2 dy &\geq \int_D |f(y) - a(y)|^2 dy \geq (\lambda r/2)^2|D| \\ &\geq C_0^{-1}(\lambda r)^2\rho^n = C_0^{-1}[2(1 + C)M]^{-n}(\lambda r)^{2+n}\end{aligned} \tag{13}$$

où C_0 ne dépend que de n. D'un autre coté, (11) dit que

$$\int_{B(x,2r)} |f(y) - a(y)|^2 dy \geq C_1 r^{2+n} \widetilde{\gamma}_2(x,r)^2$$

donc finalement $\lambda^{2+n} \leq C_0 C_1^{-1} [2(1 + C(n,2))M]^n \widetilde{\gamma}_2(x,r)^2$, ce qui contredit la définition de λ si $C(M)$ est assez grand. Donc on a (10). □

Donnons maintenant une traduction géométrique du corollaire 1, sans insister trop. Pour un ensemble fermé E dans $\mathbb{R}^n$, et lorsque la dimension entière $d < n$ a été donnée, on pose

$$\beta_q(x,r) = \inf_{P \in \mathcal{P}_d} \left\{ r^{-d} \int_{E \cap B(x,r)} r^{-q} \operatorname{dist}(y,P)^q d\mathcal{H}^d(y) \right\}^{1/q} \tag{14}$$

pour $x \in E$ et $r > 0$, où l'on a noté $\mathcal{P}_d$ l'ensemble des plans affines de dimension d dans $\mathbb{R}^n$. Pour $q = +\infty$, on note souvent directement

$$\beta(x,r) = \beta_\infty(x,r) = \inf_{P \in \mathcal{P}_d} \sup \left\{ r^{-1} \operatorname{dist}(y,P) \right\}. \tag{15}$$

Souvent on ajoute la contrainte que P contient x dans (14) et (15). Ca ne change souvent pas grand-chose.

On pourrait aussi définir $\beta_q(x,r)$ pour x hors de E, mais on ne saurait pas forcément si $\beta_q(x,r)$ est petit parce que E est ténu dans $B(x,r)$ ou parce qu'il est vraiment proche d'un plan en moyenne. En fait, le bon cadre pour utiliser les $\beta_q(x,r)$ avec la définition ci-dessus est quand E est Ahlfors-régulier de dimension d, ce qui signifie que E n'est pas réduit à un point et qu'il existe une constante $C \geq 1$ telle que

$$C^{-1} r^d \leq \mathcal{H}^d(E \cap B(x,r)) \leq C r^d \quad \text{pour } x \in E \text{ et } 0 < r < \operatorname{diam}(E). \tag{16}$$

Alors, on a également que $\beta_p(x,r) \leq C\beta_q(x,r) \leq C^2 \beta(x,r)$ pour $1 \leq p \leq q \leq +\infty$, $x \in E$, et $0 < r < \operatorname{diam}(E)$, par (12) et Hölder.

Noter que nos normalisations font que $\beta(x,r) \leq 1$, donc (dans le cas Ahlfors-régulier) $\beta_q(x,r) \leq C$.

Le corollaire 1 permet de contrôler les $\beta_q(x,r)$ lorsque E est le graphe d'une fonction Lipschitzienne.

Corollaire 2. *Soit $f : \mathbb{R}^d \to \mathbb{R}^{n-d}$ une fonction Lipschitzienne, et soit $E \subset \mathbb{R}^n$ le graphe de f. Soit $q \in [1, +\infty]$. Si $d \geq 2$, supposons que $q < \frac{2d}{d-2}$. Alors* *les nombres $\beta_q(x,r)$ définis ci-dessus vérifient la condition suivante (encore une condition de mesure de Carleson). Il existe $C = C(||\nabla f||_\infty, n, d, q)$ telle que*

(17)
$$\int_{t=0}^{r} \int_{y\in E\cap B(x,r)} \beta_q(y,t)^2 \, \frac{d\mathcal{H}^d(y)\, dt}{t} \leq C\, r^n \quad \textit{pour } x \in E \textit{ et } 0 < r < \mathrm{diam}(E).$$

[Ici, $\mathrm{diam}(E) = +\infty$*, mais on écrirait comme ça dans le cas général où E serait un ensemble Ahlfors-régulier pas forcément borné.]*

La démonstration est très facile. D'abord, sur le graphe E, la mesure de Hausdorff $\mathcal{H}^d$ est équivalente à la mesure image de la mesure de Lebesgue par le paramétrage naturel $x \to (x, f(x))$ de E par $\mathbb{R}^d$. Maintenant, si $z = (x, f(x)) \in E$ et $r > 0$ sont donnés, on se donne une fonction affine a qui réalise presque l'infimum dans la définition de $\widetilde{\gamma}_q(x,r)$ (voir (1) et les commentaires avant (10)), et on utilise le plan affine P d'équation $y = a(x)$ pour évaluer $\beta_q(z,r)$. Il s'agit alors de remarquer que pour $z' = (x', f(x')) \in E \cap B(z,r)$, $\mathrm{dist}(z', P) \leq |f(x') - a(x')|$, ce dont on déduit que

$$(18) \qquad \begin{aligned} \beta_q(z,r)^q &= r^{-d} \int_{E\cap B(z,r)} r^{-q}\, \mathrm{dist}(z',P)^q \, d\mathcal{H}^d(z') \\ &\leq C r^{-d} \int_{x'\in B(x,r)} r^{-q} |f(x') - a(x')|^q dx' \leq C \widetilde{\gamma}_q(x,r)^q \end{aligned}$$

par équivalence des deux mesures. Ensuite on intègre (18) en tenant compte de l'équivalence des deux mesures, et on obtient (13). □

Pour des ensembles E Ahlfors-réguliers de dimension d, (13) est l'une des nombreuses définitions équivalentes de ce que l'on appelle “rectifiabilité uniforme”; ainsi, le corollaire 2 dit que les graphes lipschitziens sont uniformément rectifiables (UR). Dans la définition par (13), on a mis l'accent sur une propriété géométrique de proximité à des plans affines (à mettre en parallèle avec l'existence presque-partout de plans tangents approchés pour les ensembles rectifiables), mais il y a d'autres angles d'approche. On a eu un peu moins de succès avec les histoires de densité (caractérisation de l'UR seulement en dimension 1 ou codimension 1, si je me souviens bien), ou avec les projections. Par contre, on a des caractérisations convenables par la continuité sur $L^2(E)$ d'opérateurs d'intégrale singulière, pas mal d'autres caractérisations géométriques, et quelques applications. La notion de rectifiabilité simple est très pratique parce qu'on peut prendre des unions dénombrables, mais UR est plus logique quand on veut un contrôle quantitatif, d'où l'intérêt d'UR. Voir l'exposé de Park City sur mon site pour plus de détails.

L'idée d'utiliser encore l'espace des boules (ici, $E\times]0,+\infty[$) pour mesurer la régularité de E est ici aussi très fructueuse. Elle est due à S. Semmes dans ce contexte géométrique.

20. LIPSCHITZ ET BILIPSCHITZ

On va démontrer un résultat amusant de P. Jones, N. Katz, et A. Vargas, sur l'existence de morceaux bilipschitziens dans une application lipschitzienne. On le verra comme application du théorème de Dorronsoro pour rentabiliser, mais c'est un peu abusif (il y aurait peut-être eu plus court, quoique notre preuve par Plancherel ne soit pas très lourde).

La vraie excuse est l'utilisation, une fois de plus, d'un argument combinatoire avec des cubes dyadiques, même si ça n'est pas exactement un argument de temps d'arrêt comme j'aurais plutôt souhaité pour conclure. Néanmoins, je trouve l'argument très rigolo (et il a encore plus de sel quand on connaît l'argument plus compliqué connu juste avant et qui donnait un moins bon résultat (le corollaire)).

Théorème (JKV). *Pour tout choix de n (dimension entière) et $\delta > 0$, il existe un entier $M = M(n,\delta)$, avec la propriété suivante. Soient Q_0 le cube unité de $\mathbb{R}^n$ et $f : Q_0 \to \mathbb{R}^n$ une application 1-lipschitzienne (une normalisation qui ne coûte rien). On peut trouver M ensembles compacts $K_1, \cdots K_M \subset Q_0$ tels que: pour tout $j \le M$,*

$$|f(x) - f(y)| \ge \delta |x-y| \quad \text{pour } x, y \in K_j\,, \tag{1}$$

et

$$\mathcal{H}^n\big(f\big(Q_0 \setminus \cup_j K_j\big)\big) \le 2\delta. \tag{2}$$

Des commentaires d'abord. Ceci n'est vraiment utile que si $\mathcal{H}^d(f(Q_0)) > 2\delta$, puisque dans ce cas, il reste $\mathcal{H}^d(f(\cup_j K_j)) \ge \mathcal{H}^d(f(Q_0)) - 2\delta > 0$. Dit un peu autrement, si $\delta \ge 1/2$, le théorème ne donne aucune information. Et même le corollaire suivant est intéressant.

Par ailleurs, si l'image est petite, il se peut par exemple que f soit affine très contractante, et alors, pas question de chercher un morceau bilipschitzien!

Corollaire. *Pour tout $\delta > 0$, il existe $\varepsilon = \varepsilon(n, \delta) > 0$ tel que si $f : Q_0 \to \mathbb{R}^n$ est une application 1-lipschitzienne telle que $\mathcal{H}^d(f(Q_0)) > 3\delta$, alors il existe un compact $K \subset Q_0$ tel que $|f(x) - f(y)| \geq \delta|x - y|$ pour $x, y \in K_j$ et $\mathcal{H}^d(f(K)) \geq \varepsilon$.*

Appliquer le théorème, qui dit que

$$\sum_j \mathcal{H}^d(f_j(K)) \geq \mathcal{H}^d(f(\cup_j K_j)) \geq [\mathcal{H}^d(f(Q_0)) - 2\delta] \geq \delta, \tag{3}$$

et choisir j tel que $\mathcal{H}^d(f_j(K)) \geq \delta/M$. Donc on peut prendre $\varepsilon = \delta/M$, où M sort du théorème.

Si $\mathcal{H}^d(f(Q_0)) > 0$, il est assez facile d'obtenir un ensemble de mesure positive surlequel f est bilipschitzienne. D'abord, on jette un ensemble W de petite mesure (donc il reste encore $\mathcal{H}^d(f(Q_0 \setminus W)) > 0$), et à ce prix on peut supposer que f est de classe C^1. Ensuite on utilise Sard pour dire qu'il existe un point x de densité de $Q_0 \setminus W$ en lequel la différentielle de f est de rang n (et on peut même se débrouiller en regardant la preuve de Sard pour que la différentielle soit bilipschitzienne avec un certain contrôle). Et un prend l'intersection de $Q_0 \setminus W$ avec une petite boule centrée en x. Mais ce qui est bien ici, c'est que l'on puisse prouver le corollaire avec une vraie borne inférieure.

En tout cas, il est assez facile de construire des exemples où K est forcément d'intérieur vide: faire plein de petits plis. Ceci semble interdire une démonstration trop triviale par un argument de compacité.

Passons à la démonstration du théorème. Soient donc n et $\delta > 0$ donnés. Donnons-nous aussi un très petit $\varepsilon > 0$, à choisir plus tard.

Pour chaque entier $k \geq 0$, notons Δ_k l'ensemble des cubes dyadiques contenus dans Q_0 et de côté 2^{-k}. Notons $\mathcal{B}_k$ l'ensemble des cubes $Q \in \Delta_k$, qui n'ont PAS la bonne propriété suivante:

(4)

il existe une fonction affine a_Q telle que $|f(x) - a_Q(x)| \leq \varepsilon 2^{-k}$ pour $x \in 10\sqrt{n}Q$.

On a noté $10\sqrt{n}Q$ le cube de même centre et de côté $10\sqrt{n} \cdot 2^{-k}$. On va utiliser le résultat de Dorronsorro pour démontrer le lemme suivant, qui n'utilise que le fait que f est 1-lipschitzienne.

Lemme 1. *Il existe une constante $C(\varepsilon)$, qui ne dépend que de n et ε, telle que*

$$\sum_k \sum_{Q \in \mathcal{B}_k} |Q| \leq C(\varepsilon). \tag{5}$$

C'est encore notre condition de Carleson préférée, sauf qu'ici on n'en a besoin que pour le cube unité.

Pour pouvoir appliquer le théorème de Dorronsorro tranquillement, prolongeons f à $\mathbb{R}^n$ tout entier (mais on n'aura pas besoin ensuite de connaître les valeurs de f loin de Q_0). On peut facilement trouver un prolongement C-lipschitzien (même si en fait il existe un prolongement 1-lipschitzien, plus difficile à construire). Vérifions que si $Q \in \mathcal{B}_k$, alors

$$\widetilde{\gamma}_2(x,r) \geq C_1^{-1}\varepsilon^{\frac{n+2}{2}} \tag{6}$$

pour tout $x \in Q$ et tout r tel que $100\sqrt{n}\cdot 2^{-k} \leq r < 200\sqrt{n}\cdot 2^{-k}$, et où $\widetilde{\gamma}_2$ est la variante de γ_2 introduite peu avant (10) dans le chapitre précédent, où l'on approxime seulement par des fonctions $n||f||_{lip}$-lipschitziennes.

Procédons calmement par contraposition. Supposons que $\widetilde{\gamma}_2(x,r) \leq C_1^{-1}\varepsilon^{\frac{n+2}{2}}$ pour un choix de x et de r comme ceci-dessus. Alors $\widetilde{\gamma}_\infty(x,r/2) \leq CC_1^{-1}\varepsilon$, à cause de (10) dans le chapitre précédent, et où C a le droit de dépendre de $||f||_{lip}$. Par définition, on trouve une fonction affine (et aussi C-lipschitzienne) a telle que $|f(y)-a(y)| \leq CC_1^{-1}\varepsilon r$ dans $B(x,r/2)$. Et comme cette boule contient largement $10Q$ parce que $x \in Q$ et $100\sqrt{n}\cdot 2^{-k} \leq r$, on peut prendre $a_Q = a$ dans (4), donc Q n'est pas dans $\mathcal{B}_k$. Donc on a (6).

Passons maintenant au lemme 1. Noter que les ensembles

$$E_Q = \big\{(x,r)\,;\, x \in \mathrm{int}(Q) \text{ et } 100\sqrt{n}\cdot 2^{-k} \leq r < 200\sqrt{n}\cdot 2^{-k}\big\} \tag{7}$$

associés à des cubes Q différents sont disjoints, puisque pour un (x,r) donné, r détermine le k tel que $Q \in \Delta_k$, et ensuite Q est unique puisque $x \in \mathrm{int}Q$. Noter aussi que $r \leq r_0 = 200\sqrt{n}$. Finalement,

$$\begin{aligned}\sum_k \sum_{Q\in\mathcal{B}_k} |Q| &\leq C\varepsilon^{-n-2}\sum_k\sum_{Q\in\mathcal{B}_k}\int_{E_Q}\widetilde{\gamma}_2(x,r)^2\frac{dxdr}{r} \\ &\leq C\varepsilon^{-n-2}\int_{Q_0\times(0,r_0)}\widetilde{\gamma}_2(x,r)^2\frac{dxdr}{r} \leq C(\varepsilon).\end{aligned} \tag{8}$$

par (6) et le corollaire du théorème de Dorronsorro. D'où le lemme 1. □

Remarque: il y a éventuellement un peu plus direct pour démontrer le lemme, mais c'est assez pratique pour nous de le faire comme ça. Dans l'article initial de JKV, il me semble qu'ils calculaient des coefficients d'ondelettes, et démontraient une condition de type Carleson sur les coefficients en question. Ce qui ressemble un peu.

Notons maintenant $\mathcal{C}$ la classe des cubes $Q \in \Delta = \cup_k \Delta_k$ tels que $\mathcal{H}^n(f(Q)) \leq \frac{3\delta}{2}\,|Q|$, et qui sont maximaux avec cette propriété.

Par maximalité, ces cubes sont presque disjoints (si les intérieurs de deux d'entre eux se rencontraient, l'un serait contenu dans l'autre, comme toujours pour des cubes dyadiques). Donc

$$\mathcal{H}^n(f(\cup_{Q\in\mathcal{C}} Q)) \leq \frac{3\delta}{2} \sum_{Q\in\mathcal{C}} |Q| \leq \frac{3\delta}{2}. \tag{9}$$

On choisit une grande constante N, telle que $C(\varepsilon) < \delta N/2$. On pose

$$Z = \big\{x \in Q_0\,;\, x \text{ est dans au moins } N \text{ cubes de } \cup_k \mathcal{B}_k\big\} \cup \big[\cup_{Q\in\mathcal{C}}\, Q\big] \tag{10}$$

Alors
(10)

$$\begin{aligned}\mathcal{H}^n(f(Z)) &\leq \mathcal{H}^n\Big(\big\{x \in Q_0\,;\, x \text{ est dans au moins } N \text{ cubes de } \cup_k \mathcal{B}_k\big\}\Big) + \frac{3\delta}{2}\\ &\leq \frac{1}{N}\sum_k \sum_{Q\in\mathcal{B}_k} |Q| + \frac{3\delta}{2} \leq \frac{C(\varepsilon)}{N} + \frac{3\delta}{2} \leq 2\delta\end{aligned}$$

par (9), puis le lemme 1, puis le choix de N.

On pose $E = Q_0 \setminus Z$; ainsi $\mathcal{H}^n(f(Q_0 \setminus E)) \leq 2\delta$, et il ne reste plus qu'à trouver au plus M compacts K_j qui recouvrent E et sur lesquels f est δ^{-1} bilipschitzienne, comme en (1).

Notons $\mathcal{H}$ la classe des couples ordonnés (Q_1, Q_2), où Q_1 et Q_2 sont dans $\cup_k \mathcal{B}_k$, et sont semi-adjacents au sens où ils sont dans un même Δ_k, ne sont pas égaux, mais $Q_1 \subset 10\sqrt{n}Q_2$ (disons).

Mettons un ordre $<$ sur $\mathcal{H}$, tel que $(Q_1, Q_2) < (Q_1', Q_2')$ automatiquement quand les Q_i sont dans Δ_k et les Q_i' sont dans $\Delta_{k'}$ pour un $k' > k$ (les grands cubes passent en premier). Ceci nous laisse encore beaucoup de choix, mais ça n'est pas grave.

Notons (Q_1^l, Q_2^l) le l-ième élément de $\mathcal{H}$. Notons $k(l)$ l'entier k tel que Q_1^l et Q_2^l sont dans Δ_k. Donc $k(l)$ est une fonction croissante de l.

Pour $x \in Q_0$ et $l \geq 0$, on va définir maintenant un mot $\alpha_l(x)$ de longueur finie, composé de 1 et de 0, et ceci avec les propriétés suivantes:

$$\alpha_{l+1}(x) \text{ commence par } \alpha_l(x), \tag{i}$$

(ii) $\qquad\qquad \alpha_l(x)$ est constant sur chaque cube de $\Delta_{k(l)}$.

Il va en fait s'agir de coder dans $\alpha_l(x)$ l'appartenance ou non de x aux Q_i^l précédents.

Notons $n_l(x)$ la longueur de $\alpha_l(x)$. Donc ce sera une fonction croissante de l.

On prend pour $\alpha_0(x)$ le mot vide (par exemple). Supposons ensuite $\alpha_l(x)$ défini, et définissons $\alpha_{l+1}(x)$. Il y a plusieurs cas.

(a) On garde $\alpha_{l+1}(x) = \alpha_l(x)$ si x n'est ni dans Q_1^l ni dans Q_2^l.

Il reste à définir α_{l+1} sur les Q_j^l. Notons α_1 et α_2 les valeurs constantes de $\alpha_l(\cdot)$ sur Q_1^l et Q_2^l respectivement. Notons n_1 et n_2 leurs longueurs respectives

(b) Si $n_1 = n_2$, on prolonge $\alpha_l(x)$ par 0 sur Q_1^l et par 1 sur Q_2^l

(c) Si $n_1 > n_2$, on garde $\alpha_{l+1}(x) = \alpha_l(x)$ sur Q_1^l et on prolonge $\alpha_l(x)$ par $\varepsilon \in \{0, 1\}$ sur Q_2^l, où l'on cloisit ε différent de la coordonnée numéro $n_2 + 1$ de n_1. Comme ceci, on sera sur que α_{l+1} ne commence pas pareil sur Q_2^l que sur Q_1^l.

