



HAL
open science

Théorèmes ergodiques des actions de groupes

Claire Anantharaman, Jean-Philippe Anker, Martine Babillot, Aline Bonami,
Bruno Demange, Sandrine Grellier, François Havard, Philippe Jaming,
Emmanuel Lesigne, Patrick Maheux, et al.

► **To cite this version:**

Claire Anantharaman, Jean-Philippe Anker, Martine Babillot, Aline Bonami, Bruno Demange, et al..
Théorèmes ergodiques des actions de groupes. L'Enseignement Mathématique. 41, 2010, Monographies de L'Enseignement Mathématiques, 978-2-940264-08-7. hal-00464094

HAL Id: hal-00464094

<https://hal.science/hal-00464094>

Submitted on 15 Mar 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Claire Anantharaman
Jean-Philippe Anker
Martine Babillot
Aline Bonami
Bruno Demange
Sandrine Grellier
François Havard
Philippe Jaming
Emmanuel Lesigne
Patrick Maheux
Jean-Pierre Otal
Barbara Schapira
Jean-Pierre Schreiber

**THÉORÈMES ERGODIQUES POUR
LES ACTIONS DE GROUPES**

TABLE DES MATIÈRES

Préface	1
Foreword	3
Introduction	9
1. Principes de démonstration d'un théorème ergodique	17
1.1. Le théorème de von Neumann	18
1.2. Inégalités maximales et principe de Banach	22
1.3. Les théorèmes ergodiques de Birkhoff et de Hopf-Dunford-Schwartz	26
2. Théorèmes généraux pour les groupes moyennables	33
2.1. Groupes moyennables	34
2.2. Théorèmes ergodiques pour les suites de Følner	40
2.3. Le principe de transfert	46
3. Inégalités maximales et lemmes de recouvrement	51
3.1. Moyennes sur les boules de \mathbb{R}^d	52
3.2. Mesures doublantes et non doublantes	55
3.3. Suites doublantes d'un groupe	64
3.4. Suites tempérées	68
4. Un exemple : le groupe des affinités	81
4.1. Structure du groupe des affinités	82
4.2. Suites de Følner dans le groupe des affinités	84
4.3. Inégalités maximales pour les boules hyperboliques	88
4.4. Inégalités maximales pour des familles plus générales	94
4.5. Un théorème ergodique sur le groupe S_d	102
5. Moyennes sphériques	105
5.1. Inégalités maximales pour les sphères de \mathbb{R}^d	106
5.2. Inégalités maximales pour les sphères de \mathbb{Z}^d	119

6. Théorèmes ergodiques pour une action d'un groupe libre	141
6.1. Actions par isométries sur un espace métrique compact	143
6.2. Théorèmes ergodiques en moyenne dans L^2	145
6.3. Inégalités maximales et théorèmes ergodiques ponctuels dans L^2	153
6.4. Inégalités maximales et théorèmes ergodiques dans L^p	159
6.5. Le théorème de Bufetov	168
7. Théorèmes ergodiques pour $SO^o(d, 1)$.	173
7.1. Théorème ergodique en moyenne dans L^2	175
7.2. Inégalités maximales et théorème ergodique ponctuel dans L^2	180
7.3. Inégalités maximales et théorème ergodique dans L^p	186
7.4. Estimations des fonctions sphériques	192
Postface	197
A. Régularité des représentations et mesurabilité des fonctions maximales	201
A.1. Représentations de G et de $\mathcal{M}(G)$	201
A.2. Mesurabilité des fonctions maximales	204
B. Éléments de théorie spectrale	207
B.1. C^* -algèbres commutatives	207
B.2. Le théorème spectral	208
C. Théorème de Rota pour une suite d'opérateurs de Markov	213
D. L'espace hyperbolique	219
D.1. Les différents modèles	219
D.2. Isométries de \mathbb{H}^{d+1}	221
D.3. Quelques formules	223
Bibliographie	225
Index	237

PRÉFACE

Ce livre est issu d'un séminaire qui s'est tenu pendant plusieurs années à l'Université d'Orléans. Martine Babillot est à l'origine de ce travail collectif. Son objectif était de rassembler collègues et étudiants en thèse, en faisant interagir leurs compétences complémentaires. De fait, le sujet étudié, par sa richesse et sa diversité, se prête particulièrement bien à la collaboration d'analystes de cultures différentes.