(d) Si $n_1 < n_2$ on fait pareil qu'en (c), mais en échangeant les rôles de 1 et 2.

Ceci termine notre définition des α_l par récurrence. Il reste à voir les propriétés. Les propriétés (i) et (ii) sont faciles (la seconde, par récurrence en même temps que la construction). Pour être franc, on va négliger les frontières des cubes, ou alors il faudrait les prendre non fermés et disjoints. Ca n'est pas grave; noter que si f est bilipschitzienne sur F, elle l'est aussi sur $\overline{F}$. On note que dans les procédures (a-d), on ne sépare pas des points d'un même cube de $\Delta_{k(l)}$ (et donc encore moins les cubes éventuellement plus petits de $\Delta_{k(l+1)}$.

Lemme 2. *Pour $x \in E$, les $\alpha(x)$ sont de longueur au plus CN.*

En effet, la longueur de $\alpha_l(x)$ ne change (dans la procédure plus haut) que lorsque x est dans l'un des deux Q_j^l. Il s'agit donc de savoir que ceci n'arrive pas plus de CN fois. Mais quand c'est le cas, les deux Q_j^l sont dans un $\mathcal{B}_k$. Et x n'est pas dans plus de N tels cubes (par définition de Z et E), et chaque tel cube n'intervient dans notre suite de couples (Q_1, Q_2) qu'au plus C fois, à cause de la définition de "semi-adjacent". $\square$

Passons aux K_j. Pour $x \in E$, notons $\alpha(x)$ la valeur finale de $\alpha_l(x)$. Et posons, pour tout mot α de longueur $\leq CN$,

$$K_\alpha = \{x \in E\,;\, \alpha(x) = \alpha\}. \tag{11}$$

Il n'y en a pas pas plus de 2^{CN+1}; il s'agit encore de voir que f est bilipschitzienne sur chaque K_α. [Se souvenir qu'on peut remplacer K_α par son adhérence, si nécessaire.]

Soient donc x et y dans K_α. Pour chaque k, notons $Q_k(x)$ et $Q_k(y)$ les cubes de côté 2^{-k} qui les contiennent. Choisissons k tel que $2^{-k+1}\sqrt{n} < |x-y| \leq 2^{-k+2}\sqrt{n}$. De cette manière, on est certain que $Q_k(x)$ et $Q_k(y)$ sont distincts, et aussi qu'ils sont semi-adjacents avec la définition plus haut (c'est-à-dire, que $Q_2 \subset 10\sqrt{n}Q_1$).

Si ces deux cubes formaient une paire de $\mathcal{H}$, on se serait débrouillé en (b) ou (c) pour que $\alpha(x) \neq \alpha(y)$, parce qu'ils ne commencent pas pareil. Donc au moins un des deux (notons-le Q) n'est pas dans $\mathcal{B}_k$. Ce qui signifie qu'il vérifie (4), et il existe une fonction affine $a = a_Q$ qui est $\varepsilon 2^{-k}$-proche de f sur $10Q$.

Si a est telle que $|a(x) - a(y)| \geq \frac{11\delta}{10}\,|x-y|$ et ε a été choisi assez petit, on a prouvé que $|f(x) - f(y)| \geq \delta|x-y|$ et on est content.

Sinon, il existe au moins une direction de $\mathbb{R}^n$ (celle de $x-y$) où a est $\frac{11\delta}{10}$-contractante. Par ailleurs, puisque a est si proche de f qui est 1-Lipschitzienne, on voit que a est $(1+C\varepsilon)$-Lipschitzienne. On utilise tout ceci pour montrer que l'image de Q par a est contenue dans un objet simple de diamètre $\leq (1+\varepsilon)\operatorname{diam}(Q)$ et de mesure au plus $(1+\varepsilon)^n\,\frac{11\delta}{10}|Q|$.

J'essaie quand même d'en dire un peu plus. Objet simple, c'est juste une manière de dire image d'un cube par une application affine $(1+C\varepsilon)$-Lipschitzienne. Pour ce qui est du calcul de la mesure de l'image, c'est juste $|Q|$ fois la valeur absolue du déterminant de cette application a. Qu'on peut calculer dans une base orthonormée qui commence dans une base par un premier vecteur e_1 sur lequel a est $\frac{11\delta}{10}$-contractante. Quitte à composer par une isométrie qui ramène $f(e_1)$ dans la direction de e_1 (sans changer la valeur absolue du déterminant), on peut supposer que la matrice de a a un bloc de taille 1 (en haut à gauche) $\frac{11\delta}{10}$-contractant, une partie en haut à droite qui ne compte pas pour le déterminant, et un gros bloc en bas à droite qui est encore la matrice d'une application $(1+C\varepsilon)$-Lipschitzienne. On en déduit l'estimation du volume souhaitée. Maintenant la forme simple, ça aide juste pour dire qu'un $C\varepsilon\operatorname{diam}(Q)$ voisinage de $a(Q)$, ça a une mesure qui n'est pas plus que celle de $a(Q)$, plus $C\varepsilon\operatorname{diam}^{n-1}(Q)$. Je n'ai pas d'autre moyen intelligent de faire ceci que d'avoir une borne pour la surface, et de la recouvrir par des boules de rayon $C\varepsilon\operatorname{diam}(Q)$.

Finalement, par (4), $f(Q)$ est à peine plus grand, donc $|f(Q)| < \frac{3\delta}{2}\,|Q|$ (à nouveau, si ε est assez petit). On regarde la définition de $\mathcal{C}$ juste avant

(9), on observe que donc $Q \in \mathcal{C}$, et on en déduit que $Q \subset Z$ (par (10)), ce qui est une contradiction puisque $E = Q_0 \setminus Z$, et que $x, y \in E$, alors que l'un des deux est dans Q.

Ceci termine la démonstration du théorème. $\square$

Deux mots pour finir des motivations qui pouvaient mener à un tel résultat. Il s'agissait en gros d'étudier des ensembles qui contiennent dans chaque boule une partie significative d'image par rotation de graphe lipschitzien. On peut imaginer pourquoi le théorème de JKV ci-dessus peut permettre de prouver que certains ensembles sont de ce type. [Par une variante, on est ramené à trouver des applications lipschitziennes dont l'image est assez grande, ce qui est apparament beaucoup plus facile.] Ensuite, il faut se convaincre que pour divers problèmes d'analyse, on peut obtenir des résultats pour les ensembles qui contiennent ainsi des morceaux de graphes lipschitziens, à partir des résultats correspondants pour les graphes eux-mêmes. Il se trouve que ça arrive. Et la raison pour laquelle ça a des chances de marcher est du même type que pour le théorème de John et Nirenberg: un peu d'information dans chaque boule peut donner beaucoup d'information globale.

21. LE THEOREME DE REIFENBERG

On va essayer de décrire, et de donner un gros bout de démonstration du théorème de Reifenberg dit du disque topologique. En gros, on se donne un ensemble compact E qui, dans $B(0,10)$, ressemble assez, dans chaque boule centrée sur E, à un plan de dimension d, et on en déduit que, dans $B(0,5)$, cet ensemble a un paramétrage bihöldérien par un d-plan. Voyons ça plus en détails.

L'énoncé dit que pour tout choix d'entiers $0 < d < N$ et toute petite constante $\tau \in]0, 10^{-1}]$, il existe $\varepsilon \in]0, 10^{-2}]$ tel que l'on ait l'énoncé suivant. On se donne un ensemble compact $E \subset \mathbb{R}^N$. On suppose que $0 \in E$ et que, pour tout choix de $x \in E \cap B(0,10)$ et $0 \leq r \leq 10$, il existe un d-plan affine $P = P(x,r)$ passant par x tel que

(1)
$$\mathrm{dist}(y,P) \leq \varepsilon r \text{ pour } y \in E \cap B(x,r) \text{ et } \mathrm{dist}(y,E) \leq \varepsilon r \text{ pour } y \in P \cap B(x,r).$$

On en déduit qu'il existe une application $f : B(0,3) \to \mathbb{R}^N$, avec les propriétés suivantes:

(2) $$|f(x) - x| \leq \tau \text{ pour } y \in B(0,3),$$

(3) $(1-\tau)|x-y|^{1+\tau} \leq |f(x)-f(y)| \leq (1+\tau)|x-y|^{1-\tau}$ pour $x, y \in B(0,2)$

(la condition bihöldérienne, qui d'ailleurs implique que f est un homéomorphisme de $B(0,2)$ dans son image), et, en posant $Z = P(0,10)$,

(4) $$E \cap B(0,1) \subset f(Z \cap B(0,2)) \subset E \cap B(0,3).$$

On commence par des commentaires avant de démontrer un bout du théorème.

D'abord, on voit qu'en prenant τ petit, on obtient one condition bihöldérienne aussi proche qu'on veut du bilipschitzien. Mais Bilipschitz serait faux, car un flocon de neige, avec angle de base très petit, est un exemple d'ensemble E qui vérifie les hypothèses avec $d = 1$, mais n'est pas localement bilipschitz-équivalent à un segment, pour la bonne raison que sa longueur est localement infinie.

Il découle de (2), (3), et un argument de degré topologique (comme celui qu'on a évoqué pour la surjectivité locale des applications lipschtziennes avec une dérivée de rang maximal, quand on a projeté sur un d-plan) que l'image $f(B(0,3))$ contient $B(0,2)$. Donc f est un homéomorphisme local.

Souvent on ne donne qu'une fonction $f : Z \cap B(0,3) \to E$; ici on a un peu plus, puisqu'on dit que cette application a un prolongement à l'espace, ce qui signifie que non seulement E est localement homéomorphe à un d-disque, mais qu'en plus la manière dont il est plongé respecte la topologie de $\mathbb{R}^N$. Mais dans l'esquisse de démonstration ci-dessous, on se contentera de construire la restriction de f à Z (en laissant la seconde étape de prolongement de f).

Reifenberg (Annals of math., 1964) s'intéressait à ceci pour des histoires d'existence et de régularité d'ensembles et de surfaces minimales.

J'aime le résultat pour deux raisons. D'abord, comme dans le cas de John-Nirenberg, c'est un cas où une connaissance vague de E, mais à toutes les échelles, donne une connaissance bien plus précise que prévu. Autrement dit, l'hypothèse (1) dans chaque boule ne donne qu'une idée très vague de la topologie de E (il pourrait y avoir des trous, ou des boucles, au d'autres trucs bizarres), mais en fait tous ces mauvais comportements sont exclus par la même hypothèse à d'autres échelles. Bref, si vous disposez d'une mauvaise vue qui rend tout un peu flou, mais si vous avez le moyen de zoomer où vous voulez et à toutes les échelles, vous pouvez quand même tout contrôler.

La seconde raison est que la démonstration est encore un bel exemple d'algorithme descendant.

Maintenant passons à la démonstration. Il est assez amusant que le mouvement naturel (en tout cas le mien) aurait été de chercher une application de E vers Z, et que l'application $f : Z \to E$ est en fait beaucoup plus facile à construire.

Donc on veut définir f sur Z, et puisqu'on veut que l'image soit dans E, l'idée est de pousser les points vers E. On ne peut pas le faire bien directement, mais on va le faire étape par étape, avec une étape à chaque échelle 2^{-n}, $n \geq 0$, en constatant qu'à chaque échelle E est localement proche d'un plan, ce qui nous permet de pousser les points dans cette direction.

On aura besoin de partitions de l'unité à chaque échelle. Donc, pour tout $n \geq 0$, on se donne une collection maximale $\{x_i\}$, $i \in I_n$, de points de $E \cap B(0, 10)$, avec la propriété que $|x_i - x_j| \geq 2^{-n}$ pour $i, j \in I_n$, $i \neq j$. On notera $B_i = B(x_i, 2^{-n})$.

Travaillons à n fixé pour l'instant. Par maximalité de I_n, $E \cap B(0, 10)$ est contenu dans l'union des $\overline{B}_i$. Pour chaque i, soit $\widetilde{\theta}_i$ une fonction bosse telle que $\widetilde{\theta}_i = 1$ sur $2B_i$, $0 \leq \widetilde{\theta}_i \leq 1$ sur $3B_i \setminus 2B_i$, et $\widetilde{\theta}_i = 0$ hors de $3B_i$. Le plus simple est de prendre des translations et dilatations d'une même fonction bosse de classe C^∞, de sorte que $|\nabla^k(\widetilde{\theta}_i)| \leq C_k 2^{nk}$ pour tout $k \geq 0$ et $i \in I_n$. Posons

$$E_0 = E \cap B(0, 10) \quad \text{et} \quad D_n = \left\{y \in \mathbb{R}^N \,;\, \mathrm{dist}(y, E_0) \leq 2^{-n}\right\}. \tag{5}$$

Observons que

$$E_0 \subset \bigcup_{i \in I_n} \overline{B}_i \tag{6}$$

donc

$$D_n \subset \bigcup_{i \in I_n} 2\overline{B}_i \tag{7}$$

de sorte que

$$\sum_{i \in I_n} \widetilde{\theta}_i(y) \geq 1 \text{ sur } D_n. \tag{8}$$

La somme est aussi inférieure à C, parce que les $3B_i$ sont trivialement de recouvrement borné, ce qui nous permet de définir

$$\theta_i(y) = \frac{\widetilde{\theta}_i(y)}{\sum_{j \in I_n} \widetilde{\theta}_j(y)} \text{ pour } y \in D_n. \tag{9}$$

On a alors que

$$\sum_{i \in I_n} \theta_i(y) = 1 \text{ sur } D_n \tag{10}$$

et

$$|\nabla^k \theta_i(x)| \leq C_k 2^{nk} \text{ pour } y \in D_n. \tag{11}$$

Voilà pour les partitions de l'unité locales. On aura aussi besoin de plans. Pour $i \in I_n$, on pose $P_i = P(x, 10 \cdot 2^{-n})$ (donné par (1)), et on note π_i la projection orthogonale sur P_i. On aura besoin de savoir qu'il y a une certaine cohérence sur les P_i.

Lemme 1. *Si $i, j \in I_n \cap I_{n-1}$, et si $3B_i \cap 3B_j \neq \emptyset$, alors π_i et π_j sont si proches que*

$$|\pi_i(y) - \pi_j(y)| \leq C\varepsilon 2^{-n} \text{ pour } y \in 3B_i \cup 3B_j \tag{12}$$

et, en notant $D\pi_i$ la différentielle de π_i (qui est aussi la projection orthogonale sur le plan vectoriel parallèle à P_i) et pareil pour $D\pi_j$,

$$||D\pi_i - D\pi_j|| \leq C\varepsilon. \tag{13}$$

Notons P_i' le plan vectoriel parallèle à P_i et choisissons une base orthonormée $e_1, \cdots, e_d$ de P_i'. Pour chaque k, $x_i + 2^{-n}e_k$ est dans $P_i \cap 10B_i$, donc il existe $z_k \in E$ tel que $|x_i + 2^{-n}e_k - z_k| \leq 10\varepsilon 2^{-n+1}$. En particulier $z_k \in 10B_j$, donc il existe $w_k \in P_j$ tel que $|w_k - z_k| \leq 10\varepsilon 2^{-n+1}$. De même, $x_i \in P_i \cap 10B_i$ donc il existe $z_0 \in E$, puis $w_0 \in P_j$ tel que $|w_0 - x_i| \leq 10\varepsilon 2^{-n+2}$. Ainsi la différence $f_k = 2^n(w_k - w_0)$ est dans P_j', et $|f_k - e_k| \leq 40\varepsilon$. Les f_k sont bien indépendants (si ε est assez petit); ils forment une base (presque orthogonale) de P_j'.

Maintenant on peut vérifier (13) assez facilement. On complète $e_1, \cdots, e_d$ en une base orthonormée de $\mathbb{R}^N$, et il suffit de vérifier que $|D\pi_i(e_k) - D\pi_j(e_k)| \leq C\varepsilon$ pour $1 \leq k \leq N$. Pour $k \leq d$, c'est facile, puisque $D\pi_i(e_k) = e_k$ et $|D\pi_j(e_k) - e_k| \leq \mathrm{dist}(e_k, P_j') \leq |e_k - f_k| \leq 40\varepsilon$.

Pour $k > d$, on doit vérifier que $|D\pi_j(e_k)| \leq C\varepsilon$, ou de manière équivalente que $|\langle e_k, v| \leq C\varepsilon$ pour tout $v \in P_j'$ de norme ≤ 1. C'est assez facile, il faut juste vérifier que tout v comme ceci est combinaison linéaire des f_k avec des coefficients ≤ 2. On laisse faire les détails (suggestion: montrer que la matrice de passage M de la base des f_k à une base orthonormée de P'_j est telle que $|M - I| \leq C\varepsilon$.

Maintenant qu'on a (13), (2) est facile, puisqu'il se réduit à vérifier que $|\pi_i(y) - \pi_j(y)| \leq C\varepsilon 2^{-n}$ pour $y = x_i$. Et c'est clair, puisqu'on sait que $\pi_i(x_i) = x_i$ et qu'on a vu que $\mathrm{dist}(x_i, P_j) \leq 20\varepsilon 2^{-n}$. □

Je crois qu'on peut commencer à définir notre application f. En fait, on prendra

$$f = \lim_{n \to +\infty} f_n \tag{14}$$

avec une suite $\{f_n\}$ définie par récurrence par $f_0(x) = x$ et, pour $n \geq 0$, par

$$f_{n+1} = g_n \circ f_n \tag{15}$$

pour des fonctions g_n qu'on va juste définir dans D_n. On verra bientôt que ça se compose bien. En attendant, définissons g_n avec l'idée de pousser les gens vers E. On prend simplement

$$g_n(y) = \sum_{i \in I(n)} \theta_i(y) \pi_i(y) \tag{16}$$

et c'est pour cela qu'on préfère rester dans D_n.

Pour la suite, on va essayer de ne pas mélanger trop les n et les $n+1$, mais des erreurs ne sont pas exclues. Posons $Z_0 = Z \cap B(0, 8)$. On va démontrer par récurrence sur n que f_n est bien définie dans un voisinage de Z_0, que

$$f_n(Z_0) \subset \big\{ y \in \mathbb{R}^n \,;\, \mathrm{dist}(y, E_0) \leq C_1 \varepsilon 2^{-n} \big\}, \tag{17}$$

et que

$$|g_n(y) - y| \leq C_2 \varepsilon 2^{-n} \text{ pour } y \in f_n(Z_0). \tag{18}$$

La constante C_1 ne dépendra que des constantes précédentes (celles du lemme 1 en particulier), et C_2 dépendra aussi de C_1.

On commence avec $n = 0$, pour lequel on sait que $f_0(x) = x$, donc que (17) est vrai parce que $Z = P(0, 10)$ reste tout proche de E.

Supposons maintenant que (17) est vrai à l'ordre n (c'est vrai pour $n = 0$). Commençons par observer que $f_n(Z_0) \subset \mathrm{interieur}(D_n)$ trivialement, donc que

g_n est définie dans un voisinage de $f_n(Z_0)$, et f_{n+1} est définie dans un voisinage de Z_0. Il nous faudra montrer (18) et (17) pour $n+1$.

Soit $y \in f_n(Z_0)$. Par (17), il existe $j \in I_n$ tel que $y \in 2B_j$. De plus soit $i \in I_n$ tel que $\theta_i(y) \neq 0$; alors $y \in 3B_i$, et on peut appliquer le lemme 1 à i et j. Alors

(19)

$$|g_n(y)-\pi_j(y)| = \Big|\sum_{i\in I_n} \theta_i(y)[\pi_i(y)-\pi_j(y)]\Big| \leq \sum_i \theta_i(y)|\pi_j(y)-\pi_i(y)| \leq C\varepsilon 2^{-n}$$

parce que $\sum_{i\in I_n} \theta_i(y) = 1$ et par (12).

Reste à estimer $|\pi_j(y)-y|$. Par (17), il existe $z \in E_0$ tel que $|z-y| \leq C_1\varepsilon 2^{-n}$. Evidemment, $z \in 3B_j$ puisque $y \in 2B_j$, donc (1) dit que $\text{dist}(z, P_j) \leq 10\varepsilon 2^{-n}$. Finalement, $\text{dist}(y, P_j) \leq (C_1+10)\varepsilon 2^{-n}$, $|\pi_j(y)-y| \leq \text{dist}(y, P_j) \leq (C_1+10)\varepsilon 2^{-n}$, et (19) dit que $|g_n(y)-y| \leq (C_1+10+C)\varepsilon 2^{-n}$, ce qui prouve (18).

Supposons maintenant que (18) est également vrai pour $m \leq n$, et vérifions (17) à l'ordre $n+1$. Soit $x \in Z_0$ tel que $z = f_{n+1}(x)$ et posons $y = f_n(x)$. Donc $y \in f_n(Z_0)$ et $z = g_n(y)$. En sommant (18) et ses prédécesseurs, il vient

$$|z-x| = |f_{n+1}(x)-x| \leq 2C_2\varepsilon \tag{20}$$

et en particulier $z \in B(0,9)$ (car $Z_0 \subset B(0,8)$).

Reprenons les notations ci-dessus, en particulier soit $j \in I_n$ tel que $y \in 2B_j$. On a encore (19). De plus, $\pi_j(y) \in P_j \cap 2B_j$ (puisque P_j passe par x_j et $y \in 2B_j$), donc $\text{dist}(\pi_j(y), E) \leq 10\varepsilon 2^{-n}$ (par (1)). Donc $\text{dist}(z, E) = \text{dist}(g_n(y), E) \leq (10+C)\varepsilon 2^{-n}$ par (19), ce qui donne (17) avec $C_1 = 2(10+C)$, car $z \in B(0,9)$ (donc z est aussi proche de E_0 que de E).

Ceci termine notre construction des f_n, et la vérification de (17) et (18).

Notons qu'en sommant (18), on trouve que

$$|f_m(x)-f_n(x)| \leq 2C_2\varepsilon 2^{-n} \quad \text{pour } m > n \text{ et } x \in Z_0 \tag{21}$$

donc que la suite $\{f_n\}$ converge, uniformément sur Z_0, vers une limite qu'on notera f. En plus,

$$|f(x)-x| \leq 2C_2\varepsilon \quad \text{pour } x \in Z_0 \tag{22}$$

de sorte que l'on a prouvé (2) (pour la restriction de f à Z_0, qui est la seule chose qu'on définira ici).

A cause de (17), $f(Z_0) \subset E$, ce qui donne aussitôt la seconde moitié de (4).

Il nous reste à vérifier (3) (les estimations bihöldériennes) et la première moitié de (4) (la surjectivité locale). On va avoir besoin d'un second énoncé à vérifier par récurrence.

Posons $\Gamma_n = f_n(Z_0)$. Cette fois, on va montrer que pour tout n et tout $i \in I_n$ tel que

$$|x_i| \leq 3 + 2^{-n+1} \tag{23}$$

Γ_n passe à distance $\leq 2^{-n-10}$ de x_i et coïncide sur $3B_i$ avec le graphe d'une fonction $C_3\varepsilon$-lipschitzienne φ_i définie sur P_i et à valeurs dans $P_i^\perp$.

C'est assez clair pour f_0. Soit $n \geq 0$, supposons la chose vraie pour n, et passons à $n+1$. Soit $i \in I_{n+1}$ vérifiant (23).

Choisissons $k \in I_n$ tel que $x_i \in \overline{B}_k$; c'est possible parce que $x_i \in E_0$ et par (6). Notons que si x_i vérifie (23), alors x_k le vérifie aussi (à l'ordre n). Donc l'hypothèse de récurrence dit que sur $3B_k$ (qui au passage contient largement $3B_i$), Γ_n est un bon graphe lipschitzien qui passe près de x_k.

Noter que $\Gamma_{n+1} = g_n(\Gamma_n)$, et que sur Γ_n la fonction g_n bouge les points de moins de $C_2\varepsilon 2^{-n}$ (par (18)). C'est aussi une fonction de classe C^∞, et on va s'intéresser à sa dérivée Dg_n dans $3B_k$. La formule (16) donne (dans D_n qui par (17) contient un petit voisinage de Γ_n)

$$Dg_n(y) = \sum_{j\in I_n} \theta_j(y) D\pi_j(y) + \sum_{j\in I_n} D\theta_j(y)[\pi_j(y) - \pi_k(y)], \tag{24}$$

où l'on a profité de ce que $\sum_{j\in I_n} D\theta_j(y) = 0$ pour retirer la "constante" $\pi_k(y)$. On ne regarde que quand $y \in 3B_k$, ce qui permet d'appliquer le lemme 1 à k et à tous les $j \in I_n$ qui ont une contribution non nulle. Ainsi, la norme de la seconde somme est $\leq C\varepsilon$. Comme $|D\pi_j(y) - D\pi_k(y)| \leq C\varepsilon$ (toujours par le lemme 1), on obtient finalement que

$$|Dg_n(y) - D\pi_k(y)| \leq C\varepsilon \text{ sur } 3B_k, \tag{25}$$

avec une constante C qui ne dépend pas de C_3. C'est bien, parce que ceci nous dit que g_n va avoir tendance à raplatir encore les points dans la direction de P_k (ou, ce qui revient en gros au même à cause du lemme 1, dans la direction de π_j). Les observations qui suivent sont peut-être un peu pénibles; le but en est de vérifier au passage la surjectivité de la projection sur l'image du graphe.

Par hypothèse de récurrence, Γ_n coïncide sur $3B_k$ avec le graphe d'une certaine fonction $C_3\varepsilon$-lipschitzienne $\varphi_k : P_k \to P_k^\perp$. On paramètre le graphe par la fonction $\psi_k(z) = z + \varphi_k(z)$. Ensuite, on note $D = P_k \cap \rho B_k$, avec ρ juste un peu plus petit que 3, par exemple, $\rho = 29/10$. Notre hypothèse sur le fait que Γ_n passe très près de x_k, plus la description lipschitzienne, font que $\psi_k(z) \in 3B_k$ pour $z \in D$, ce qui permet d'utiliser (25).

Donc on trouve que sur D, la différentielle de $g_k \circ \psi_k$ est $C\varepsilon$-proche de l'identité (qui coïncide avec $D\pi_k$ sur l'espace tangent P'_k à P_k). Et c'est encore pareil pour la différentielle de la projection $\pi_i \circ g_k \circ \psi_k$ puisque le lemme 1 dit que $D\pi_i - D\pi_k$ est petit. En intégrant, on trouve aussi que sur D,

$$|\pi_i \circ g_k \circ \psi_k(y) - y - a| \leq C\varepsilon 2^{-n}, \tag{26}$$

où l'on a posé $a = \pi_i \circ g_k \circ \psi_k(x_k) - x_k$.

Vérifions que $|a| \leq C\varepsilon 2^{-n}$. Notons que $\psi_k(x_k) \in B_k$ parce que Γ_n passe près de x_k, donc $\mathrm{dist}(\psi_k(x_k), P_k) \leq C\varepsilon 2^{-n}$ (par (17) et (1)), donc $|\psi_k(x_k) - x_k| \leq C\varepsilon 2^{-n}$ (parce que x_k est la projection orthogonale de $\psi_k(x_k)$ par définition de ψ_k), donc $|g_k \circ \psi_k(x_k) - x_k| \leq C\varepsilon 2^{-n}$ par (18) et parce que $\psi_k(x_k) \in \Gamma_n$, et finalement $|a| \leq C\varepsilon 2^{-n}$ par (17) et (1). Donc

$$|\pi_i \circ g_k \circ \psi_k(y) - y| \leq C\varepsilon 2^{-n} \quad \text{pour } y \in D. \tag{27}$$

Par un argument de degré (ou d'homotopie) vu plus haut, on en déduit que l'image de D par $h = \pi_i \circ g_k \circ \psi_k$ contient $P_i \cap 3B_i$. [On se souvient que $3B_i$ tombe nettement à l'intérieur de $3B_k$, et que $D = P_k \cap \rho B_k$ avec $\rho = 29/10$.]

De plus, le fait que la différentielle de h est sur D est $C\varepsilon$-proche de l'identité implique que h est injective, avec un inverse 2-lipschitzien h^{-1} défini sur un voisinage de $P_i \cap 3B_i$ dans P_i.

Finalement on pose $\psi_i = g_k \circ \psi_k \circ h^{-1}$ sur le voisinage de $P_i \cap 3B_i$. C'est une application Lipschitzienne, donc aussi $\varphi_i = \pi_i^\perp \circ g_k \circ \psi_k \circ h^{-1}$. De plus φ_i est $C\varepsilon$-Lipschitzienne sur $P_i \cap 3B_i$, sa dérivée étant obtenue en composant celle de $\psi_k \circ h^{-1}$ (qui est 3-Lipschitzienne), puis la composée de Dg_k (qui par (25) est très proche de $D\pi_k$ et de $D\pi_i$) et de $D\pi_i^\perp$. En plus C ne dépend toujours pas de C_3 (puisqu'il vient directement de (25) et du lemme 1). On aura juste utilisé l'hypothèse de récurrence pour prouver que h^{-1} est définie et 2-Lipschitzienne.

Noter que ψ_i est bien le paramétrage naturel du graphe G de φ_i, au sens où $\pi_i \circ \psi_i$ est bien l'identité, et $\pi_i^\perp \circ \psi_i = \varphi_i$. On aura fini quand on aura montré que G coïncide avec Γ_{n+1} sur $3B_i$, et qu'il passe près de x_i.

La partie facile: $G \cap 3B_i \subset \Gamma_{n+1}$, parce que si $z \in G$, alors $z = \psi_i(y)$ pour un $y \in P_i \cap 3B_i$, alors $z = g_n \circ \psi_k \circ h^{-1}(y)$. Mais $\psi_k(h^{-1}(y)) \in \Gamma_n$, donc $z \in \Gamma_{n+1}$.

Pour la réciproque, on doit faire un peu de chasse aux diagrammes. On se donne $z \in \Gamma_{n+1} \cap 3B_i$, et on doit vérifier que $z \in G$. Soit $y \in \Gamma_n$ tel que $g_n(y) = z$. Par (18), y est très proche de $3B_i$, donc il est dans $3B_k$. Ainsi, $y = \psi_k(w)$ pour un $w \in P_k \cap 3B_k$.

En fait, comme y est très proche de $3B_i$, $w \in D$ (sinon $\psi_k(w)$ tomberait trop loin de $3B_i$; se souvenir que $3B_i$ tombe bien à l'intérieur de $3B_k$, et que $|\psi_k(y)| \leq (2^{-10} + C\varepsilon)2^{-n}$ par hypothèse de récurrence). On pose $\xi = h(w)$; donc $w = h^{-1}(\xi)$ (par injectivité), et alors $z = g_n(y) = g_n(\psi_k(w)) = g_n \circ \psi_k \circ h^{-1}(\xi) = \psi_i(\xi)$. Comme $z \in 3B_i$ (et $\pi_i \circ \psi_i = h \circ h^{-1}$ est l'identité sur le domaine de définition de h^{-1}), on en déduit que $\xi \in 3B_i$, et donc $z \in G$ comme souhaité!

Ce qui termine notre vérification par récurrence sur les graphes lipschitziens.

On déduit de la description ci-dessous que tout point $x \in E \cap B(0,3)$ est à distance au plus 2^{-n+1} de $\Gamma_n = f_n(Z_0)$. En passant à la limite (et en utilisant la compacité de $\overline{Z}_0$), on trouve que $x \in f(\overline{Z}_0)$, et ceci termine la démonstration de (4).

Reste à voir que (3) est vrai. On se donne x et y dans $Z \cap B(0,2)$ et on essaie de suivre ce qui arrive à $x_n - y_n$, où $x_n = f_n(x)$ et $y_n = f_n(y)$. Si $|x-y| \geq 2^{-3}$, disons, alors (22) dit que $\big||f(y)-f(x)| - |x-y|\big| \leq C\varepsilon$, donc on n'a pas de problème pour (3). Supposons donc que $|x-y| \leq 2^{-3}$.

Pour les premières valeurs de n, $|x_n - y_n| \leq 2^{-n-2}$. Dans ce cas, il existe un $k \in I_n$ tel que x_n et y_n soient tous les deux dans $2B_k$. Comme $x \in B(0,2)$ et $|x_n - x| \leq C\varepsilon 2^{-n}$ par (22), on a bien (23) pour k, donc x_n et y_n sont sur le graphe de la fonction φ_k. Alors les calculs faits plus hauts sont valides, et en particulier (25).

Pour ne pas avoir à intégrer directement sur un graphe, utilisons le paramétrage par la fonction ψ_k, dont la différentielle est $C\varepsilon$-proche de l'identité (restreinte à P_k'), de sorte que la différentielle de $g_n \circ \psi_k$ est aussi $C\varepsilon$-proche de l'identité. [Se souvenir que $\pi_k = I$ sur P_k'.] On intègre sur un segment qui mène de $\pi_k(x_n)$ à $\pi_k(y_n)$ et on trouve que

$$\begin{aligned}\big||g_n(x_n)-&g_n(y_n)| - |\pi_k(x_n)-\pi_k(y_n)|\big| \\ &= \big||g_n\circ\psi_k\circ\pi_k(x_n) - g_n\circ\psi_k\circ\pi_k(y_n)| - |\pi_k(x_n)-\pi_k(y_n)|\big| \\ &\le C\varepsilon|\pi_k(x_n)-\pi_k(y_n)|\end{aligned}$$

puisque $x_n = \psi_k\circ\pi_k(y_n)$ et pareil pour y_n. On a aussi que $|x_n - y_n| = |\psi_k\circ\pi_k(x_n)-\psi_k\circ\pi_k(y_n)|$ est $C\varepsilon|\pi_k(x_n)-\pi_k(x_n)|$-proche de $|\pi_k(x_n)-\pi_k(x_n)|$, puisque φ_k est $C\varepsilon$-lipschitzienne. Donc finalement, puisque $|x_{n+1}-y_{n+1}| = |g_n(x_n)-g_n(y_n)|$,

$$\big||x_{n+1}-y_{n+1}| - |x_n - y_n|\big| \le C\varepsilon|x_n-y_n|. \tag{28}$$

Autrement dit, en posant $d_n = |x_n - y_n|$, on a que

$$d_{n+1}/d_n \in [1-C\varepsilon, 1+C\varepsilon]. \tag{29}$$

Par contre, à chaque étape 2^{-n} est carrément divisé par 2. Donc il vient un moment où $d_n > 2^{n-2}$. A ce moment, on décide d'utiliser l'estimation (18), et en itérant, il vient $|f(x)-x_n| \le C\varepsilon 2^{-n}$, et pareil pour y, de sorte qu'en comparant $\big||f(x)-f(y)| - d_n\big| \le C\varepsilon d_n$ (pour cet n là, que du coup on va appeler m). Bref, $|f(x)-f(y)|$ est très proche de d_m, et il ne reste plus qu'à voir que

$$(1-\tau/2)|x-y|^{1+\tau} \le d_m \le (1+\tau/2)|x-y|^{1-\tau} \text{ pour } x,y\in B(0,2) \tag{30}$$

Pour estimer d_m, on note qu'à cause de (24), on a $(1-C\varepsilon)^m \le d_m/|x-y| \le (1+C\varepsilon)^m$, donc $|\log_2(d_m/|x-y|)| \le C\varepsilon m$. En même temps, 2^{-m} est de l'ordre de d_m, donc $|\log_2(d_m) - m| \le 2$. De sorte que, pour commencer par une majoration,

$$\log_2(d_m) \le \log_2(|x-y|) + C\varepsilon m \le \log_2(|x-y|) + C\varepsilon\log_2(d_m) + C\varepsilon,$$

qui donne $\log_2(d_m) \le (1-C\varepsilon)^{-1}\log_2(|x-y|) + C\varepsilon$, pous en prenant des puissances, $d_m \le (1+C\varepsilon)|x-y|^{1-C\varepsilon}$, qui bien sûr implique la seconde partie de (30). L'autre morceau se fait pareillement.

Tout ceci termine notre vérification pour ce qui est du paramétrage de E: pour étendre ce paramétrage à l'espace ambient, on doit compléter notre recouvrement de E par un recouvrement avec des boules plus petites de la partie de $\mathbb{R}^N$ qui est loin de E, et modifier notre définition des f_n et des g_n en ajoutant des termes correctifs (pour ne rien changer sur Z, mais éviter d'écraser sur E les points qui ne sont pas dans Z par des applications qui ressembleraient trop à des projections). □

22. COURBURE DE MENGER ET NOYAU DE CAUCHY

Pas de notes sur ce sujet, mais plan d'exposé possible:
- définition de la courbure
- la formule magique
- le calcul approché de $T(1)$ en termes de courbure pour les mesures à croissance linéaire
- la conséquence en termes de caractérisation par la courbure
- un calcul pour dire que la courbure est dominée pour un graphe lipschitzien
- le fait que ca caractérise les ensembles uniformément rectifiables (dans le cas Ahlfors-régulier)

23. REPRESENTATIONS CONFORMES

Observation. Si $f : \mathbb{C} \to \mathbb{C}$ est analytique, avec $f'(0) \neq 0$, et si deux courbes γ_1 et γ_2 de classe C^1 passent par 0 et y font un angle α, alors $f(\gamma_1)$ et $f(\gamma_2)$ se croisent en $f(0)$ et y font le même angle α.

C'est facile à vérifier, puisque f, vue commme application de $\mathbb{R}^2$ dans $\mathbb{R}^2$ a une différentielle, dont l'action est la composée d'une rotation et d'une homotétie.

Une application conforme est une application d'un domaine de $\mathbb{R}^n$ dans $\mathbb{R}^n$, qui ne change pas les angles (de courbes là où elles se croisent). Pour $n > 2$, il se trouve qu'il y a très peu d'applications conformes, alors que quand $n = 2$, les applications (provenant par l'identification de $\mathbb{R}^2$ avec $\mathbb{C}$ de fonctions) analytiques dont la dérivée ne s'annule pas le sont.

On appellera ici représentation conforme une application analytique bijective d'un domaine de $\mathbb{C}$ dans un autre. [Noter que l'injectivité force la dérivée à ne pas s'annuler, et d'ailleurs que les angles sont doublés, ou plus, en un point où la dérivée s'annulle.]

On va rappeler divers fait sur les fonctions analytiques (en dimension 1), pour arriver au théorème de Riemann sur l'existence de représentations conformes.

a. Lemme de Schwarz

On notera $\mathbb{D}$ le disque unité ouvert dans $\mathbb{C}$.

Lemme. *Soit $f : \mathbb{D} \to \mathbb{D}$ une application analytique, avec $f(0) = 0$. Alors $|f(z)| \leq |z|$ sur $\mathbb{D}$, et en particulier $|f'(0)| \leq 1$. De plus, si $|f'(0)| = 1$ ou si $|f(z)| = |z|$ pour un $z \in \mathbb{D}$, alors $f(z) = \lambda z$ pour un $\lambda \in \mathbb{C}$ (qui est donc de module 1).*

Démonstration classique. On vérifie que $g(z) = f(z)/z$, prolongée par $g(0) = f'(0)$, est une fonction analytique sur $\mathbb{D}$, pour cause de singularité effaçable en 0. Son module dans $B(0, 1-\varepsilon)$ est inférieur à $(1-\varepsilon)^{-1}$ à cause du principe du maximum et puisque $|f(z)| \leq 1$ sur $\mathbb{D}$. Donc en fait $|g(z)| \leq 1$ partout, et en particulier $|f'(0)| = |g(0)| \leq 1$.