Amos Nevo, qui a beaucoup contribué à l'impressionnant développement de la théorie ergodique des actions de groupes, a fait plusieurs séjours à Orléans. Son aide a été précieuse et nous l'en remercions ici.

Au moment de la disparition de Martine Babillot, la rédaction de cet ouvrage était déjà bien entamée. Elle a été poursuivie ensuite, malheureusement à un rythme beaucoup moins soutenu, l'absence de l'enthousiasme communicatif de Martine se faisant bien sûr ressentir. Nous lui dédions ce livre. Sans elle, il n'aurait pas existé.

Orléans, Janvier 2009

FOREWORD

by Amos NEVO

The present book is devoted to ergodic theorems for group actions, and offers the reader a valuable exposition of this subject, which is novel in its approach and stands out in at least three respects.

The first is the authors' remarkable success in presenting the development of the subject matter. The book starts at the very foundation of the subject and gives a thorough exposition of many of the classical arguments in ergodic theory. Building upon this foundation, the book then proceeds to give an equally thorough exposition of several advanced and contemporary topics, while keeping its lucid and readable style. Indeed, many of the results presented and explained in the book are of recent and sometimes very recent vintage, at the forefront of current research. Furthermore, the book succeeds in leading the reader towards a view of a host of open problems which are the subject of active ongoing research efforts.

The second is the range of topics discussed in the book, which taken together gives the reader an unusually broad perspective on the subject matter. Two major distinctions arising naturally in ergodic theory are those between groups of polynomial volume growth and groups of exponential volume growth, and between amenable groups and non-amenable groups. The subjects discussed in the book span both of these divides. The reader will find a discussion of the classical arguments associated with polynomial volume growth, the doubling condition, and the associated geometric covering arguments. Later on one finds a discussion and some novel results on the group of affinities of Euclidean space, serving as an introduction to the difficult (and wide open) subject of ergodic theorems for solvable groups of exponential volume growth. Further along the reader will find a detailed presentation of the classical arguments of amenable ergodic theory, including ergodic theorems for asymptotically invariant sequences, culminating in the recent ergodic theorem for tempered sequences. The presentation then moves to the subject of ergodic theory of non-amenable groups, focusing in particular on ergodic theorems for radial averages on the free groups. Here both the spectral approach and Markov operators approach to the proof of the

pointwise ergodic theorem are presented. Finally the subject of ergodic theorems for radial averages on Lie groups is taken up. The presentation focuses on an exposition of ergodic theorems for sphere averages in actions of the isometry groups of hyperbolic spaces, including the requisite preliminaries on the structure theory of the groups and on the spectral theory of spherical functions.

The third concerns the unusually wide range of analytic methods that the book utilizes and demonstrates. Indeed, classical amenable ergodic theory has emphasized reliance on geometric comparison arguments associated with the growth and shape of sets in the group, on the associated covering arguments based on the doubling condition and its generalizations, and on the use of asymptotically invariant sets and the transference principle. Of course, none of these arguments apply directly to the ergodic theory of non-amenable groups. One of the main points in the approach the book describes to the non-amenable case is that one needs to deploy a much wider range of analytic techniques than those used in the amenable case. These techniques include arguments employing unitary representation theory, methods of spectral analysis, real and complex interpolation arguments, and singular integrals techniques such as Tauberian theory and the method of square functions. The book gives the first detailed systematic account of the use of such methods in the proofs of ergodic theorems, and thus makes an important contribution to the development of this emerging aspect of ergodic theory.

Let us demonstrate some of the points mentioned above by discussing several examples, to which the results discussed in the book directly apply, and which yield interesting comparisons with classical amenable ergodic theory. Full references to the results and conjectures mentioned below can be found in the main text.

1) *Euclidean sphere averaging.* — Let f be a function on the Euclidean plane which is doubly periodic, namely invariant under a lattice subgroup L (or equivalently, a function lifted from the factor torus $T = \mathbb{R}^2/L$). Fix a line ℓ with slope incommensurable with the lattice directions, and a point x in the plane. Averaging f on symmetric intervals on the line through x parallel to ℓ , Birkhoff's theorem asserts that these averages converge for almost every x to the average of f on a fundamental parallelogram, for any $f \in L^1(T)$.