En cas d'égalité, le principe du maximum (appliqué à g sur un disque fermé assez grand) dit que g est constante, donc $f(z) = \lambda z$. $\square$

b. Transformations de Möbius

Pour θ réel et $z_0 \in \mathbb{D}$, on définit $\tau = \tau_{\theta,z_0}$ par $\tau(z) = e^{i\theta}\,\dfrac{z - z_0}{1 - \overline{z}_0 z}$.

Lemme. *τ_{θ,z_0} est une fonction analytique bijective de $\mathbb{D}$ dans $\mathbb{D}$. Et d'ailleurs toute fonction analytique bijective de $\mathbb{D}$ dans $\mathbb{D}$ est un τ_{θ,z_0}.*

D'abord, τ_{θ,z_0} est analytique sur $\mathbb{D}$. Elle a une extension continue au bord (et même une extension analytique un peu au-delà).

On peut calculer l'inverse de τ_{0,z_0}. On veut résoudre $w = \frac{z-z_0}{1-\overline{z}_0 z}$, donc $(1 - \overline{z}_0 z)w = z - z_0$ ou encore $z(1 + \overline{z}_0 w) = w + z_0$, et donc on trouve que pour $w \neq -1/\overline{z}_0$, il y a une seule solution, à savoir $z = \tau_{0,-z_0}(w)$.

Noter que quand $|z| = 1$, le module du dénominateur est $|1 - \overline{z}_0 z| = |\overline{z} - \overline{z}_0| = |z - z_0|$ (en multipliant par $\overline{z}$), donc $|\tau(z)| = 1$ pour $|z| = 1$, et τ envoie le cercle dans le cercle. En fait, sur le cercle, l'inverse de τ_{0,z_0} est bien défini (calcul plus haut), et est aussi une transformation de Möbius, qui donc envoie le cercle dans le cercle. Bref, la restriction de τ au cercle est une bijection du cercle.

Ensuite, $\tau(\mathbb{D}) \subset \mathbb{D}$, vu que $\tau(0) = -e^{i\theta} z_0 \in \mathbb{D}$ et par le lemme de traversée des frontières: si γ est un segment qui mène de 0 à un autre point du disque, l'image de γ par τ ne peut rencontrer le cercle (puisque les soultions de $\tau(z) = w$, avec w dans le cercles, sont toutes dans le cercle.

Donc dans la formule d'inversion, comme l'inverse aussi envoie le disque dans lui-même, on trouve que τ est est bijective sur $\mathbb{D}$, avec pour inverse la même composition de rotation et de $\tau_{0,-z_0}$. Que ne ferait-on pas pour éviter un petit calcul!

Il reste à vérifier que toutes les bijections holomorphes de $\mathbb{D}$ sont des τ_{θ,z_0}. Soit φ un tel isomporphsime. Soit $z_0 = f(0)$ et posons $\psi = \tau_{0,z_0} \circ \varphi$. C'est encore un isomorphisme, et maintenant $\psi(0) = 0$. Par Schwarz, $|\psi'(0)| \leq 1$. Par Schwarz appliqué à ψ^{-1}, $|\psi'(0)| \geq 1$. Par le cas d'égalité, $\psi(z) = e^{i\theta} z$ pour un $\theta \in \mathbb{R}$. Donc $\varphi = \tau_{0,-z_0} \circ \psi$, ce qui donne sauf erreur $\tau(z) = \dfrac{e^{i\theta} z + z_0}{1 + \overline{z}_0 e^{i\theta} z} = e^{i\theta} \dfrac{z + e^{-i\theta} z_0}{1 + \overline{z}_0 e^{i\theta} z} = \tau_{\theta,z_1}(z)$, avec $z_1 = -e^{-i\theta} z_0$. □

Exercice (peu intéressant?) On sait donc à cause du lemme que la composée de deux transformées de Möbius est une transformée de Möbius. Le vérifier à la main.

c. Lemme de Schwarz, version invariante

Pour pas plus cher que le lemme de Schwarz, on a le résultat suivant, qui coincide avec le lemme de Schwarz quand $z_0 = f(z_0) = 0$.

Lemme. *Soit $f : \mathbb{D} \to \mathbb{D}$ une application analytique. Alors*

$$(a) \qquad \left| \frac{f(z) - f(z_0)}{1 - \overline{f(z_0)} f(z)} \right| \leq \left| \frac{z - z_0}{1 - \overline{z}_0 z} \right| \quad \textit{pour } z, z_0 \in \mathbb{D}$$

et

$$(b) \qquad \frac{|f'(z)|}{1 - |f(z)|^2} \leq \frac{1}{1 - |z|^2} \quad \textit{pour } z \in \mathbb{D}.$$

Si de plus il y a égalité avec un $z \neq z_0$ dans (a), ou un $z \in \mathbb{D}$, alors f est une transformation de Möbius.

Pour la démonstration, on pose $F = \tau_{0,f(z_0)} \circ f \circ \tau_{0,-z_0}$. Noter que F est analytique, envoie $\mathbb{D}$ dans $\mathbb{D}$, et $F(0) = \tau_{0,f(z_0)} \circ f(z_0) = 0$, donc $|F(w)| \leq |w|$ and $|F'(0)| \leq 1$.

La première conclusion s'écrit aussi, en posant $w = \tau_{0,z_0}(z)$, $|\tau_{0,f(z_0)} \circ f(z)| \leq |\tau_{0,z_0}(z)|$, ce qui est (a). Et (b) pour z_0 est la limite de (a) quand z tend vers z_0, ou s'obtient en disant que $|F'(0)| \leq 1$. D'où aussi la gestion du cas limite, qui se produit seulement quand F est une rotation. □

d. Familles normales

On va noter $H(\Omega)$ l'ensemble des fonctions holomorphes (ou analytiques, je dirai les deux indifféremment) définies sur l'ouvert Ω (souvent supposé connexe) et à valeurs dans $\mathbb{C}$.

Définition. *Une partie $\mathcal{F}$ de $H(\Omega)$ est appelée famille normale si pour toute suite $\{f_k\}$ dans $\mathcal{F}$, on peut extraire une sous-suite qui converge, uniformémént sur tout compact de Ω.*

La notion dépend évidemment de Ω; il sera plus facile d'être une famille normale sur un ouvert plus petit. Ainsi, les z^k, $k \geq 0$, sont une famille normale sur $B(0,1)$ (toutes les sous-suites convergent vers 0, mais pas sur un ouvert srtictement plus grand.

Donc une famille normale est une famille séquenciellement relativement compacte pour la topologie de la convergence uniforme sur tout compact. La notion est très pratique, à cause du fait que la topologie en question est équivalente à beaucoup d'autres, et aussi parce qu'il est facile de trouver des familles normales.

Il existe une variante de la définition pour des fonctions méromorphes (à valeurs dans la sphère de Riemann), mais on n'en parlera pas ici.

Proposition. *Si pour tout $K \subset \Omega$ compact, il existe $C_K \geq 0$ tel que $|f(z)| \leq C_K$ pour $z \in K$ et $f \in \mathcal{F}$, alors $\mathcal{F}$ est une famille normale.*

Plan de démonstration. On commence par recouvrir Ω une famille de boules. On va noter $\mathcal{D}$ l'ensemble des $x \in \Omega$ dont les coordonnées sont rationnelles, on numérote les $x \in \mathcal{D}$, et on note x_k le k-ième élément de $\mathcal{D}$. Bref, $\mathcal{D} = \{x_k\,;\, k \geq 1\}$. Pour $k \geq 1$, on pose $r_k = \frac{1}{3}\operatorname{dist}(x_k, \mathbb{C} \setminus \Omega)$, sauf si $\Omega = \mathbb{C}$ auquel cas on prend $r_k = 1$. Et enfin on pose $B_k = B(x_k, r_k)$.

Par hypothèse, il existe $C_k \geq 0$ tel que $|f(z)| \leq C_k$ pour $z \in \overline{B}(x_k, 2r_k)$. Du coup, on a aussi que

$$(lip) \qquad \text{chaque } f \in \mathcal{F} \text{ est } 10C_k r_k^{-1}\text{-Lipschitzienne sur } B_k,$$

par exemple en utilisant la formule de Cauchy basée sur le cercle $\partial B(x_k, 2r_k)$, et en dérivant sous le signe intégral.

Soit $\{f_n\}$ une suite dans $\mathcal{F}$. Comme pour tout k, $|f_n(x_k)| \leq C_k$, la suite $\{f_n(x_k)\}$ est bornée, et on peut en extraire une sous-suite convergente. Par le procédé diagonal, on trouve une sous-suite, qu'on notera encore $\{f_n\}$, telle que

$f(x_k) = \lim_{n\to+\infty} f_n(x_k)$ existe pour tout k. Par (lip), f est $10C_k r(x_k)^{-1}$-Lipschitzienne sur $\mathcal{D}\cap B_k$, et admet donc une extension lipschitzienne à B_k, que l'on notera encore f. On vérifie aiément que

$$(r) \qquad \bigcup_{k\geq 1} B_k = \Omega$$

(à cause du choix spécifique de r_k; ce serait faux si on avait pris des r_k trop petits). Du coup, il existe en fait un unique prolongement de f (initialement définie sur $\mathcal{D}$) à Ω.

En fait, par l'argument usuel utilisant la démonstration du théorème de Montel, la suite $\{f_n\}$ converge vers f uniformément sur B_k: pour tout $\varepsilon > 0$, on choisit une partie finie $A \subset \mathcal{D}\cap B_k$, telle que tout $x \in B_k$ est à distance $\leq \varepsilon$ de A. Alors, pour tout n assez grand pour que $|f(a) - f_n(a)| \leq \varepsilon$ pour $a \in A$, on a automatiquement que pour $x \in B_k$, $|f(x) - f_n(x)| \leq |f(x) - f(a)| + |f(a) - f_n(a)| + |f_n(a) - f_n(x)| \leq 10C_k r_k^{-1}\varepsilon + \varepsilon + 10C_k r_k^{-1}\varepsilon$.

D'où la convergence uniforme sur B_k. En conséquence, f est analitique sur chaque B_k, donc sur Ω.

Finalement, pour la convergence uniforme sur un compact $K \subset \Omega$, on se contente de recouvrir K par un nombre fini de B_k (c'est possible par (r) et par compacité), et d'utiliser la convergence uniforme sur chaque B_k. $\square$

Notons encore que la convergence uniforme sur tout compact de Ω, entraîne bien d'autres choses, comme la convergence uniforme sur tout compact des dérivées, etc. Ainsi, si $\mathcal{F}$ est une famille normale sur Ω, $\mathcal{F}' = \{f' \,;\, f \in \mathcal{F}\}$ est une famille normale sur Ω.

e. Domaines simplement connexes

Définition. *L'ensemble $\Omega \subset \mathbb{R}^n$ est dit simplement connexe quand tout lacet tracé dans Ω est homotope (pami les lacets tracés dans Ω) à un lacet trivial.*

Le plus souvent, on utilisera ceci pour un ouvert Ω, d'ailleurs supposé connexe. [S'il ne l'est pas, on demande simplement que toutes les composantes connexes soient simplement connexes.] Lacet trivial signifie lacet constant. Donc, on demande que pour tout $\gamma : \mathbb{T} \to \Omega$ continu (et où $\mathbb{T}$ est le tore $\frac{\mathbb{R}}{\mathbb{Z}}$ (ou le cercle si vous préférez)), il existe $F : \mathbb{T}\times[0,1] \to \Omega$, tel que $F(x,0) = \gamma(x)$ pour $x \in \mathbb{T}$, et $F(x,1) = a$ pour un $a \in \Omega$.

Evidemment les spécialistes auront reconnu la condition que $\pi_1(\Omega) = \{0\}$.

Passons aux domaines (connexes) simplement connexes de $\mathbb{C}$, où la simple connexité nous permet de trouver des primitives. On utilisera la convention standard que "domaine" signifie "ouvert connexe".

Proposition. *Soit $\Omega \subset \mathbb{C}$ un domaine simplement connexe, et soit $f \in H(\Omega)$. Alors il existe $g \in H(\Omega)$ telle que $g' = f$.*

Contrexemple où Ω n'est pas simplement connexe: prendre $\Omega = B(0,2) \setminus \overline{B}(0,1/2)$, et noter que $f(z) = 1/z$ n'a pas de primitive, car il faudrait que localement, $f(e^{it} = t + Cte$.

Rappel rapide sur la démonstration. On se donne $z_0 \in \Omega$, et on essaie la formule

$$g(w) = \int_\gamma f(z)dz = \int_{[0,1]} f(\gamma(t))\gamma'(t)dt \tag{p}$$

où la première formule est un abus de langage qui donne une idée de ce qu'on fait, et où (pour la seconde formule) $\gamma : [0,1] \to \Omega$ est un paramétrage de classe C^1 par morceaux d'un chemin (noté abusivement γ dans la première formule) tel que $\gamma(0) = z_0$ et $\gamma(1) = w$.

La vérification importante à faire est que le résultat dans (p) ne dépend pas du chemin choisi. Une fois cette vérification faite, on montre aisément que $g(z)$ est holomorphe dans Ω, de dérivée f. [Noter simplement que pour w' proche de w, $g(w') - g(w) = \int_{[0,1]} f(w + t(w'-w))(w'-w)dt$, diviser par $w' - w$, et prendre la limite.]

Pour la vérification, on doit juste montrer que si γ est un lacet de classe C^1 par morceaux, $\int_\gamma f(z)dz = 0$. Notons également γ son paramétrage par $\mathbb{T}$. Comme Ω est simplement connexe, il existe $F : \mathbb{T} \times [0,1] \to \Omega$ tel que $F(x,0) = \gamma(x)$ pour $x \in \mathbb{T}$ et $F(x,1) = a$ pour un $a \in \Omega$.

Supposons d'abord que γ et F sont tous deux de classe C^1. On observe alors que

$$\frac{\partial}{\partial s}\int_{\mathbb{T}} f(F(t,s))\,\frac{\partial F(t,s)}{\partial t}dt = \int_{\mathbb{T}} f'(F(t,s))\,\frac{\partial F(t,s)}{\partial s}\,\frac{\partial F(t,s)}{\partial t}dt + \int_{\mathbb{T}} f(F(t,s))\,\frac{\partial^2 F(t,s)}{\partial s \partial t}$$

puis on intègre par parties (par rapport à t), en observant que justement la dérivée de $f(F(t,s))$ est $f'(F(t,s))\,\frac{\partial F(t,s)}{\partial t}$. Donc on trouve 0, ce qui signifie que $\int_\gamma f(z)dz = \int_{\mathbb{T}} f(F(t,1))\,\frac{\partial F(t,1)}{\partial t}dt = 0$, comme souhaité.

Pour passer au cas général on utilise un argument de régularisation de son choix. Par exemple, mais je pense qu'il y a plus court (et que votre démonstration préférée du théorème de Morera le fait) on peut d'abord régulariser γ, et se ramener au cas où γ est de classe C^1.

Ensuite, remplacons d'abord F par $\widetilde{F}$, où $\widetilde{F}(t,s) = F(t,0)$ pour $-\infty \leq t \leq 1/3$, $\widetilde{F}(t,s) = F(t, 3s-1)$ pour $1/3 \leq t \leq 2/3$, et $\widetilde{F}(t,s) = F(t,1) = Cte$ pour $2/3 \leq t < +\infty$. Juste pour donner de la marge. Ensuite, on remplace $\widetilde{F}$ par une convolution $F_\varepsilon = \widetilde{F} * \varphi_\varepsilon$, avec ε petit et où φ_ε est par exemple obtenue à partir d'un produit de fonctions bosses. On laisse le lecteur vérifier que la convolution a bien un sens sur $\mathbb{T} \times \mathbb{R}$ (ceci revient à vérifier que la convolution est périodique quand la fonction est périodique). La formule

$$\frac{\partial}{\partial s} \int_{\mathbb{T}} f(F_\varepsilon(t,s)) \, \frac{\partial F_\varepsilon(t,s)}{\partial t} dt = 0$$

vaut encore, comme plus haut. Pour $s = 0$, on trouve $\int_{\mathbb{T}} f(F_\varepsilon(t,s)) \, \frac{\partial F_\varepsilon(t,s)}{\partial t} dt = 0$ comme plus haut (puisque $F_\varepsilon(t,0)$ est constante, donc $\frac{\partial F_\varepsilon(t,0)}{\partial t} = 0$). Pour $s = 1$ on trouve

$$\int_{\mathbb{T}} f(F_\varepsilon(t,1)) \, \frac{\partial F_\varepsilon(t,1)}{\partial t} dt,$$

qui tend bien vers $\int_{\mathbb{T}} f(\gamma(t)) \, \gamma'(t) dt$ quand ε tend vers 0, puisque pour ε petit, $F_\varepsilon(t,1)$ n'est plus que la convolution en une variable de $\gamma(t)$. Donc $\int_{\mathbb{T}} f(\gamma(t)) \, \gamma'(t) dt = 0$, et la proposition s'en déduit. $\square$

Corollaire 1 (existence d'un logarithme). *Si Ω est un domaine simplement connexe et $f \in H(\Omega)$ ne s'annulle pas, alors il existe $g \in H(\Omega)$ telle que $e^g = f$ sur Ω.*

Souvent utilisé directement avec $f(z) = z$, d'ailleurs. La démonstration est facile. On choisit une primitive G de $1/f$, et on note que la dérivée de $e^{-G} f$ est $-G' e^{-G} f + e^{-G} f' = 0$. Donc (par connexité de Ω), $e^{-G} f$ est constante. Autrement dit, $e^G = Cf$. On prend $g = G + \lambda$, où la constante λ est choisie telle que $e^\lambda = 1/C$. $\square$

Corollaire 2 (existence d'une racine). *Si Ω est un domaine simplement connexe et $f \in H(\Omega)$ ne s'annulle pas, alors il existe $h \in H(\Omega)$ telle que $h^2 = f$ sur Ω.*

Prendre $h = e^{g/2}$, où g est comme au corollaire 1. $\square$

En fait l'existence (automatique) de logarithmes, ou de racines, est équivalente à la simple connexité de Ω. Voir Rudin par exemple, mais on ne s'en servia pas.

f. Théorème de représentation de Riemann

Théorème (Riemann). *Soit Ω un domaine simplement connexe du plan, non vide et différent de $\mathbb{C}$. Soit $z \in \Omega$. Alors il existe une unique application analytique bijective $f : \Omega \to \mathbb{D}$, telle que $f(z_0) = 0$ et $f'(z_0) > 0$.*

Les autres application analytiques bijective de $\Omega \to \mathbb{D}$ (encore appelées représentations conformes) sont obtenues par composition avec des applications de Möbius du disque: si $g : \Omega \to \mathbb{D}$ est une représentation conforme, on note que $g \circ f^{-1}$ est un ismoprphisme (conforme) du disque, donc est une application de Möbius τ. Et donc $g = g \circ f^{-1} \circ f = \tau \circ f$.

L'unicité vient de ce que si de plus $g(z_0) = f(z_0) = 0$, il vient $\tau(0) = 0$, donc $\tau(z) = \lambda z$. Si de plus $g'(z_0)$ et $f'(z_0)$ son positif, $\tau'(0) = g'(z_0)/f'(z_0) > 0$, ce qui donne $\lambda = 1$.

C'est donc l'existence qui est la plus délicate et intéressante. Noter que quand $\Omega = \mathbb{C}$, on ne peut espérer trouver f, puisque par Liouville les fonctions holomorphes bornées sur $\mathbb{C}$ sont constantes. Revenons à $\Omega \neq \mathbb{C}$. On va obtenir l'unicité par compacité, en maximisant $f'(z_0)$ chez les $f \in H(\Omega)$ injectives telles que $f(\Omega) \subset \mathbb{D}$. On pose donc

$$(a) \qquad \mathcal{F} = \big\{ f \in H(\Omega) \,:\, f \text{ est injective et } f(\Omega) \subset \mathbb{D} \big\}.$$

Lemme 1. $\mathcal{F} \neq \emptyset$.