It seems equally natural to consider the average of f on a circle of radius t or an annulus of radii t and $t + 1$, and center x . When do these averages converge to the average of f on a fundamental parallelogram, for almost all center points x ? The answer is that convergence holds when f is in $L^p(T)$, as long as $p > 2$, and this range is sharp: there are L^2 functions for which the averages diverge for almost every x . This radial ergodic theorem is ultimately based on a deep maximal inequality due to Bourgain, and one of the ideas introduced in its proof is to start the analysis in L^4 , and proceed from there by intricate analytic smoothing techniques.

In the case of triply periodic functions on \mathbb{R}^3 , and averaging on spheres or on spherical shells (namely the domain between concentric spheres of radii between t and $t+1$), the answer is that convergence holds almost everywhere for f in L^p , $p > 3/2$, and this range is sharp. This result is ultimately based on Stein's fundamental spherical maximal inequality, which starts with a Fourier transform estimate in L^2 and works its way down to $L^{3/2}$ via interpolation. The reader will find a complete account of ergodic theorems for sphere averages in \mathbb{R}^d , $d \geq 3$, and the attendant analytic techniques in Chapter 5.

2) *Discrete sphere averaging.* — Given the foregoing discussion, it is natural to consider the problem of establishing ergodic theorems for actions of \mathbb{Z}^d , where the averages σ_r are uniformly distributed on the discrete spheres, namely on

$$S_r^d = \{(m_1, \dots, m_d) \in \mathbb{Z}^d; m_1^2 + \dots + m_d^2 = r\}.$$

This challenging problem has been completely resolved recently, with the strong maximal inequality holding if and only if $f \in L^p$, $p > \frac{d}{d-2}$ and $d \geq 5$. A complete and accessible account of this intricate and analytically demanding maximal inequality is given in Chapter 5. We remark that this line of development can be viewed as higher-dimensional generalizations of some of the pioneering ergodic theorems established by Bourgain for sparse sets of integers, such as the squares and the primes. The latter results have already been given very good expositions in the literature, and the present book provides the first such exposition for the higher dimensional case.

3) *Non-Euclidean sphere averaging.* — Let \mathcal{L} be the space of unimodular lattices in \mathbb{R}^2 . \mathcal{L} can be identified with the homogeneous space $SL_2(\mathbb{R})/SL_2(\mathbb{Z})$, and thus carries an invariant probability measure. Letting $\alpha(L) = \|v\|^{-1}$ where $v \neq 0$ is a vector of minimal length in the lattice L , $\alpha(L)$ is in $L^p(\mathcal{L})$, where $1 \leq p < 2$. Let us lift the function $\alpha(L)$ from the space of unimodular lattices \mathcal{L} to $SL_2(\mathbb{R})$. We can then average it with respect to probability measures on $SL_2(\mathbb{R})$, for example those lifted from circle or annulus averages on the hyperbolic plane $\mathbb{H}^2 = SO_2(\mathbb{R}) \backslash SL_2(\mathbb{R})$. The averages of $\alpha(L)$ on such sets are of great interest and appear in various arguments in dynamics on the modular surface (with similar averages appearing in higher dimensions). Do these averages converge, starting from almost every lattice? The answer here is quite subtle, since the circle averages on \mathcal{L} satisfy the pointwise ergodic theorem in L^p only for $p > 2$, but the annuli satisfy such a theorem for all $p > 1$ (and all measure-preserving actions). Thus the annuli averages of $\alpha(L)$ converge almost surely (as the radius increases) to a constant, equal to the average of α on the space of unimodular lattices. Note that the problem of whether the circle averages on $SL_2(\mathbb{R})$ converge in L^p , $p > 2$, for every measure-preserving ergodic action is still open.

A complete discussion of the pointwise ergodic theorem for sphere averages on $SO^0(d, 1)$ for $d \geq 3$ is given in Chapter 7, together with an exposition regarding

spherical functions and the Lie structure theory involved. The problem of whether the annuli (and ball) averages converge pointwise for every $f \in L^1$ is still open.

4) *Polynomial versus exponential volume growth.* — Taking the preceding discussion in parts 1) and 3) together, we encounter an interesting phenomenon which manifests the basic difference between Euclidean and non-Euclidean spaces. Whereas Euclidean annuli have the same convergence properties as Euclidean *circles*, non-Euclidean annuli have the same convergence properties as non-Euclidean *balls*. Note that an annulus of radii t and $t + 1$ occupies only a fraction of size at most c/t of the volume of the Euclidean ball of radius t , but occupies a fixed proportion of the volume of the non-Euclidean ball of radius t . Thus the difference between polynomial volume growth in the Euclidean case, and exponential volume growth in the non-Euclidean case manifests itself in the ergodic properties of annulus averages in the two cases. The reader is invited to explore the problem of whether this dichotomy persists further in the context of nilpotent versus exponential-solvable connected Lie groups.