Si Ω ne rencontrait pas un disque $B(a, r)$, ça serait facile: on prendrait $f(z) = \frac{r}{z-a}$ et le tour serait joué. On va se ramener à ce cas par une astuce.

Notons que par hypothèse in existe $w \in \mathbb{C} \setminus \Omega$. On applique le corollaire 2 pour trouver une fonction racine h, analytique sur Ω, telle que $g^2(z) = z - w$. On va voir que prendre la racine permet d'écarter les bords pour faire de la place. Penser au cas où $\Omega = \mathbb{C} \setminus \mathbb{R}_+$.

Notons que g est injective: si z_1 et z_2 sont tels que $g(z_1) = g(z_2)$, alors $z_1 - w = z_2 - w$ (on prend le carré), donc $z_1 = z_2$.

Le même raisonnement dit aussi que si z_1 et z_2 dans Ω sont tels que $g(z_1) = -g(z_2)$, alors $z_1 = z_2$, donc $g(z_1) = 0$, ce qui est impossible car alors $0 = g(z_1)^2 = z_1 - w$, et $w \notin \Omega$. Bref, $g(\Omega) \cap -g(\Omega) = \emptyset$.

Mais g n'est pas constante (elle est injective), donc elle est ouverte, et en particulier $g(\Omega)$ contient un disque $B(a,r)$. Mais alors $B(-a,r)$ ne rencontre pas $g(\Omega)$, et on en profite pour prendre $f(z) = \frac{r}{g(z)+a}$, qui est bien définie, holomorphe, et à valeurs dans $\mathbb{D}$ par ce qu'on vient de dire. Il reste à voir qu'elle est injective, mais c'est clair parceque g l'est. Notons au passage que $g'(z_0) \neq 0$ puisque $2g(z_0)g'(z_0) = 1$ (vu que $g(z)^2 = z - w$), donc aussi $f'(z_0) \neq 0$. $\square$

Lemme 2. *Posons* $M = \sup\{|f'(z_0)|\,;\, f \in \mathcal{F}\}$. *Alors* $0 < M < +\infty$, *et il existe* $f \in \mathcal{F}$ *telle que* $f'(z_0) = M$.

D'abord, $0 < M < +\infty$. Pour la première inégalité, on vient de trouver $f \in \mathcal{F}$ telle que $f'(z_0) \neq 0$, donc $M > 0$. Vérifions maintenant que $M \leq d^{-1}$, où d est la distance de z_0 à la frontière de Ω. Ainsi $B(z_0, d) \subset \Omega$, et si $f \in \mathcal{F}$, la fonction $g(z) = f(z_0 + dz)$ est analytique de $\mathbb{D}$ dans $\mathbb{D}$, donc le lemme de Schwarz (dans sa version délocalisée) donne $f'(z_0) = d^{-1}g'(0) \leq d^{-1}$. Ainsi, $M \leq d^{-1}$, comme promis.

Pour tout $k \geq 0$, choisissons $f_k \in \mathcal{F}$ telle que $|f_k'(z_0)| \geq M - 2^{-k}$. Quitte à composer par une rotation, supposons même que ($f'(z_0)$ est un réel positif et) $f'(z_0) \geq M - 2^{-k}$. Puisque les fonctions de $\mathcal{F}$ sont uniformément bornées (par 1), $\mathcal{F}$ est une famille normale (sur Ω), ce qui nous permet de choisir une sous-suite (que l'on appellera encore $\{f_k\}$), qui converge uniformément sur tout compact de Ω vers une limite $f \in H(\Omega)$.

Notons que $f'(z_0) = \lim_{k\to+\infty} f_k'(z_0) = M$ (les dérivées aussi convergent uniformément sur les compact, et on sait que $M \leq |f_k'(z_0)| \geq M - 2^{-k}$ pour tout k).

Il nous reste à prouver que $f \in \mathcal{F}$. D'abord, $|f(z)| \leq 1$ sur Ω, mais comme f n'est pas constante (puisque $f'(z_0) = M > 0$), elle n'atteint pas son maximum sur Ω, et donc $f(\Omega) \subset \mathbb{D}$.

Reste la partie en principe délicate: f est injective. Supposone que z_1, $z_2 \in \Omega$ sont distincts, mais $f(z_1) = f(z_2) = \xi$. Soit ρ petit, et pour $j = 1, 2$, considérons le lacet γ_j défini par $\gamma_j(\theta) = f(z_j + \rho\theta)$ pour $\theta \in \mathbb{S}^1$ (le cercle unité). Pour ρ assez petit, c'est un lacet d'indice n_j par rapport à ξ, où n_j est le numéro de la première dérivée non nulle de f en z_j. [C'est une propriété classique des fonctions analytiques non constantes que toutes leurs dérivées ne peuvent être nulles en un même point.] Ce qui signifie que γ_j tourne n_j fois autour de ξ, sans d'ailleurs passer par ξ. C'est facile quand on connaît la définition et les premières popriétés de l'indice, à savoir que l'indice de la courbe $\theta \to \xi + (\rho\theta)^{n_j}$ est n_j, et que l'indice reste le même quand on modifie un lacet γ par homotopie, sans qu'il soit autorisé à passer par ξ. Dans le cas qui nous occupe, $\gamma_j(\theta) = f(z_j + \rho\theta) = \xi + c_j(\rho\theta)^{n_j}(1 + o(\rho^{n_j}))$, ce qui donne une homotopie évidente en écrasant le terme $o(\rho^{n_j})$.

On choisit donc un ρ assez petit pour que l'indice de γ_j soit $n_j > 0$, et aussi que les $B(z_j, \rho)$ soient contenues dans Ω et disjointes. Ensuite, on note que pour k assez grand, $\widetilde{\gamma}_j(\theta) = f_k(z_j + \rho\theta)$ définit encore un lacet d'indice n_j par rapport à ξ. Et on utilise maintenant le fait que quand c'est le cas, ξ est dans l'image de $B(z_j, \rho)$ par f_k. Autrement, on définirait une homotopie qui va de $\widetilde{\gamma}_j$ à une constante en posant $g_t(\theta) = f_k(z_j + t\rho\theta)$ pour $0 \leq t \leq 1$, dont l'image ne rencontrerait pas ξ, ce qui est impossible parce que l'indice du lacet final constant par rapport à ξ est forcément nul. Pour un peu plus sur le même sujet (et par exemple qu'en fait l'équation $f_n(z) = \xi$ ci dessus a n_j racines dans $B(z_j, \rho)$, voir le théorème de Rouché, par exemple dans le livre de Rudin.

Notre contradiction vient de ce que $\xi \in f_k(B(z_1, \rho)) \cap f_k(B(z_1, \rho))$, contredisant l'injectivité de f_k. Bref, $f \in \mathcal{F}$. $\square$

Lemme 3. *Soit f comme au lemme 2. Alors $f(z_0) = 0$ et $f(\Omega) = U$.*

On en déduira le théorème, puiqu'on sait déjà que f est analytique et injective.

On va à nouveau avoir besoin des transformations de möbius τ_α définies pour $\alpha \in \mathbb{D}$ par $\tau_\alpha(z) = \frac{z-\alpha}{1-\overline{\alpha}z}$.

D'abord, si $f(z_0) = \alpha \neq 0$, on va voir que $g = \tau_\alpha \circ f$ fait mieux que f. C'est encore un élément de $\mathcal{F}$, puisque τ_α est une bijection holomorphe de $\mathbb{D}$, et de plus $|g'(z_0)| = |\tau'_\alpha(f(z_0))f'(z_0)| = |\tau'_\alpha(\alpha)f'(z_0)| > |f'(z_0)|$ car $|f'(z_0)| = M > 0$ et $\tau'_\alpha(\alpha) = (1 - |\alpha|^2)^{-1} > 1$. Ceci contredit la définition de f, et donc $f(z_0) = 0$.

Ensuite supposons que $f(\Omega) \neq U$, et montrons qu'il il existe $h \in \mathcal{F}$ telle que $|h'(z_0)| > |f'(z_0)|$; on en déduira une contradiction avec la définition de f, et la surjectivité de f et le lemme suivront.

Soit maintenant $\alpha \in \mathbb{D} \setminus f(\Omega)$, et utilisons à nouveau $\tau_\alpha(z) = \frac{z-\alpha}{1-\overline{\alpha}z}$. Alors $\tau_\alpha \circ f$ ne prend pas la valeur 0 (se souvenir que $\tau_\alpha : \mathbb{D} \to \mathbb{D}$ est bijective, et noter que $\tau_\alpha(\alpha) = 0$). Donc, puisque Ω est simplement connexe, on peut trouver $g \in H(\Omega)$ telle que $g^2 = \tau_\alpha \circ f$ sur Ω.

Notons $s(z) = z^2$. Ainsi, on vient de dire que $s \circ g = \tau_\alpha \circ f$ sur Ω. Posons encore $\beta = g(z_0)$, et essayons $h = \tau_\beta \circ g$. L'idée est peut-être encore que prendre la racine risque d'augmenter la taille de l'image, donc d'augmenter $h'(z_0)$, mais je n'en suis pas totalement certain et on va faire la vérification algébriquement. Notons que h est bien holomorphe à valeurs dans $\mathbb{D}$. De plus, g est injective, car si $g(z_1) = g(z_2)$, alors $g^2(z_1) = g^2(z_2)$, ce qui donne $z_1 = z_2$ puisque $g^2 = \tau_\alpha \circ f$ est injective. Donc $h \in \mathcal{F}$.

Il reste à calculer les dérivées. Noter que

$$f = \tau_\alpha^{-1} \circ s \circ g = [\tau_\alpha^{-1} \circ s \circ \tau_\beta^{-1}] \circ [\tau_\beta \circ g] = F \circ h$$

où l'on a posé $F = \tau_\alpha^{-1} \circ s \circ \tau_\beta^{-1}$. Vérifions que $F(0) = 0$. D'abord, $\tau_\beta^{-1}(0) = \beta$ puisque $\tau_\beta(\beta) = 0$. Ensuite, $s(\beta) = \beta^2 = g(z_0)^2 = \tau_\alpha \circ f(z_0) = \tau_\alpha(0)$ par diverses définition et puisque $f(z_0) = 0$. Finalement, $F(0) = \tau_\alpha^{-1} \circ \tau_\alpha(0) = 0$, comme annoncé. C'est bien, parce que le lemme de Schwarz dit que $|F'(0)| < 1$ (puisqu'en effet F envoie $\mathbb{D}$ dans lui-même et n'est même pas injective (à cause de s), donc n'est pas linéaire). Finalement,

$$|f'(z_0)| = |F'(h(z_0))|\,|h'(z_0)| = |F'(0)|\,|h'(z_0)| < |h'(z_0)|$$

$h(z_0) = \tau_\beta \circ g(z_0) = 0$ par définition de β, puis parce que $f'(z_0) = M > 0$ par définition de f; on en déduit la contradiction recherchée. $\square$

24. LA CLASSE S ET LE THEOREME DE DISTORTION DE KOEBE

On va d'abord s'intéresser aux propriétés de régularité intérieure des applications $f : \mathbb{D} \to \mathbb{C}$ conformes. Le livre de Peter L. Duren, Univalent functions (Grundlehren 259, chez Springer) semble bien écrit, et contient bien bien plus d'informations qu'on peut imaginer. D'ailleurs une bonne partie de ce qui est ci-dessous en vient. On peut aussi utiliser le livre de Pommerenke. [Voir les références en page 1.]

a. Les classes S, Σ, et Σ'

On notera

$$S = \big\{ f \in H(D)\,;\, f \text{ est injective, } f(0) = 0, \text{ et } f'(0) = 1 \big\} \tag{1}$$

la classe des applications $f : \mathbb{D} \to \mathbb{C}$ conformes, normalisées. Noter qu'on se ramène aisément au cas où $f(0) = 0$ et $f'(0) = 1$ par translation et multiplication par une constante. C'est utile de normaliser, car comme on le verra, S est ainsi une famille normale, c.-à.-d. (séquentiellement) compact pour la topologie de la convergence uniforme sur tout compact. Ici S signifie Schlicht, ou encore Univalent, mais Schlicht est beaucoup plus smart.

Voici quelques exemples simples. D'abord,

$$f(z) = \frac{1+z}{1-z} \tag{2}$$

n'est pas dans S, mais envoie $\mathbb{D}$ sur le demi-plan de droite $\big\{w\,;\, Re(w) > 0\big\}$, avec la normalisation $f(0) = 1$. Vérification laissée au lecteur. Ceci seraait utile si on voulait passer des représentations conformes basées sur le disque aux représentations conformes basées sur le demi plan. Les choses y sont parfois plus simples. La fonction de Koebe est

$$k(z) = z(1-z)^{-2} = z + 2z^2 + 3z^3 + 4z^4 + \dots \tag{3}$$

qui envoie le disque sur le complémentaire de $\big\{x \in \mathbb{R}\,;\, x \leq -1/4\big\}$. Pour le voir, on note que si f est comme plus haut, $f(z)^2 - 1 = [(1+z)^2 - (1-z)^2]\,(1-z)^{-2} = 4z(1-z)^{-2}$ donc $k(z) = \frac{1}{4}\,f(z)^2 - \frac{1}{4}$. Ensuite on voit que k est la composée de f, puis de $w \to w^2$ qui envoie le demi-plan sur le complémentaire de l'axe réel négatif, et on retire encore $-1/4$ pour que $k(0) = 0$. Fonction très utile pour toutes sortes de problèmes d'extremas.

De temps en temps, on préfère étudier la version de S près de l'infini, à savoir

(4)

$$\Sigma = \big\{g \text{ analytique et injective sur } A = \{z \in \mathbb{C}\,;\, |z| > 1\}\,;\, g(z) = z + b_0 + b_1 z^{-1} + \dots\big\}$$

ou plutôt son petit frère

$$\Sigma' = \big\{g \in \Sigma\,;\, 0 \notin g(A)\big\}. \tag{5}$$

On aura besoin de savoir que si $f \in S$, alors $g \in \Sigma'$, où g est définie par inversion:

$$g(z) = f(1/z)^{-1} \text{ pour } z \in A. \tag{6}$$

Vérification immédiate que la normalisation est correcte, après développement de $f(w)$ en série près de l'origine. Mais réciproquement, si $g \in \Sigma'$, $f(z) = g(1/z)^{-1}$ est analytique sur $\mathbb{D}\backslash\{0\}$, puis en 0 (utiliser de développement limité donné plus haut pour vérifier que la singularité en 0 est effaçable), et enfin $f \in S$ (l'injectivité va de soi, et $f(0) = 0$ et $f'(0) = 1$ correspondent à nos contraintes sur le développement limité. Noter aussi que si g prenait la valeur 0, f ne serait pas bien définie.

Encore une remarque utile, concernant un moyen malin de construire de nouvelles fonctions univalentes. Commençons par le cas de S. Soit donc $f \in S$; on va définir

$$F(z) = f(z^2)^{1/2} \tag{7}$$

et vérifier que $F \in S$ aussi. On écrit $f(z) = zg(z)$, avec donc g analytique non nulle sur $\mathbb{D}$, et on note G une racine de $g(z^2)$. C'est possible, puisque $\mathbb{D}$ est simplement connexe. Noter que $g(0) = 1$; quitte à multipler G par -1, on peut supposer que $G(0) = 1$ aussi, et maintenant on pose $F(z) = zG(z)$; bien sûr $F(z)^2 = z^2 G(z)^2 = z^2 g(z^2) = f(z^2)$, comme souhaité. Si $F(z_1) = F(z_2)$, les carrés sont égaux aussi, donc (par (7) et puisque f est injective), donc $z_1^2 = z_2^2$ et $z_1 = \pm z_2$. En plus, $g(z^2)$ est paire, donc G aussi est paire (près de 0, $G(z)$ est obtenu à partir de $g(z^2)$ en prenant la seule racine proche de 1; ensuite on passe à $\mathbb{D}$ tout entier par prolongement analytique. Bref G est paire, donc F est impaire, de sorte que $z_1 \neq z_2$ donne $z_1 = -z_2$, donc $F(z_1) = 0$ et finalement $z_1 = z_2 = 0$ quand même. Donc F est injective. Et $F(0) = 0$ et $F'(0) = 1$ par construction, donc $F \in S$.

On peut faire la même chose pour $g \in \Sigma'$. Cette fois on veut définir

$$G(z) = g(z^2)^{1/2}, \tag{8}$$

et le plus simple est de se ramener au cas précédent par une inversion, en prenant $F(w) = g(w^{-2})^{-1/2}$ pour $w \in \mathbb{D}$ puis $G(z) = F(z^{-1})^{-1}$. On pourrait aussi procéder directement, en écrivant $G(z) = zh(z)$, puis en extrayant une racine H de $h(z^2)$ sur le domaine A (qui n'est pas simplement connexe, mais ici $h(\infty) = 1$, ce qui permet de tracvailler sur $A \cup \{\infty\}$, ou justement de manière équivalente sur $\mathbb{D}$ par inversion), et en prenant $G(z) = zH(z)$ sur A. Noter qu'il est important que g ne prenne pas la valeur 0, pour pouvoir extraire une racine carrée.

b. Théorème de l'aire

Evdemment le résultat ci-dessus utilise de manière importante l'injectivité de g.

Théorème (de l'aire). *Pour $g(z) = z + b_0 + \sum_{n=1}^{+\infty} b_n z^{-n}$ dans la classe Σ, on a*

$$\sum_{n=1}^{\infty} n|b_n|^2 \leq 1 \tag{9}$$

avec égalité si et seulement si $\mathbb{C} \setminus g(A)$ est de mesure de Lebesgue nulle.

Posons $C_r = g(\partial B(0,r))$ pour $r > 1$; comme g est injective (et donc sa dérivée ne s'annule pas), C_r est un lacet régulier, qui borne un domaine borné Ω_r (le complémentaire de $g(\mathbb{C} \setminus B(0,r))$). Et l'aire de Ω_r est, par Green appliqué à la fonction z, dont la divergence est 2,

$$A(r) = \int\int_{\Omega_r} d\lambda(z) = \frac{1}{2}\int_{C_r} z \cdot n \tag{10}$$

où n est la normale extérieure à Ω_r. On note que $z \cdot n = Re(\overline{z}n)$ en notations complexes, et on trouve

$$A(r) = \frac{1}{2}\int_{C_r} Re(\overline{z}n) = Re\Big(\frac{1}{2i}\int_{C_r} \overline{z}dz\Big). \tag{11}$$

Ensuite on pose

$$z = g(re^{i\theta}) = re^{i\theta} + b_0 + \sum_{n=1}^{+\infty} b_n r^{-n} e^{-in\theta}, \tag{12}$$

et, puisque $g'(w) = 1 - \sum_{n=1}^{+\infty} nb_n w^{-n-1}$

$$dz = ire^{i\theta} g'(re^{i\theta})d\theta = ire^{i\theta}d\theta - i\sum_{n=1}^{+\infty} nb_n r^{-n} e^{-in\theta} d\theta. \tag{13}$$

Dans les deux cas, la série converge normalement (convergence d'une série de Laurent pour des rayons strictement plus grands que le rayon de convergence, ou, de manière équivalente après inversion, convergence de la série entière strictement au-dessous du rayon de convergence). L'intégrale dans (11) est donc une somme double, dont seuls les termes de fréquence nulle restent, de sorte que

$$A(r) = Re\Big(\frac{1}{2i}\int_{C_r} \overline{z}dz\Big) = \pi r^2 - \pi\sum_{n=1}^{+\infty} n|b_n|^2 r^{-2n} \tag{14}$$

(et la partie imaginaire était nulle de toute façon). Evidemment, $A(r) \geq 0$, donc on sait déjà que $\sum_{n=1}^{+\infty} n|b_n|^2 r^{-2n} \leq \pi r^2$ pour $r > 1$. Logiquement,

$A(r)$ est une fonction décroissante de r (et cela se voit sur la formule). Sa limite est $\pi - \pi \sum_{n=1}^{+\infty} n|b_n|^2$, mais aussi $|\mathbb{C} \setminus g(A)|$ (mesure d'une intersection dénombrable décroissante d'ensembles Ω_r de mesure finie), de sorte que

$$(15) \qquad 1 - \sum_{n=1}^{+\infty} n|b_n|^2 = \frac{1}{\pi} \lim_{r \to 1} A(r) = \frac{1}{\pi} |\mathbb{C} \setminus g(A)|,$$

ce dont on déduit le théorème. □

Noter que ceci implique aussi que $|\mathbb{C} \setminus g(A)| \leq \pi$, donc n'est pas trop gros, ce qu'on retrouvera dans le théorème de Koebe.