5) *Ergodic theorems for general groups of transformations.* — The mean and pointwise ergodic theorems of von Neumann and Birkhoff apply to an arbitrary measure-preserving transformation on a probability space. Choosing for example a rotation acting on the unit sphere $S^2 \subset \mathbb{R}^3$, the existence of a rotation axis implies that the dynamics is very constrained. If we are interested in generating an equidistributed group orbit on the sphere then we must choose *two* orthogonal matrices in $SO_3(\mathbb{R})$, which generate a dense subgroup. Viewing this choice as a representation of the free group on two generators, we are ineluctably led to consider the ergodic theory of free groups, acting by isometries on a compact metric space, or more generally by measure-preserving transformations on a probability space. How should we define averages on the free group which will yield analogs of the classical ergodic theorems? The suggestion of using the averages on word-length spheres in the free group and the corresponding averaging operators on functions was raised by Arnold and Krylov who proved the equidistribution theorem on S^2 . The same averages were used later on by Guivarc'h, who proved the mean ergodic theorem for general measure-preserving actions. The next step of course is to ask whether these averages, acting on a function f , converge to the integral of f with respect to the invariant probability measure, for almost every orbit? The current answer in the literature is that this holds when f is in L^p for $p > 1$, and even for f in the Orlicz space $L \log L$. This line of development is given a complete account in Chapter 6. Note that the problem of showing that almost sure convergence holds for every f in L^1 is still open.

The examples and problems mentioned above are primarily of an analytic nature, and attest to the close connection and mutual influence between ergodic theory and analysis, and particularly the theory of singular integrals. They also attest to the fact

that the natural problem of developing ergodic theorems for singular geometric averages (such as spheres) necessarily leads us to employ a variety of analytic techniques in L^p -spaces. This is primarily a consequence of the nature of the averaging problems themselves, which exhibit different behavior in different L^p spaces, a fact which can be quite surprising. Sometimes the L^p analysis is necessitated simply because it is the best we can do at this point in time, for example having to settle for a result in $L^p, p > 1$, when the L^1 result eludes us. In general, the discussion the book provides shows that ergodic theory, particularly when dealing with non-amenable groups, can benefit greatly from analytic techniques, many of which involve extensive use of interpolation arguments in L^p .

The previous examples also serve to briefly demonstrate to the reader, through the prism of ergodic theorems, two further important facts. The first is the remarkable diversity of ergodic theory and its complex interactions with other branches of mathematics, including real and harmonic analysis, integral geometry, operator algebras, representation theory, and aspects of group theory and number theory, among others. The second is the vitality of the theory, which offers a wide array of compelling open problems, together with a varied and sophisticated toolbox which can be used to approach at least some of them.

Therefore it is natural to expect that the present book will play an important role by offering a thorough, readable, and wide-ranging exposition of ergodic theorems for group actions to a new generation of ergodic theorists.

In closing, I would like to recall that the origin of the collaborative effort that produced the present book was Martine Babillot's initiative to organize a study group at the University of Orléans devoted to its subject matter. I had the privilege of visiting Orléans several times during this period, and I would like to take this opportunity to thank all members of the department for their hospitality.

The preparation of the book was greatly disrupted by Martine's tragic and untimely death. Looking back on the development of this project, one is struck by how much the book ultimately reflects the very traits that made Martine such a valued colleague and such a valued friend. It is the breadth and inclusiveness of her mathematical interests, together with her enthusiasm, friendliness and ability to bring people together, that are primarily responsible for the success of this interdisciplinary collaborative effort. I join the authors and all of Martine's friends in the hope that this valuable book will constitute a lasting tribute to her memory.

Amos Nevo, November 2009

INTRODUCTION

À la suite des travaux de précurseurs comme Poincaré, la *théorie ergodique* des systèmes dynamiques est née vers 1930 avec les théorèmes de von Neumann et de Birkhoff, qui donnèrent un cadre mathématique rigoureux pour expliquer l'hypothèse d'ergodicité en mécanique statistique. Cette hypothèse, formulée par Boltzmann et Gibbs, affirme que, dans l'évolution d'un système de particules, presque toutes les trajectoires s'équirépartissent sur les surfaces d'énergie constante dans l'espace des phases.