Exercice. Montrer que S est compacte (séquentiellement si vous préférez), en utilisant le théorème de l'aire.

c. Théorème du quart de Keobe

Corollaire (Bieberbach). *Pour $f(z) = z + \sum_{k \geq 2} a_k z^k$ dans S, on a $|a_2| \leq 2$, avec égalité seulement quand f est la conjugaison de la fonction de Koebe par une rotation.*

On part de $f \in S$, construit $F(z) = f(z^2)^{1/2}$ come en (7), puis on applique une inversion pour obtenir $G(w) = F(w^{-1})^{-1} = f(w^{-2})^{-1/2}$ dans la classe Σ'. Mais si $f(z) = z + a_2 z^2 + o(z^2)$, $f(w^{-2}) = w^{-2} + a_2 w^{-4} + o(w^{-4}) = w^{-2}(1 + a_2 w^{-2} + o(w^{-2}))$ et $G(w) = w^{-1}(1 - \frac{a_2}{2} w^{-2} + o(w^{-2}))$, donc $b_0 = 0$ et $b_1 = a_2/2$ avec les notations plus haut, de sorte que $|a_2| = 2|b_1| \leq 2$, par le théorème de l'aire.

S'il y a égalité, on trouve que $b_n = 0$ pour $n \geq 2$, donc $G(w) = w - e^{i\theta} w^{-1} = w(1 - e^{i\theta} w^{-2})$, donc

$$f(w^{-2}) = G(w)^{-2} = w^{-2}(1 - e^{i\theta} w^{-2})^{-2} = e^{-i\theta}[e^{i\theta} w^{-2}(1 - e^{i\theta} w^{-2})^{-2}] = e^{-i\theta} k(e^{i\theta} w^{-2})$$

Et comme $w \to w^{-2}$ est une surjection de A sur $\mathbb{D}$, il vient $f(z) = e^{-i\theta} k(e^{i\theta} z)$ sur $\mathbb{D}$, comme promis. □

Ceci a incité Bieberbach à conjecturer que k est aussi extrémale pour le choix du n-ième coefficient, c.-à.-d. que $|a_n| \leq n$ pour tout n, et ceci a finalement été démontré par De Branges.

On s'intéresse à une conséquence fort utile du théorème de l'aire.

Théorème (Koebe One-Quarter Theorem). *Si $f \in S$, alore $f(\mathbb{D})$ contient $B(0,1/4)$.*

A nouveau, k est extrémale, puisque c'est une représentation conforme de $\mathbb{D}$ sur le complémentaire de $]-\infty,-1/4]$. Evidemment le théorème est faux pour les fonctions qui sont analytiques non injectives.

Démonstration. Soit $f \in S$, et supposons que $\omega \in \mathbb{C}$ n'est pas dans son image. On veut vérifier que $|\omega| \geq 1/4$. Posons

$$g(z) = \frac{\omega f(z)}{\omega - f(z)} \quad \text{pour } z \in \mathbb{D} \tag{16}$$

C'est bien défini (le dénominateur ne s'annule pas), g est injective parce que f l'est, et $g(0) = 0$ et $g'(0) = 1$ par calcul. Donc $g \in S$, et on va lui appliquer le théorème de Bieberbach. Or, près de 0,

$$\begin{aligned} g(z) &= \frac{\omega(z + a_2 z^2 + o(z^2))}{\omega - z + o(z)} = z\,\frac{1 + a_2 z + o(z)}{1 - \frac{z}{\omega} + o(z)} \\ &= z\,[1 + a_2 z + o(z)]\,[1 + \frac{z}{\omega} + o(z)] = z + (a_2 + \frac{1}{\omega})z^2 + o(z^2) \end{aligned} \tag{17}$$

Et on en déduit que $|a_2 + \frac{1}{\omega}| \leq 2$, donc $|\frac{1}{\omega}| \leq 2 + |a_2| \leq 4$ et finalement $|\omega| \geq 1/4$, comme annoncé. □

d. Théorèmes de distortion

Le théorème de Bienerbach peut être vu comme une version améliorée du lemme de Schwarz, dont on va voir des versions délocalisées et des conséquences pour les représentations conformes. On note encore τ_a la transformée de Möbius définie par

$$\tau_a(z) = \frac{z - a}{1 - \overline{a}z}, \tag{18}$$

et on va utiliser le fait que pour $f \in S$ et $a \in D$,

$$f_a(z) = \frac{f \circ \tau_a^{-1}(z) - f(a)}{(1 - |a|^2) f'(a)} \tag{19}$$

est encore dans S. Vérification facile: $f_a(z) = 0$ car $\tau_a(a) = 0$, et $f_a'(0) = 1$ parce que $(f \circ \tau_a^{-1})'(0) = f'(a)/\tau_a'(0) = (1 - |a|^2) f'(a)$ (voir ci-dessous). Pour la seconde dérivée, notons $\tau = \tau_{-a}$, et calculons

$$(f \circ \tau_a^{-1})''(z) = (\tau' f' \circ \tau)'(z) = \tau''(z) f' \circ \tau(z) + \tau'(z)^2 f'' \circ \tau(z) \tag{20}$$

et comme

$$\tau_a'(z) = [1 - \overline{a}z + \overline{a}(z-a)](1-\overline{a}z)^{-2} = [1-|a|^2](1-\overline{a}z)^{-2} \tag{21}$$

puis

$$\tau_a''(z) = 2\overline{a}[1-|a|^2](1-\overline{a}z)^{-3} \tag{22}$$

on trouve que $\tau''(z) = -2\overline{a}[1-|a|^2](1+\overline{a}z)^{-3}$, qui vaut $-2\overline{a}$ en 0. Finalement, on calcule $f_n''(0)$ à partir de (19), puis on utilise (20); on trouve que

$$\begin{aligned}(1-|a|^2)f'(a)f_a''(0) &= (f\circ\tau_a^{-1})''(0) = -2\overline{a}(f'\circ\tau)(0) + (1-|a|^2)^2 f''\circ\tau(0)\\ &= -2\overline{a}(1-|a|^2)f'(a) + (1-|a|^2)^2 f''(a)\end{aligned} \tag{23}$$

et donc

$$f_a''(0) = -2\overline{a} + \frac{(1-|a|^2)f''(a)}{f'(a)}.$$

Or le coefficient a_2 correspondant, qui est la moitié de $f_a''(0)$ est de module ≤ 2 par Bieberbach. Donc,

$$\left|\frac{(1-|a|^2)f''(a)}{f'(a)} - 2\overline{a}\right| \leq 4. \tag{24}$$

On multiplie par a, et on divise par $1-|a|^2$ pour obtenir ceci:

Théorème 1. *Pour $f \in S$ et $a \in \mathbb{D}$,*

$$\left|\frac{af''(a)}{f'(a)} - 2\frac{|a|^2}{1-|a|^2}\right| \leq \frac{4|a|}{1-|a|^2}. \tag{25}$$

On aime bien f''/f', parce que c'est la dérivée de $\log(f')$, donc quand c'est petit f' est assez stable. Passons à une estimation de f', en gros obtenue à partir du théorème 1 par intégration.

Théorème 2. *Pour $f \in S$ et $z \in \mathbb{D}$, et en notant $r = |z|$,*

$$\frac{1-r}{(1+r)^3} \leq |f'(z)| \leq \frac{1+r}{(1-r)^3}. \tag{26}$$

Le théorème 1 donne, rien qu'en prenant la partie réelle,

$$\frac{2r^2-4r}{1-r^2} \leq Re\Big(\frac{zf''(z)}{f'(z)}\Big) \leq \frac{2r^2+4r}{1-r^2}. \tag{27}$$

Noter que f' ne s'anule pas, ce qui permet d'en prendre un logarithme dans le domaine simplement connexe $\mathbb{D}$, et si $z = re^{i\theta}$,

$$r\frac{\partial}{\partial r}(\log(f'(z)) = re^{i\theta}\frac{f''(z)}{f'(z)} = \frac{zf''(z)}{f'(z)}.$$

En prenant la partie réelle, on voit que $r\frac{\partial}{\partial r}(|\log f'(z)|)$ est le terme central de (27). Donc

$$\frac{2r-4}{1-r^2} \leq \frac{\partial}{\partial r}(|\log f'(z)|) \leq \frac{2r+4}{1-r^2}. \tag{28}$$

Comme $|\log f'(0)| = 0$, on intègre et on trouve que

$$\log\Big(\frac{1-r}{(1+r)^3}\Big) \leq \log|f'(z)| \leq \log\Big(\frac{1+r}{(1-r)^3}\Big)$$

puisque $2r-4 = 3(r-1)-(r+1)$ et pareillement $2r+4 = 3(r+1)-(r-1)$. $\square$

Et on va intégrer une fois de plus pour estimer f.

Théorème 3. *Pour $f \in S$ et $z \in \mathbb{D}$, et en notant toujours $r = |z|$,*

$$\frac{r}{(1+r)^2} \leq |f(z)| \leq \frac{r}{(1-r)^2}. \tag{29}$$

La borne supérieure est facile: on note que si $z = re^{i\theta}$, $f(z) = \int_0^r e^{i\theta}f'(\rho e^{i\theta})d\rho$, on intègre la seconde inégalité de (26), et on trouve $|f(z)| \leq \int_0^r \frac{1+\rho}{(1-\rho)^3}\,d\rho \leq \frac{r}{(1-r)^2}$.

La surprise est qu'il faut se donner du mal pour la borne inférieure, qui correspond à une borne supérieure pour l'inverse.

Si $|f(z)| \geq 1/4$, on note simplement que $r \to \frac{r}{(1+r)^2}$ est une fonction croissante de r, qui vaut $1/4$ en $r=1$, donc l'inégalité est vraie. Donc on peut supposer que $|f(z)| < 1/4$, et alors le théorème de Koebe dit que f^{-1}

est définie sur le segment $[0, f(z)]$. On note γ l'image de ce segment par f^{-1}, et on écrit que

$$f(z) = \int_\gamma f'(\zeta)d\zeta$$

Mais il se trouve que $f'(\zeta)d\zeta$ pointe toujours dans la même direction (celle de $f(z)$). En fait (pour faire les calculs brutalement), si $f(z) = \rho e^{i\theta}$, on paramètre γ par $\gamma(t) = f^{-1}(t\rho e^{i\theta})$, alors

$$\gamma'(t) = \rho e^{i\theta}(f^{-1})'(t\rho e^{i\theta}) = \frac{\rho e^{i\theta}}{f'(\gamma(t))} \tag{30}$$

et du coup $d\zeta = d(\gamma(t))\gamma'(t)dt = \frac{\rho e^{i\theta}}{f'(\gamma(t))}dt$ donc $f'(\zeta)d\zeta = \rho e^{i\theta}dt$. Ainsi

$$\begin{aligned} |f(z)| &= \int_\gamma |f'(\zeta)|dl(\zeta) \geq \int_\gamma \frac{1-|\zeta|}{(1+|\zeta|)^3}dl(\zeta) \\ &\geq \int_\gamma \frac{1-|\zeta|}{(1+|\zeta|)^3}\,\frac{d}{dl}(|\zeta|)\,dl(\zeta) \geq \int_0^r \frac{1-t}{(1+t)^3}dt \end{aligned} \tag{31}$$

où l'on a noté dl la longueur sur γ. Au cas où ce calcul vous déplairait, essayons de le justifier en revenant au paramétrage par γ. On écrit donc $f(z) = \rho e^{i\theta}$ et on pose $\gamma(t) = f^{-1}(t\rho e^{i\theta})$ pour $t \in [0,1]$; un peu plus loin dans le calcul on posera $r(t) = |\gamma(t)|$. Noter que par (30) $\gamma'(t)f'(t) = \rho e^{i\theta}$, donc

$$\begin{aligned} |f(z)| = \rho &= \int_0^1 |\gamma'(t)f'(t)|dt \geq \int_0^1 \frac{1-|\gamma(t)|}{(1+|\gamma(t)|)^3}|\gamma'(t)|dt \\ &= \int_0^1 \frac{1-r(t)}{(1+r(t))^3}|\gamma'(t)|dt \geq \int_0^1 \frac{1-r(t)}{(1+r(t))^3}r'(t)dt \end{aligned}$$

où l'on peut même dire que $r(t)$ est C^1 parce que $\gamma(t) \neq 0$ sur $]0,1]$, vu que $f(0) = 0$ et f est injective. Mais bien sûr cà n'est pas cela qui nous aurait arrêté autrement. En tout cas, on pose $r(t) = \rho$ dans l'intégrale, et on se souvient que pour les intégrales de Riemann, il n'est pas nécessaire que $r(t)$ soit injective pour faire le changelent de variable, puis on note qu'une primitive de $\frac{1-r}{(1+r)^3}$ est $\frac{r}{(1+r)^2}$. On trouve que $|f(z)| \geq \frac{r(1)}{(1+r(1))^2} = \frac{|z|}{(1+|z|)^2}$, comme promis. $\square$

Il se trouve que quand, dans les théorèmes ci-dessus, il y a égalité pour un $z \neq 0$, la fonction f est une conjuguée de fonction de Koebe par une rotation. On laisse la vérification au lecteur.

Il y a une dernière estimée mentionnée dans le livre de Duren, qui porte sur $|zf'(z)/f(z)|$, mais on va essayer de faire sans.

Par contre, disons deux mots de comment on peut se servir de ces estimations. Soit $f : U \to V$ conforme bijective, entre deux domaines du plan différents de $\mathbb{C}$, et soit $x \in U$. Vérifions que

$$\frac{1}{4}\,|f'(x)|\,\mathrm{dist}(x, \partial U) \leq \mathrm{dist}(f(x), \partial V) \leq 4|f'(x)|\,\mathrm{dist}(x, \partial U). \tag{32}$$

On pose $D = \mathrm{dist}(x, \partial U)$ et on applique le théorème de Koebe à la fonction $g(z) = \dfrac{f(x + Dz)}{Df'(x)} - f(x)$, qui est bien dans S parceque $g'(z) = 1$ par normalisation. On trouve que $g(\mathbb{D})$ contient $B(0, 1/4)$, ce qui signifie que $f(B(x, D))$ contient $B(f(x), D|f'(x)|/4)$. En particulier, $\mathrm{dist}(f(x), \partial V) \geq D|f'(x)|/4$, la première inégalité de (32). Mais le même résultat appliqué à f^{-1} donne $\mathrm{dist}(x, \partial U) \geq \mathrm{dist}(f(x), \partial V)|f'(x)|^{-1}/4$, ce qui est la seconde inégalité.

Supposons que l'on définisse la distance hyperbolique sur U par

$$d_U(x, y) = \inf\Big\{ \int_\gamma \mathrm{dist}(\zeta, \partial U)^{-1}\,;\, \gamma \text{ est un chemin } C^1 \text{ dans } U \text{ qui va de } x \text{ à } y \Big\}$$

(ce qui n'est sans doute qu'équivalent à la formule usuelle), alors les représentations conformes préservent la distance hyperbolique à une constante 4 près. En effet, si $\gamma : [0, 1]$ est un chemin C^1 dans U qui mène de x à y, et si $f : U \to V$ est conforme (et bijective), alors $f \circ \gamma$ est un chemin C^1 dans V qui mène de $f(x)$ à $f(y)$, et par exemple

$$\begin{aligned} d_V(f(x), f(y)) &\leq \int_{f\circ\gamma} \mathrm{dist}(\zeta, \partial V)^{-1} = \int_{[0,1]} \mathrm{dist}(f \circ \gamma(t), \partial V)^{-1} |(f \circ \gamma)'(t)| dt \\ &= \int_{[0,1]} \mathrm{dist}(f \circ \gamma(t), \partial V)^{-1} |f'(\gamma(t))||\gamma'(t)| dt \\ &\leq 4 \int_{[0,1]} \mathrm{dist}(\gamma(t), \partial U)^{-1} |\gamma'(t)| dt, \end{aligned}$$

donc en prenant la borne inférieure sur γ, $d_V(f(x), f(y)) \leq 4d_U(x, y)$. On obtiendrait que $d_V(f(x), f(y)) \geq d_U(x, y)/4$ en considérant f^{-1}.

Ce genre de considération est assez utile dès qu'on manipule des représentations conformes.

25. EXTENSION AU BORD DE LA REPRESENTATION CONFORME

On va démontrer un gros bout du résultat suivant, qui je crois est un théorème de Carathéodory (voir le théorème 2).

Théorème 1. *Soit $\Omega \subset \mathbb{C}$ un domaine borné simplement connexe, dont la frontière est une courbe simple γ, et soit $f : \Omega \to \mathbb{D}$ une représentation conforme. Alors f s'étend par continuité en un homéomorphisme de $\overline{\Omega}$ dans $\overline{\mathbb{D}}$, dont la restriction à $\partial\Omega$ est un homéomorphisme de $\partial\Omega$ dans le cercle unité.*

On verra des choses un peu plus précises, mais quand même on évitera de parler de certaines pathologies à la frontière. On va suivre la démonstration donnée dans Rudin. On commence par un lemme sur H^∞.

Lemme 1. *Soit $f : \mathbb{D} \to \mathbb{C}$ une fonction analytique bornée. Alors la limite $\widetilde{f}(\theta) = \lim_{r\to 1} f(re^{i\theta})$ existe pour presque-tout $\theta \in \mathbb{R}$, et de plus,*

$$f(re^{i\varphi}) = \frac{1}{2\pi} \int_{[0,2\pi]} P_r(\theta - \varphi)\widetilde{f}(\theta)d\theta \tag{1}$$

pour $0 \le r < 1$ et $\varphi \in \mathbb{R}$, où $P_r(\theta)$, le noyau de Poisson, est donné par

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2}. \tag{2}$$

Pour l'instant, on laisse la démonstration du lemme 1 en suspens, que vous connaissez sans doute. Noter que le lemme est vrai non seulement pour les fonctions analytiques bornées comme ci-dessus, mais pour les fonctions de H^1 (c.-à.-d., analytiques et telles que $||f||_{H_1} = \frac{1}{2\pi} \sup_{0<r<1} \int_{[0,2\pi]} |f(re^{i\theta})|d\theta$ est fini). Mais là c'est un peu plus délicat. Voir le prochain chapitre.

Et maintenant on va regarder comment se comporte la représentation conforme $f : \Omega \to \mathbb{D}$ près des "points frontière simples".

Definition. *Un point a de $\partial\Omega$ est dit simple si pour toute suite $\{z_n\}$ de points de Ω qui tend vers a, on peut trouver un chemin continu $\gamma : [0, 1[\to \Omega$, et des temps $0 < t_1 < \cdots t_n$, avec $\lim_{n\to+\infty} t_n = 1$, $\gamma(t_n) = z_n$ pour tout n, et $\lim_{t\to 1} \gamma(t) = a$.*

Bref, on peut trouver une courbe (à savoir $\gamma([0, 1[)$ dans Ω, qui passe par tous les z_n (dans l'ordre) et se termine en a. Ce n'est pas le cas, par exemple, si $\partial\Omega$ contient un segment, avec accès des deux côtés, et a est au milieu du segment. Plus grave, on peut construire des morceaux de frontière, avec des peignes, qui ne sont pas accessibles du tout. Et dans ce qui suit, on ne s'occupera pas de ceux-ci.

Dans ce qui suit, Ω est simplement connexe borné et $f : \Omega \to \mathbb{D}$ est une représentation conforme. Noter que supposer Ω borné est souvent moins grave qu'il n'y paraît, parce que si $\mathbb{C} \setminus \Omega$ contient un disque $B(a, r)$, on peut considérer l'image de Ω par $z \to \frac{1}{z-a}$, qui est borné. On peut aussi essayer d'écarter par une application racine.

Lemme 2. *Si a est un point frontière simple de Ω, $f(a) = \lim_{z\in\Omega\,;\,z\to a} f(z)$ existe et est dans ∂D.*

On notera $g : \mathbb{D} \to \Omega$ la réciproque de f. C'est une application analytique bornée injective.

Supposons que la limite n'existe pas. Donc il existe une suite $\{z_n\}$ dans Ω, qui tend vers a, mais telle que $\{f(z_n)\}$ n'ait pas de limite. Comme cette dernière est une suite bornée, elle a au moins deux points d'adhérence, x_0 et x_1, donc en fait, quitte à passer à une suite extraite, on peut supposer que $\{f(z_{2n})\}$ converge vers x_0 et $\{f(z_{2n+1})\}$ converge vers x_1.

La définition d'un point simple donne un chemin γ associé à la suite $\{z_n\}$, dont l'image $\Gamma = f \circ \gamma$ oscille entre x_0 et x_1, ce qui donnera bientôt une contradiction.

Posons $K_r = g(\overline{B}(0, r))$; c'est un compact contenu dans Ω, donc, puisque $\lim_{t\to 1} \gamma(t) = a$, il existe $t(r)$ tel que $\gamma(t) \notin K_r$ pour $t \geq t(r)$. Autrement dit,

$$\lim_{t\to 1} |\Gamma(t)| = 1. \tag{3}$$

Noter que ceci vaut aussi sans supposer que $\lim_{z\in\Omega\,;\,z\to a} f(z)$ n'existe pas, et donnera le fait que $|f(a)| = 1$. Noter aussi que $|x_0| = |x_1| = 1$, par (3).

Soient maintenant J_1 et J_2 les composantes connexes de $\partial B(0, 1)\setminus\{x_0, x_1\}$, en choisissons $\theta_1 \in J_1$ et $\theta_2 \in J_2$. Supposons d'abord que pour tout choix de θ_1, le segment $[0, r\theta_1]$ rencontre $\Gamma([t_{2n}, t_{2n+1}])$ pour une infinité de valeurs de n.

Alors les points d'intersections (notons-les y_n) tendent vers le cercle (à cause de (3)). Pour presque tout $\theta_1 \in J_1$, $g(y_n)$ (pour cette sous-suite) tend vers la limite radiale de g au point θ (qui existe grâce au lemme 1). Mais par ailleurs, $y_n = f \circ \gamma(s_n)$ pour un $s_n \in [t_{2n}, t_{2n+1}]$, donc $g(y_n) = \gamma(s_n)$ tend vers a par définition de γ. On a donc montré que la limite radiale de $g - a$ est 0 presque-partout sur J_1. Posons

$$G(z) = \prod_{k=1}^{m} [g(ze^{2ik\pi/m}) - a] \tag{4}$$

où m est un entier qu'on choisit assez grand pour que l'union des $e^{2ik\pi/m} J_1$, $1 \leq k \leq m$, contienne $\partial\mathbb{D}$. C'est encore une fonction analytique bornée sur $\mathbb{D}$, et par choix de m, ses valeurs au bord sont nulles presque partout. Par le lemme 1, comme $G(z)$ se calcule à partir de ses valeurs au bord, $G(z) = 0$ sur $\mathbb{D}$. Donc l'une au moins des $g(ze^{2ik\pi/m}) - a$ a une infinité non dénombrable de zéros, donc est nulle. Alors, $g = a$, ce qui contredit son injectivité.

Donc il existe au moins un $\theta_1 \in J_1$ tel que le segment $[0, r\theta_1]$ ne rencontre $\Gamma([t_{2n}, t_{2n+1}])$ que pour un nombre fini de valeurs de n. Par le même argument, on peut trouver $\theta_2 \in J_2$ tel que $[0, r\theta_2]$ ne rencontre $\Gamma([t_{2n}, t_{2n+1}])$ que pour un nombre fini de valeurs de n. Cependant, vérifions que pour n assez grand, l'arc $\Gamma([t_{2n}, t_{2n+1}])$ doit rencontrer $[0, r\theta_1] \cup [0, r\theta_1]$; on en déduira une contradiction. C'est facile: on peut construire (à la main!) une fonction continue sur $\overline{\mathbb{D}}$ qui est > 0 du côté de x_0 de $[0, r\theta_1] \cup [0, r\theta_1]$, et < 0 du côté de x_1. Alors $\Gamma(t)$ change de signe entre $f(z_{2n}) = f \circ \gamma(t_{2n}) = \Gamma(t_{2n})$ (qui tend vers x_0) et $f(z_{2n+1}) = \Gamma(t_{2n+1})$ (qui tend vers x_1). Cette contradiction finale montre l'existence de $f(a)$. Le fait que $|f(a)| = 1$ vient de ce que, si $\{z_n\}$ est une suite dans Ω qui tend vers a, $f(z_n)$ tend vers $f(a)$, et en même temps, $|f(z_n)|$ tend vers 1 par la preuve de (3). $\square$

Lemme 3. *Si a_1, $a_2 \in \partial\Omega$ sont des points simples distincts, alors $f(a_1) \neq f(a_2)$ (où bien sûr les $f(a_j)$ sont définis au lemme 2).*

Choisissons, pour $j = 1, 2$, une suite $\{z_{j,n}\}$ dans Ω qui tend vers a_j, et un chemin γ_j comme dans la définition d'un point frontière simple, et posons $\Gamma_j = f \circ \gamma_j$. Ainsi, $\lim_{t\to 1} \gamma_j(t) = a_j$ et en particulier il existe $t_0 < 1$ tel que

$$|\gamma_1(t_1) - \gamma_2(t_2)| \geq \frac{1}{2} |a_1 - a_2| \tag{5}$$

pour $t_1, t_2 \geq t_0$.

Supposons que $f(a_1) = f(a_2) = \xi$. Notons que $\Gamma_1([0, t_0]) \cup \Gamma_2([0, t_0])$ est compact dans Ω, donc ne rencontre pas $B(\xi, \rho)$ pour ρ assez petit. Cependant, $\lim_{t\to 1} \Gamma_j(t) = f(a_j) = \xi$ pour $j = 1, 2$ (puisque $\lim_{t\to 1} \gamma_j(t) = a_j$), donc Γ_j

rencontre $\partial B(\xi,\rho)$ en au moins un point w_j (par le lemme de passage des frontières), et du coup $w_j = \Gamma_j(t_j)$ pour un $t_j \geq t_0$. En particulier,

$$|g(w_1) - g(w_2)| = |\gamma_1(t_1) - \gamma_2(t_2)| \geq \frac{1}{2}\,|a_1 - a_2|$$

par (5). Mais $|g(w_1)-g(w_2)| \leq \int_{\partial(\rho)} |g'(z)|dl(z)$, en notant $\partial(\rho) = \partial B(0,1) \cap \partial B(\xi,\rho)$ l'arc de $\partial B(\xi,\rho)$ qui est contenu dans le disque. Par Cauchy-Schwarz et puisque la longueur de $\partial(\rho)$ est inférieure à $\pi\rho$,

$$|g(w_1) - g(w_2)|^2 \leq \pi\rho \int_{\partial(\rho)} |g'(z)|^2 dl(z).$$

En divisant par ρ puis en intégrant pour $0 < \rho < \rho_0$ (assez petit), il vient (6)

$$+\infty = \frac{1}{4}\,|a_1 - a_2|^2 \int_0^{\rho_0} \frac{d\rho}{\rho} \leq \pi \int_0^{\rho_0} \int_{\partial(\rho)} |g'(z)|^2 dl(z) = \pi \int_{\Omega \cap B(\xi,\rho_0)} |g'(z)|^2 d\lambda(z)$$

Or il se trouve que l'intégrale de droite converge, puisque pour tout $r < 1$

$$\int_{B(0,r)} |g'(z)|^2 d\lambda(z) = \int_{B(0,r)} |\det J_g(z)| d\lambda(z) = |g(B(0,r)| \leq |\Omega| < +\infty$$

où $J_g(z)$ est le déterminant jacobien en z de l'application g, et parce que Ω est borné. Ceci termine la démonsration du lemme 3. $\square$

Il reste à prouver un morceau du théorème. On ne vérifiera pas que $\partial\Omega$ est une courbe de Jordan, alors tout point de $\partial\Omega$ est simple, et on va continuer à partir de là. Donc on suppose que tous les points de $\partial\Omega$ sont simples, et le lemme 2 donne une extension continue $\overline{\Omega}$ dans $\overline{D}$.

Cette extension est continue: si $\{z_n\}$ dans $\overline{\Omega}$ tend vers z, on peur remplacer chaque $z_n \in \partial\Omega$ par un $z'_n \in \Omega$ tel que $|z_n - z'_n| + |f(z_n) - f(z'_n)| \leq 2^{-n}$, et obtenir que $f(z) = \lim_{n\to+\infty} f(z'_n) = \lim_{n\to+\infty} f(z_n)$.

Elle est injective, par le lemme 3 et le fait que $|f(z)| = 1$ pour $z \in \partial\Omega$. Elle est surjective, car l'image est compacte et contient $\mathbb{D}$. Enfin la réciproque est continue, comme toujours pour les réciproques de bijections continues sur des compacts. Et cette réciproque est une extension de g, bien sûr, donc est conforme sur $\mathbb{D}$. $\square$

Un point gênant de notre hypothèse de simplicité est qu'elle écarte le cas où $\partial\Omega$ est lisse, mais avec accès des deux côtés. On peut souvent se débarasser de ce problème en dédoublant la frontière, par application d'une application racine.

Notons également que si Ω_1 et Ω_2 sont bornés, simplement connexes, avec des frontières composées de points simples, les représentations conformes de Ω_1 sur Ω_2 s'étendent également en des homéomorphismes des adhérences: on se contente de passer par le disque unité et composer deux applications conformes.

Et maintenant, un résultat sur les courbes de Jordan (pour ce qui suit, j'utilise Pommerenke).

Théorème 2 (Carathéodory). *Soit $f : \mathbb{D} \to \Omega$ une représentation conforme. Alors f a une extension continue et injective (donc qui est un homéomorphisme de $\overline{D}$ sur $\overline{\Omega}$) si et seulement si $\partial\Omega$ est une courbe de Jordan (un lacet simple).*

Ici on peut voir Ω comme un domaine de la sphère de Riemann $\widehat{\mathbb{C}}$, mais pour simplifier, on va supposer que Ω est un domaine borné de $\mathbb{C}$. On ne va peut-être pas tout démontrer, mais je vais essayer d'en écrire (recopier) un bout.

Théorème 3. *Soit $f : \mathbb{D} \to \Omega$ une représentation conforme, avec Ω borné. Les propriétés suivantes sont équivalentes.*

$$(a) \qquad f \text{ a une extension continue à } \overline{\mathbb{D}};$$

$$(b) \qquad \partial\Omega \text{ est (l'image) d'un arc (continu);}$$

$$(c) \qquad \partial\Omega \text{ est localement connexe;}$$

$$(d) \qquad \mathbb{C} \setminus \Omega \text{ est localement connexe.}$$

C'est bien, parce que ca sera assez facile, modulo un peu de topologie qu'on passera.

En particulier, (b) implique (a) dit que si on peut paramétrer $\partial\Omega$ par un arc (pas de forcément manière injective!), la valeur au bord de f le fait automatiquement.

Cette fois, on autorise des courbes bordées des deux côtés par Ω, qui donneront des paramétrages non injectifs.

Un ensemble fermé E est dit localement connexe quand, pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que, pour tout choix de $a, b \in E$ tels que $|b-a| \leq \delta$, il existe un compact connexe $H_{a,b}$ de diamètre $\leq \varepsilon$ qui contient a et b, et $H_{a,b} \subset E$. On exclue ainsi les bords de peignes. Peut-être qu'on devrait dire "uniformément localement connexe". Et on aura besoin de savoir que

(7) si $\gamma : [0,1] \to \mathbb{R}^n$ est continu, $\gamma([0,1])$ est localement connexe.

En fait, c'est encore vrai pour l'image $\gamma(I)$ d'un compact localement connexe. On se donne $\varepsilon > 0$. On commence par trouver δ_1 tel que $|f(x) - f(y)| \leq \varepsilon/2$ dès que $|x-y| \leq \delta_1$ (par uniforme continuité de f). Puis on recouvre I par un nombre fini de boules (ouvertes) B_j de rayon δ_2, où δ_2 est assez petit pour que si $p, q \in \overline{B}_j$, il existe un compact connexe $H_{p,q}$ de diamètre $\leq \delta_1$, contenu dans I et qui les contienne.

Les $K_j = f(\overline{B}_j)$ sont compacts, de diamètres $\leq \varepsilon/2$, et recouvrent $f(I)$. Prenons $\delta < \delta_1$ assez petit pour que si K_j et K_k sont disjoints, ils soient à distance $> \delta$. Et donnons-nous $a, b \in f(I)$. Donc $a \in K_j$ et $b \in K_k$ pour certains j et k. Et par définition de δ, K_j rencontre K_k en un point c. Alors $a = f(x)$, $b = f(y)$, $c = f(z) = f(w)$, avec $x, z \in \overline{B}_j$ et $y, w \in \overline{B}_k$. Le connexe recherché est $f(H_{x,z}) \cup f(H_{y,w})$. $\square$

Commençons maintenant la démonstration du théorème 3.

Pour "(a) implique (b)", on utilise l'arc $t \to f(e^{2i\pi t})$. On sait que $w = f(e^{2i\pi t}) \in \overline{\Omega}$; vérifions que $w \in \partial\Omega$. Sinon, $w \in \Omega$, et il existe $z \in \mathbb{D}$ tel que $w = f(z)$. Alors, f étant ouverte, si B est un petit disque centré en z contenu dans $\mathbb{D}$, $f(B)$ contient $B(w, \rho)$ pour un $\rho > 0$. Par injectivité, $f(\mathbb{D} \setminus B)$ ne rencontre plus $B(w, \rho)$, et w ne peut lui être adhérent. Cette contradiction montre que $f(\partial\mathbb{D}) \subset \partial\Omega$. Mais $f(\partial\mathbb{D})$ contient $\partial\Omega$, car $f(\overline{D})$ est un fermé qui contient Ω, donc $f(\overline{D})$ contient $\partial\Omega$, et par ailleurs $f(D) = \Omega$ ne rencontre pas $\partial\Omega$. Bref, l'image de $t \to f(e^{2i\pi t})$ est bien $\partial\Omega$.

Pour "(b) implique (c)", on utilise juste (7).

Pour "(c) implique (d)", on se donne $\varepsilon > 0$, et on choisit $\delta > 0$ comme dans la locale connexité de $\partial\Omega$. On se donne $a, b \in \mathbb{C} \setminus \Omega$, avec $|b-a| \leq \delta$. Si $[a,b]$ ne rencontre pas $\partial\Omega$, il est contenu dans $\mathbb{C} \setminus \Omega$, et c'est le compact connexe désiré. Sinon, on prend l'union de $[a,c]$, où c est le premier point de $[a,b]$ qui se trouve dans $\partial\Omega$ (existence facile), de $b, d]$, où d est le premier point de $[b,a]$ qui se trouve dans $\partial\Omega$, et d'un compact connexe de $\partial\Omega$ qui contient c et d. On vérifie que c'est un connexe et que son diamètre est au plus 3ε.

Il reste “(d) implique (a)”, le morceau qui demande une vraie démonstration. On fixe $z \in \mathbb{D}$, disons avec $|z| > 1/2$, et on veut contrôler le diamètre de l'image des petits disques centrés en z.

Notons $C(r) = \mathbb{D} \cap \partial B(z, r)$ pour r petit. Ici, c'est simple, $C(r)$ est un arc de cercle. On veut trouver des rayons tels que $l(r)$, la longueur de $f(C(r))$, soit petite. Or

$$l(r) = \int_{C(r)} |f'(r)| dr \leq \Big\{ \pi r \int_{C(r)} |f'(z)|^2 dl(z) \Big\}^{1/2}$$

(C'est en gros le calcul fait au-dessus de (6)), donc pour ρ petit,

$$\int_\rho^{\sqrt{\rho}} l(r)^2 \frac{dr}{\pi r} \leq \int_\rho^{\sqrt{\rho}} \int_{C(r)} |f'(z)|^2 dl(z) \leq |f(\mathbb{D} \cap B(z, \sqrt{\rho}))| \leq |\Omega|$$

par le calcul de Jacobien de (6). Ce qui permet, par Markov, de trouver $r \in [\rho, \sqrt{\rho}]$ tel que

$$l(r)^2 \leq |\Omega| \Big[\int_\rho^{\sqrt{\rho}} \frac{dr}{\pi r} \Big]^{-1} \leq c |\Omega| \log(1/\rho)^{-1}$$

Bref, ce qui nous intéresse est que $r \in [\rho, \sqrt{\rho}]$ et $l(r) \leq \delta(\rho)$, avec un $\delta(\rho)$ qui tend vers 0 et qui ne dépend pas de z.

Maintenant (en gardant z, ρ, et r), on note γ un paramétrage de $C(r)$ par $[0, 1]$ à vitesse constante, puis $\Gamma = f \circ \gamma([0, 1])$. A vrai dire, supposons que $C(r)$ est vraiment un arc de cercle; le cas où $|z| < 1 - r$ et $C(r)$ est un cercle est plus facile, et laissé au lecteur. Comme Γ est de longueur est finie, les limites $a = \lim_{t \to 0^+} f \circ \gamma(t)$ et $b = \lim_{t \to 1^-} f \circ \gamma(t)$ existent. Elles sont dans $\partial\Omega$, pour la même raison que $f(e^{2i\pi t}) \in \partial\Omega$ ci-dessus, et de plus $|b - a| \leq l(r) \leq \delta(\rho)$. Par locale connexité, on peut trouver $H \subset \mathbb{C} \setminus \Omega$, compact, connexe, qui contienne a et b, et de diamètre $\leq \varepsilon(\rho)$, où $\varepsilon(\rho)$ tend également vers 0 avec ρ, et ne dépend pas de z non plus. On note $G = \Gamma \cup H$. Il est contenu dans un disque B de rayon $\delta(\rho) + \varepsilon(\rho)$. On va montrer que

$$(8) \qquad |f(z) - f(z')| \leq 2\delta(\rho) + 2\varepsilon(\rho) \text{ pour } z \in \mathbb{D} \cap B(z, r).$$

Sinon, l'un au moins des deux points $f(z)$, $f(z')$ est hors du disque B ci-dessus. On va dire que c'est $f(z')$, mais la démonstration serait la même dans l'autre cas.

Noter que comme $f(B(0,1/3))$ contient un disque de rayon $r_0 > 0$ centré en $f(0)$, on voit que si $\rho \leq 10^{-2}$, $C(r)$ ne rencontre pas $B(0,1/3)$, donc $\mathrm{dist}(\Gamma, f(0)) \geq r_0$. et $\mathrm{dist}(G,0) \geq r_0/2$, et pareil pour B. De sorte que G, qui est contenu dans B, ne sépare pas $f(0)$ de $f(z')$.

Il est temps d'appliquer le théorème suivant, dont on ne fera pas la démonstration (mais il est conseillé de faire quelques dessins pour se convaincre qu'on a du mal à trouver des contre-exemples). Voir le livre de Newman.

Théorème (Janiszewski). *Soient A_1 et A_2 deux ensembles fermés de $\mathbb{C}$, tels que $A_1 \cap A_2$ est connexe. Soient a, b, deux points de $\mathbb{C} \setminus (A_1 \cup A_2)$. Si ni A_1 ni A_2 ne séparent pas a de b, alors $A_1 \cup A_2$ non plus.*

On applique ceci à $A_1 = G$ et $A_2 = \mathbb{C} \setminus \Omega$. Comme A_2 non plus de sépare pas $f(0)$ de $f(z')$ (les deux sont dans Ω qui est connexe), et que $A_1 \cap A_2 = H$ est connexe, on voit que $A_1 \cup A_2$ ne sépare pas $f(0)$ de $f(z')$. Donc on peut les joindre par une courbe $\xi \subset \Omega \setminus \Gamma$. L'image ar f^{-1} est une courbe qui va de z' à 0, dans $\mathbb{D}$ et sans rencontrer $C(r)$. Ca, c'est impossible parce que $z' \in B(z,r)$.