La théorie ergodique s'est beaucoup développée et interagit avec de nombreuses branches des mathématiques, comme la théorie des probabilités, l'analyse fonctionnelle, la théorie des nombres, la géométrie et la topologie différentielle. Nous mettrons ici en avant ses liens avec la théorie des groupes et, au travers des inégalités maximales, avec l'analyse fine.

L'objet de la théorie ergodique est l'étude des systèmes dynamiques du point de vue de la théorie de la mesure. Un *système dynamique mesuré* est un espace mesuré (X, \mathcal{B}, m) muni d'une transformation τ (ou d'un semi-groupe à un paramètre $(\tau_t)_{t \geq 0}$) qui préserve la mesure m . Le plus souvent la mesure m est une probabilité. En théorie ergodique on s'intéresse à la description des propriétés statistiques des trajectoires des points sous l'itération de la transformation et on établit une classification des systèmes dynamiques mesurés.

Un point de départ de l'étude de ces propriétés est l'établissement de *théorèmes ergodiques*, qui décrivent le comportement asymptotique des moyennes "temporelles"

$$A_n f(x) = \frac{1}{n} \sum_{k=0}^{n-1} f(\tau^k x),$$

où f est une fonction mesurable sur X , à valeurs réelles. Parmi les premiers résultats théoriques importants figurent, comme nous l'avons dit, ceux de von Neumann et de Birkhoff. Pour une fonction f de carré intégrable, von Neumann a montré que $(A_n f)$ converge en moyenne quadratique vers une fonction laissée invariante par τ , tandis

que Birkhoff a, de son côté, établi un théorème de convergence presque partout. Dans le cas où le système dynamique possède la propriété d'irréductibilité naturelle appelée ergodicité, la limite de la suite $(A_n f)$ est la valeur moyenne de la fonction f pour la probabilité m : le théorème ergodique donne ainsi une description mathématique de l'hypothèse ergodique, c'est-à-dire la coïncidence entre moyennes spatiale et temporelle, quand le temps est grand. Les lois des grands nombres du Calcul des Probabilités sont des cas particuliers de théorèmes ergodiques.⁽¹⁾

La transformation τ de l'espace (X, \mathcal{B}, m) définit une action du semi-groupe \mathbb{N} par $(n, x) \in \mathbb{N} \times X \mapsto nx = \tau^n x$. Quand la transformation est inversible, elle définit une action du groupe \mathbb{Z} . Dans le présent ouvrage, nous nous intéressons principalement aux extensions des théorèmes ergodiques dans le cadre plus général d'une action de groupe. Pour pouvoir énoncer l'analogie des deux théorèmes ergodiques, de Birkhoff et von Neumann, il nous faut introduire l'analogie des moyennes temporelles qui y interviennent. Si μ est une mesure de probabilité sur G et si $f \in L^1(X, m)$, nous définissons la moyenne $\mu \cdot f$ par la formule

$$\mu \cdot f(x) = \int_G f(gx) d\mu(g).$$

Ainsi, la fonction $A_n f$ ci-dessus n'est autre que la moyenne de f relativement à la mesure de probabilité uniforme μ_n sur $\{0, 1, \dots, n-1\}$. La question suivante est au cœur de nos préoccupations : pour quelles suites (μ_n) de mesures de probabilité sur G , et pour quels espaces L^p , les théorèmes de von Neumann et de Birkhoff restent-ils vrais ? Il est également naturel de se pencher sur le même problème avec des familles $(\mu_r)_{r>0}$ indexées par \mathbb{R}_+^* . De telles moyennes interviennent dans l'étude des systèmes dynamiques et sont aussi étroitement reliées à des questions fondamentales en analyse réelle et harmonique. Le théorème de différentiation de Lebesgue en est une illustration élémentaire. Son cadre est celui de l'action de \mathbb{R}^d sur lui-même par translations. En prenant pour β_r la probabilité uniforme sur la boule de centre 0 et de rayon r , on a

$$\beta_r \cdot f(x) = \frac{1}{|B(0, r)|} \int_{B(0, r)} f(g+x) dg = \frac{1}{|B(x, r)|} \int_{B(x, r)} f(g) dg,$$

où $|B(x, r)|$ désigne le volume de la boule de centre x et de rayon r . Rappelons que le théorème de différentiation de Lebesgue dit que, pour tout $f \in L^1(\mathbb{R}^d)$, et pour

⁽¹⁾La formulation des théorèmes de von Neumann et de Birkhoff s'étend naturellement dans diverses directions. En notant P l'application $f \mapsto f \circ \tau$, remarquons que $A_n f$ s'écrit $\frac{1}{n} \sum_{k=0}^{n-1} P^k f$. On peut alors, plus généralement, se donner un opérateur borné P de $L^p(X, m)$, $p \geq 1$, et étudier la suite $(\frac{1}{n} \sum_{k=0}^{n-1} P^k f)$. On peut aussi considérer un opérateur borné (ou même un semi-groupe d'opérateurs) d'un espace de Banach plus général qu'un espace L^p . Dans cet ouvrage, nous aborderons très peu cet aspect, qui est traité en détail dans le livre de Krengel [Kre85]. Par ailleurs, on peut remplacer la transformation "déterministe" τ par une règle de transition aléatoire : c'est le domaine des théorèmes ergodiques pour chaînes de Markov, qui ne sera pas étudié ici.

presque tout $x \in \mathbb{R}^d$, on a $\lim_{r \rightarrow 0} \beta_r \cdot f(x) = f(x)$. Cet aspect local ne sera pas étudié dans notre ouvrage et nous nous limiterons à l'étude du comportement des moyennes $\mu_r \cdot f$ quand r tend vers $+\infty$. Nous ferons toutefois appel aux outils qui sont traditionnellement utilisés dans les généralisations du théorème de Lebesgue, à savoir les fonctions maximales et les lemmes de recouvrement.

On distingue deux types de théorèmes ergodiques, les théorèmes ergodiques en moyenne, où on étudie la convergence dans L^p , et les théorèmes ergodiques ponctuels, qui traitent de la convergence presque partout. Les théorèmes en moyenne étendent le théorème de von Neumann et sont généralement plus accessibles que les théorèmes ponctuels. Ils se démontrent, le plus souvent, par des méthodes de point fixe ou bien en utilisant la théorie spectrale dans L^2 . La démonstration des théorèmes ponctuels passe en général par deux étapes : on obtient d'abord la convergence presque partout pour les fonctions d'une famille dense dans $L^p(X, m)$, puis à l'aide d'une inégalité maximale on étend le résultat à tout $L^p(X, m)$. Précisons que ces inégalités concernent l'opérateur maximal $f \mapsto f_\mu^* = \sup_{r>0} |\mu_r \cdot f|$. Ce dernier est dit de type fort (p, p) (ou borné sur L^p) s'il existe $C > 0$ tel que $\|f_\mu^*\|_p \leq C\|f\|_p$ pour tout $f \in L^p(X, m)$. Si on a seulement $m(\{f_\mu^* > \alpha\|f\|_p\}) \leq (C^p/\alpha^p)$ pour tout $\alpha > 0$ et tout $f \in L^p(X, m)$, on dit que l'opérateur maximal est de type faible (p, p) .

Notre objectif est de présenter un panorama des théorèmes ergodiques pour des actions de groupes. Nous partons des résultats classiques des débuts de la théorie ergodique. Nous traitons après le cas des actions de groupes moyennables, qui fut longtemps le cadre naturel de cette théorie. Ensuite nous exposons des travaux réalisés essentiellement par Nevo ces quinze dernières années. Consacrés à des actions de groupes non moyennables comme les groupes libres, les groupes hyperboliques ou les groupes de Lie semi-simples non compacts, ces travaux marquent un tournant dans l'évolution de la théorie.

Concernant un sujet aussi vaste, nous avons dû faire des choix. Deux préoccupations nous ont guidés dans notre travail :

- d'une part être accessible à tout lecteur familier avec les bases de l'analyse fonctionnelle, de la théorie de la mesure et des probabilités ;
- d'autre part dégager un certain nombre d'idées fondamentales et les mettre en œuvre dans des contextes variés, tout en donnant des pistes pour des lectures complémentaires.

Venons-en à une description plus détaillée du contenu du livre. Nous commençons dans le chapitre 1 par un exposé élémentaire du théorème de von Neumann puis du théorème de Birkhoff sous sa forme plus générale due à Hopf et à Dunford et Schwartz.