Donc on a prouvé (8). Et il est facile d'en déduire la continuité uniforme de f sur $\mathbb{D}$, puis le théorème 3. $\square$

26. PETITE PAUSE ET RAPPELS SUR H^P

On va rappeler ici quelques résultats simples sur la théorie élémentaire des espaces de Hardy H^p, pas toujours avec démonstration.

1. Noyau de Poisson.

Commençons par parler du noyau de Poisson, défini par

$$P_r(\theta) = \frac{1 - r^2}{1 - 2r\cos\theta + r^2} \tag{1}$$

pour $0 \leq r < 1$ et $\theta \in \mathbb{R}$.

D'abord, on calcule: en notant $z = re^{i\theta}$,

(3)

$$\begin{aligned} 1 + \sum_{n=1}^{+\infty} 2r^n \cos(n\theta) &= 1 + 2Re\Big(\sum_{n=1}^{+\infty} z^n\Big) = Re\Big(1 + \frac{2z}{1-z}\Big) = Re\Big(\frac{1+z}{1-z}\Big) \\ &= Re\Big(\frac{(1+z)(1-\overline{z})}{(1-z)(1-\overline{z})}\Big) = \frac{Re\big((1+z)(1-\overline{z})\big)}{1 - 2r\cos\theta + r^2} \\ &= \frac{1-r^2}{1 - 2r\cos\theta + r^2} = P_r(\theta). \end{aligned}$$

Noter que tout converge quand $r < 1$. Aussi, $\frac{1+z}{1-z}$ est analytique, donc harmonique (se souvenir que $\Delta = \partial\overline{\partial}$), donc aussi sa partie réelle. Ainsi $P_r(\theta)$ est une fonction harmonique de z. Un petit coup de dérivation sous le signe somme, dont la vérification est laissée au lecteur (mais tout se passe très bien quand on reste dans un $B(0, r_0)$ avec $r_0 < 1$ donne donc ceci.

Lemme 1. *Si μ est une mesure finie sur $[0, 2\pi]$, en en notant encore $z = re^{i\theta}$,*

$$P\mu(z) =: \frac{1}{2\pi} \int_{[0,2\pi]} P_r(\theta - \varphi) d\mu(\varphi) \tag{4}$$

est une fonction harmonique sur $\mathbb{D}$.

Ensuite, notons que $P_r(\theta) \geq 0$, et $\frac{1}{2\pi} \int_{[0,2\pi]} P_r(\theta) d\theta = 1$ (regarder le membre de gauche de (3)). On va en déduire ceci.

Lemme 2. *Pour f continue sur $\overline{D}$ et harmonique sur $\mathbb{D}$,*

$$f(re^{i\theta}) = \frac{1}{2\pi} \int_{[0,2\pi]} P_r(\theta - \varphi) f(e^{i\varphi}) d\varphi \tag{5}$$

pour $0 \leq r < 1$ et $\theta \in \mathbb{R}$.

Le membre de droite est harmonique sur $/D$, et on vérfie aisément qu'il tend vers $f(e^{i\varphi}$ quand $z \in \mathbb{D}$ tend vers $e^{i\varphi}$ (exercice sur les approximations de l'unité: couper l'intégrale en deux, suivant que $|\theta - \varphi| \leq \varepsilon$ ou pas). Donc la différence avec $f(z)$ est une fonction harmonique, qui a une extension continue sur $\overline{\mathbb{D}}$ et nulle au bord. Le princpe du maximum dit que cette foncytion est nulle. $\square$

Un peu plus délicate est l'étude de la limite au bord de la fonction $P\mu$ de (4). Il se trouve qu'elle a pour presque tout $\theta \in [0, 2\pi]$ une limite radiale

$$\widetilde{f}(e^{i\theta}) = \lim_{r \to 1} f(re^{i\theta}), \tag{6}$$

et que $f(t)dt$ est la partie absolument continue de $d\mu$. C'est (sauf erreur de ma part) ce qu'on appelle le théorème de Fatou. A nouveau, ceci n'utilise que le fait que les P_r sont une approximation de l'unité, et un argument de densité semblable à ceux qu'on a fait plus haut.

Exercice. Retrouver le lemme 2 à partir des ingrédients suivants. D'abord, $f(0) = \frac{1}{2\pi}\int_{[0,2\pi]} f(e^{i\varphi})d\varphi$ (formule de la moyenne). Ensuite, si $\varphi : \mathbb{D} \to \mathbb{D}$ (par exemple) est analytique, et $f : \mathbb{D} \to \mathbb{R}$ est harmonique, alors $f \circ \varphi$ est harmonique. On applique ceci à une transformée de Möbius, et on devrait retrouver (5).

2. Espace h^p.

On va se contenter de parler de $1 \leq p < +\infty$. D'abord, il y les espaces de fonctions harmoniques, qu'on appellera h^p. On dira que $f \in h^p(\mathbb{D})$ si elle est harmonique sur $\mathbb{D}$, et si

$$(7) \qquad ||f||_{h^p} = \sup_{0<r<1} \left\{ \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^p dt \right\}^{1/p}$$

est fini. Il se trouve que, par sous-harmonicité de $|f|$, la fonction $r \to \left\{ \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^p dt \right\}^{1/p}$ est croissante, de sorte que la borne supérieure dans (7) est aussi une limite croissante quand r tend vers 1.

Notons aussi que

$$(8) \qquad h^p \subset h^q \quad \text{pour } 1 \leq p \leq q < +\infty$$

par Hölder, donc on s'attend à avoir parfois un peu plus de mal avec h^1.

Soit $f \in h^1$. D'abord,

$$(9) \qquad f(\rho z) = \frac{1}{2\pi} \int_{[0,2\pi]} P_r(\theta - \varphi) f(\rho e^{i\varphi}) d\varphi$$

par le lemme 2 appliqué à $f(\rho \cdot)$. Puisque $f \in h^1$, les $f(\rho e^{i\cdot})$ sont dans une boule de L^1, on peut donc trouver une suite $\{\rho_n\}$ qui tend vers 1, telle que les $(\rho_n \cdot)$ aient une limite faible μ (qui est une mesure finie, de variation totale $||\mu|| \leq ||f||_{H^1}$). En particulier (9) et la définition d'une limite faible disent que $f = P\mu$.

Si de plus $f \in h^p$ pour un $p > 1$, alors μ est absolument continue, et vient d'une fonction de L^q (cette fois, on utilise le fait que les $f(\rho \cdot)$ définissent des formes linéaires bornées sur L^q).

Dans les deux cas, on ne perd pas d'information, car par exemple, si μ est une mesure finie, il est facile de vérifier que pour tout r, $\theta \to P\mu(re^{i\theta})$, qui est la convolution de μ avec la fonction continue positive d'intégrale 1 P_r, est fans L^1 avec une norme $\leq ||\mu||$. Autrement dit, $P\mu \in h^1$, avec $||P\mu||_{H^1} \leq ||\mu||$.

Ceci dit, on sait bien que $f \in h^1$ a des limites radiales (et même non tangentielles) presque-partout au bord, mais ces limites ne déterminent pas f, puisque dans le cas où $\mu = \delta_0$ et $f(re^{i\theta}) = P_r(\theta)$, la limite est nulle presque-parrtout. Sauf dans le cas des fonctions analytiques, qui vient maintenant!

3. Le théorème de Riesz et Riesz sur H^1.

On note H^p l'espace des fonctions f analytiques sur $\mathbb{D}$, et telles que

$$(10)\qquad ||f||_{H^p} = \sup_{0<r<1} \left\{ \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^p dt \right\}^{1/p}$$

est fini. Donc $f \in H(\mathbb{D})$ est dans H^p quand ses parties réelle et imaginaire sont dans h^p. Et la chose vraiment curieuse est que les mesures avec une partie singulière ne donnent pas de fonctions de H^1.

Théorème. (F. et M. Riesz). *Soient μ une mesure finie, et $P\mu$ comme en (4). Si $P\mu$ est analytique sur $\mathbb{D}$, alors μ est absolument continue, et en fait*

$$d\mu(t) = g(e^{it}),$$

où l'on a noté g la limite radiale de $P\mu$, dont on sait qu'elle existe presque-partout sur le cercle.

On a vu (mais pas redémontré en entier) que $P\mu \in H^1$, qu'elle a une limite radiale g, et même que g est la partie absolument continue de μ, mais il reste à voir que $d\mu(t) = g(t)dt$. On utilisera le lemme suivant.

Lemme 3. *Soit $f \in H^p(\mathbb{D})$, avec $1 \leq p < +\infty$, et notons encore f sa limite radiale (qui existe presque-partout sur $\partial\mathbb{D}$). Alors*

$$(11)\qquad \lim_{r\to 1} \int_{[0,2\pi]} |f(re^{i\theta})|^p d\theta = \int_{[0,2\pi]} |f(e^{i\theta})|^p d\theta$$

et

$$(12)\qquad \lim_{r\to 1} \int_{[0,2\pi]} |f(e^{i\theta}) - f(re^{i\theta})|^p d\theta = 0.$$

Juste quelques commentaires sur ce lemme. En fait, il vaut encore pour $0 < p < 1$, mais l'analyticité des fonctions y est importante.

On commence par observer que quand $p = 2$, $\int_{[0,2\pi]} |f(re^{i\theta})|^2 d\theta$ se calcule par Parseval, et donne $c\sum_n |a_n|^2 r^{2n}$, de sorte que (11) est facile dans ce cas.

Pour le cas général, on sait déjà que

$$\int_{[0,2\pi]} |f(e^{i\theta})|^p d\theta \leq \liminf_{r\to 1} \int_{[0,2\pi]} |f(re^{i\theta})|^p d\theta$$

par Fatou. Pour l'autre inégalité, on utilise le fait que $f \in H^p$ peut s'écrire $f = gB$, où B est un produit de Blaschke qui contient tous les zéros de f (et est de module 1 au bord), et $g \in H^p$ est sans zéro. Cela permet d'appliquer (11) à $g^{2/p}$ et d'en déduire (11) pour f:

$$\begin{aligned}\limsup_{r\to 1}\int_{[0,2\pi]}|f(re^{i\theta})|^p d\theta &\le \limsup_{r\to 1}\int_{[0,2\pi]}|f(re^{i\theta})|^p d\theta\\ &= \int_{[0,2\pi]}|g(e^{i\theta})|^2 d\theta = \int_{[0,2\pi]}|f(e^{i\theta})|^p d\theta.\end{aligned}$$

Ensuite, on peut déduire (12) de (11) par un argument de théorie de la mesure, basé sur le lemme d'Egoroff. Voir par exemple Duren (sur H^p). $\square$

Retour au théorème de F. et M. Riesz. On a vu que $f = P\mu \in H^1$, et on écrit que pour $\rho < 1$, on a (9). On fize z et fait tendre ρ vers 1. Comme (11) dit que $f(\rho e^{i\cdot})$ tend vers $f(e^{i\cdot})$ dans L^1, on trouve que f est aussi l'intégrale de Poisson de ses valeurs au bord. Ainsi, si $d\nu = d\mu - f dt$, on trouve que $P\nu = 0$ dans $\mathbb{D}$. On veut en déduire que $\nu = 0$, et le théorème suivra.

Or si π est une fonction test sur le tore $\mathbb{R}/2\pi\mathbb{Z}$, l'intégrale de Poisson $P(\psi * \nu)$ de la convolution $\psi * \nu$ est nul (c'est la convolution du noyau de Poisson). Or $\psi * \nu$ est continue, et est donc obtenue comme limite au bord de $P(\psi * \nu)$. C'est donc que $\psi * \nu$ est nulle pour tout ψ. Et donc aussi ν (qui est une limite faible d'icelles: tester sur une fonction continue). Fin de démonstration pour le théorème de Riesz et Riesz. $\square$

On en done généralement la version suivante: une mesure sur le cercle, qui est de type analytique (ses coefficients de Fourier négatifs sont nuls), est en fait absolûment continue (donnée par une fonction de L^1). La relation entre les deux est qu'alors l'extension harmnique est analytique sur le disque. C'est assez facile de voir ce dernier fait sur les fonctions de L^2, où la partie positive du spectre de Fourier contribue à une fonction (qui a une extension) analytique sur le disque, à savoir $\sum_{n\ge 0} a_n z^n$, et l'autre morceau a une extension $g(z) = \sum_{n<0} a_n z^{-n}$ analytique sur $\{|z| > 1\}$, et dont $g(1/\overline{z}) = \sum_{n<0} a_n r^{-n} e^{in\theta}$ est l'extension harmonique de $\sum_{n<0} a_n e^{in\theta}$. La somme des deux est l'extension harmonique donnée par le noyau de Poisson.

27. REGULARITE AU BORD DANS LE CAS DE JORDAN RECTIFIABLE

En principe, plus on ajoutera de régularité sur $\partial\Omega$, plus on peut espérer avoir de régularité pour la représentation conforme $f : \mathbb{D} \to \Omega$. Donnons encore un (dernier sans doute) exemple classique.

Théorème 1. *Soit $f : \mathbb{D} \to \Omega$ conforme, et supposons que $\partial\Omega$ est une courbe de Jordan. Alors la longueur de $\partial\Omega$ est finie si et seulement si $f' \in H^1$.*

Ici H^1, est comme au paragrahe précédent l'espace de Hardy des fonctions analytiques sur $\mathbb{D}$ telles que

$$+\infty > ||f||_{H^1} = \sup_{0<r<1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{it})|^p dt \tag{1}$$

Commençons par la partie intéressante: si la longueur de $\partial\Omega$ est finie, alors $f' \in H^1$, en utilisant le minimum d'information sur H^1.

Donc on suppose que $\partial\Omega$ est une courbe de Jordan de longueur finie, et on sait que f a une extension continue à $\overline{\mathbb{D}}$, telle que $f : \partial\mathbb{D} \to \partial\Omega$ est un homéomorphisme.

Longueur finie signifie que

$$L = \sup \Big\{ \sum_{k=1}^{n} |x_k - x_{k-1}| \,;\, x_0 < x_1 < x_2 \cdots < x_n = x_0 \text{ est une subdivision de } \partial\Omega \Big\}$$

est fini. Notre définition de subdivision signifie qu'on a implicitement choisi un sens de parcours sur $\partial\Omega$, et que le x_k sont bien rangés sur $\partial\Omega$. Donc on n'en fait le tour qu'une fois. Et si on choisissait l'orientation inverse, on obtiendrait le même résultat. Et encore, on a bien sûr copié la définition standard de longueur d'une courbe, sauf qu'ici on a un lacet simple. En tout cas, on peut utiliser le paramétrage de $\partial\Omega$ par f, et obtenir

$$L = \sup \Big\{ \sum_{k=1}^{n} |f(e^{it_k}) - f(e^{it_{k-1}})| \,;\, 0 = t_0 < x_1 < x_2 \cdots < t_n = 2\pi \Big\} \tag{2}$$

Evidemment, un autre paramétrage équivalent aurait donné le même L, et on vient juste de dire que la fonction $t \to f(e^{it})$ est à variation bornée sur $[0, 2\pi]$, de variation totale L. Ou encore, qu'elle est l'intégrale d'une mesure finie $\partial\mu$ (à valeurs complexes puisque $\partial\Omega$ est tracée dans $\mathbb{C}$), dont la variation totale est $| \int_0^{2\pi} d\mu| = L$. Il sera utile d'écrire

$$f(e^{i\theta}) = \int_0^{\theta} d\mu(t) \tag{3}$$

où, comme f est continue, on sait que μ n'a pas d'atomes et donc on n'a pas besoin de faire la différence entre $\int_{[0,2\pi]} d\mu$ et $\int_{[0,2\pi[} d\mu$, par exemple.

Notons encore $P_r(\theta) = \frac{1-r^2}{1-2r\cos\theta+r^2}$ le noyau de Poisson. On sait (parce que f est harmonique sur $\mathbb{D}$ et continue sur le disque fermé) que

$$f(re^{i\theta}) = \frac{1}{2\pi}\int_{[0,2\pi]} P_r(\theta-\varphi)f(e^{i\varphi})d\varphi \tag{4}$$

On fixe r et on dérive par rapport à θ. On trouve que

$$ire^{i\theta}f'(\rho e^{i\theta}) = \frac{\partial}{\partial\theta}(f(re^{i\theta}) = \frac{1}{2\pi}\int_{[0,2\pi]} \frac{\partial}{\partial\theta}P_r(\theta-\varphi)f(e^{i\varphi})d\varphi \tag{5}$$

Ensuite $\frac{\partial}{\partial\theta}P_r(\theta-\varphi) = -\frac{\partial}{\partial\varphi}P_r(\theta-\varphi)$, et on utilise (5), plus une intégration par partie justifiée par Fubini (on risque d'en parler au moment de Sobolev; en attendant, voir mes autres notes, l'exemple près de (5) du chapitre 6 sur Sobolev, en notant que f pourrait aussi y être l'intégrale d'une mesure), pour dire que

$$ire^{i\theta}f'(re^{i\theta}) = \frac{1}{2\pi}\int_{[0,2\pi]} P_r(\theta-\varphi)d\mu(\varphi), \tag{6}$$

ou en d'autres termes que

$$izf'(z) \text{ est l'intégrale de Poisson de } d\mu. \tag{7}$$

Alors, les frères Riesz disent que $zf'(z) \in H^1$, et que μ est absolument continue et est donnée par les valeurs au bord de $izf'(z)$. Ce qui est un bon pas vers le fait que f reste conforme au bord (en un sens à préciser).

La réciproque est plus facile. Si $f' \in H^1$, montrons que $L \leq 2\pi||f'||_{H^1}$. Donnons-nous une subdivision comme en (2), où l'on a utilisé le paramétrage de $\partial\Omega$ par f. Choisissons $r < 1$ assez grand pour que

$$\begin{aligned}\sum_{k=1}^{n}|f(e^{it_k}) - f(e^{it_{k-1}})| &\leq \varepsilon + \sum_{k=1}^{n}|f(re^{it_k}) - f(re^{it_{k-1}})| \leq \varepsilon + \sum_{k=1}^{n}\int_{t_k}^{t_{k-1}}|f'(re^{it})|dt \\ &= \varepsilon + \int_0^{2\pi}|f'(re^{it})|dt \leq \varepsilon + 2\pi||f'||_{H^1}\end{aligned}$$

(où l'on a utilisé, bien que ça ne soit pas nécessaire, le fait que $\int_0^{2\pi} |f'(re^{it})|dt$ est une fonction croissante de r.

Ceci est vrai pour toute subdivision; on en déduit que $L \leq 2\pi ||f'||_{H^1}$ en prenant la borne supérieure, et ceci termine notre démonstration du théorème. $\square$

Voir Pommerenke pour pas mal de résultats de régularité des représentations conformes en fonction de la régularité du domaine. En particulier, une étude plus précise des valeurs au bord de la représentation conforme, en des points qui ne sont pas simples (tout sur les "prime ends"). Ou le fait que si $f : \mathbb{D} \to \Omega$ est conforme et $\partial\Omega$ est une otrbe de Jordan, alors $\partial\Omega$ est un arc analytique si et seulement si f se prolonge en une fonction analytique univalente dans un disque $B(0,r)$, avec $r > 1$. Ou encore le fait que lorsque la courbe de Jordan $\partial\Omega$ est de classe $C^{m+\alpha}$, avec $m \geq 1$ et $0 < \alpha < 1$, alors $f^{(m)}$ a une extension continue jusqu'au bord, qui est α-Höldérienne aussi. Mais ceci est faux pour $\alpha = 0$: pour les domaines C^1, f n'est pas forcément C^1.