Dans le chapitre 2, nous passons du cas des actions du groupe \mathbb{Z} au cas des actions d'un groupe moyennable G . Historiquement, les actions des groupes \mathbb{Z}^d et \mathbb{R}^d furent d'abord étudiées, puis on s'est progressivement aperçu que les moyennes sur des suites de Følner conduisent aux mêmes résultats que les moyennes sur des boules de \mathbb{R}^d par

exemple, du moins en ce qui concerne les théorèmes ergodiques en moyenne. Plus généralement, nous considérons une famille (μ_r) asymptotiquement invariante (ou de Følner) de mesures de probabilité sur G , c'est-à-dire telle que $\lim_{r \rightarrow +\infty} \delta_g * \mu_r - \mu_r = 0$ pour tout $g \in G$ (ici $r \in \mathbb{N}$ ou $r \in \mathbb{R}_+^*$) ⁽²⁾. Alors des techniques classiques en analyse fonctionnelle montrent que, pour tout $f \in L^p(X, m)$, la famille $(\mu_r \cdot f)$ converge dans L^p vers une fonction G -invariante lorsque r tend vers $+\infty$ (sous réserve que m soit finie si $p = 1$). Nous obtenons facilement aussi la convergence m -presque partout pour toutes les fonctions f d'une partie dense dans $L^p(X, m)$. La convergence presque partout pour toutes les fonctions de $L^p(X, m)$ n'est pas toujours satisfaite et dépend de propriétés de (μ_r) plus subtiles que l'invariance asymptotique. Suivant le principe de Banach exposé dans le chapitre 1, il suffit, pour étendre cette convergence à tout $L^p(X, m)$, de montrer que l'opérateur maximal $f \mapsto f_\mu^* = \sup_r |\mu_r \cdot f|$ est de type faible (p, p) . Lorsque G est moyennable, ceci peut être établi grâce au principe de transfert de Wiener et Calderón. Le chapitre 2 se termine par la démonstration de ce principe qui permet de se limiter, pour l'étude d'inégalités maximales, au cas de l'action de G sur lui-même par translations.

Dans le chapitre 3, nous ne considérons que des mesures μ_r absolument continues par rapport à la mesure de Haar à gauche λ sur un groupe localement compact G . Plus précisément, on se donne une famille (F_r) de parties mesurables de G de mesure de Haar $|F_r|$ finie et μ_r est la mesure de densité $(1/|F_r|)\mathbf{1}_{F_r}$. Nous présentons plusieurs techniques de recouvrement permettant d'établir des inégalités maximales pour l'opérateur $\varphi \mapsto \varphi_\mu^*$ relatif à l'action à gauche de G sur lui-même. Nous commençons par le lemme classique de recouvrement de Vitali pour les boules de \mathbb{R}^d , d'où l'on déduit le théorème ergodique de Wiener relatif aux moyennes sur ces boules.

Dans la deuxième section du chapitre 3, nous quittons momentanément le cadre des actions de groupes pour étudier les moyennes centrées et non centrées sur les boules de \mathbb{R}^d , en effectuant ces moyennes non plus suivant la mesure de Lebesgue, mais suivant une mesure de Radon quelconque sur \mathbb{R}^d . Ce sujet intervient dans de nombreux problèmes classiques en analyse, comme l'étude de la différentiation des intégrales, l'étude du comportement au bord des fonctions harmoniques, ou encore la théorie des opérateurs intégraux de Calderón-Zygmund. Si la mesure considérée conserve la même propriété de doublement de volume que la mesure de Lebesgue, l'opérateur maximal est encore de type faible $(1, 1)$, que les boules soient centrées ou non. Dans le cas où la mesure d'une boule n'est plus comparable à celle de la boule de rayon double, le lemme de recouvrement de Besicovitch permet d'établir que l'opérateur maximal relatif aux moyennes sur les boules centrées est de type faible $(1, 1)$. En revanche, pour les moyennes non centrées nous donnons un exemple, dû à Sjögren, qui montre la nécessité, dans ce cas, d'hypothèses supplémentaires sur la mesure de Radon considérée.

⁽²⁾L'existence d'une telle famille équivaut à la moyennabilité de G .

Après un examen rapide de l'opérateur maximal relatif à certains réseaux dyadiques, dont l'étude se ramène au lemme maximal de Doob pour les martingales, nous revenons au cas de l'action d'un groupe sur lui-même. Nous terminons le chapitre 3 par le théorème maximal d'E. Lindenstrauss relatif aux suites tempérées de G , c'est-à-dire aux suites (F_n) de parties de G telles que

$$\sup_n \frac{|\bigcup_{1 \leq i \leq n-1} F_i^{-1} F_n|}{|F_n|} < +\infty.$$

À l'aide d'un lemme de recouvrement probabiliste, on montre que l'opérateur maximal relatif aux moyennes sur ces F_n est de type faible $(1, 1)$. Tout groupe moyennable possède des suites de Følner tempérées. On peut donc obtenir des théorèmes ergodiques dans le cadre général des actions de groupes moyennables à condition de moyenner par rapport à de telles suites. Ce résultat généralise le théorème de Wiener et de nombreux théorèmes ergodiques obtenus après lui pour les actions de certains groupes moyennables (essentiellement les groupes à croissance polynomiale) ; on pense en particulier aux travaux de Calderón, Tempelman et Emerson.

En dehors du cas des groupes à croissance polynomiale, les suites tempérées sont généralement obtenues de façon abstraite et n'ont pas de signification géométrique. L'étude de théorèmes ergodiques et d'inégalités maximales relatifs à des moyennes sur des sous-ensembles liés à la géométrie d'un groupe résoluble (et donc moyennable) à croissance exponentielle est un problème très intéressant, largement ouvert. Le chapitre 4 est consacré à l'étude d'un exemple, celui du groupe S_d des affinités de \mathbb{R}^d . Ce groupe résoluble, à croissance exponentielle, possède une métrique naturelle, invariante à gauche, héritée de son identification avec l'espace hyperbolique \mathbb{H}^{d+1} . Les boules hyperboliques centrées sur l'élément neutre ne forment pas une famille de Følner. Nous commençons par donner des exemples explicites et simples de telles familles. Ensuite nous montrons que l'opérateur maximal relatif aux moyennes sur les boules hyperboliques (pour l'action par translations à droite de S_d sur S_d) est de type faible $(1, 1)$, et nous examinons la validité des inégalités maximales pour les moyennes sur d'autres exemples naturels de familles de parties de S_d . Ce groupe n'étant pas unimodulaire, on s'aperçoit qu'il y a lieu de distinguer entre actions à gauche et à droite de S_d sur lui-même. Notre étude nous permet finalement d'établir des théorèmes ergodiques pour certaines familles de Følner dont les éléments, géométriquement simples, forment des parallélépipèdes dans un système de coordonnées approprié de S_d .

Nous quittons ensuite le cadre classique de l'étude de la convergence de $(\mu_r \cdot f)$, où (μ_r) est asymptotiquement invariante, et de nouvelles techniques vont être mises en œuvre, faisant notamment jouer un rôle central à l'espace L^2 , l'analyse de Fourier, la théorie de Littlewood-Paley-Stein et la théorie spectrale. La première partie du chapitre 5 est consacrée aux moyennes sphériques sur \mathbb{R}^d . L'opérateur maximal associé est de type fort (p, p) si $p > d/(d-1)$ et $d \geq 2$, et les restrictions sur p et d sont optimales. Pour $d = 2$, le résultat est dû à Bourgain. On a alors $p > 2$, ce qui

rend difficile le recours à l'analyse de Fourier et nécessite d'autres outils. Nous nous contentons de donner une démonstration du résultat, dû à Stein, pour $d \geq 3$. Elle repose sur une comparaison des moyennes sur les sphères et sur les boules faisant apparaître une fonctionnelle de Littlewood-Paley, et utilise une décomposition de Littlewood-Paley sur des couronnes dyadiques.

Pour les moyennes sphériques, les théorèmes ergodiques en moyenne se démontrent de façon élémentaire. Il suffit de considérer le cas L^2 , où l'on utilise une méthode spectrale faisant seulement intervenir le comportement à l'infini de la transformée de Fourier de la mesure de probabilité uniforme σ sur la sphère unité de \mathbb{R}^d . Nous nous intéressons ensuite au théorème ergodique ponctuel, valable pour $d \geq 2$ et $p > d/(d-1)$. Il est dû à Jones dans le cas $d \geq 3$ et à Lacey pour $d = 2$. Sa démonstration est nettement plus délicate, surtout dans le cas $d = 2$ que nous écartons encore de notre exposition.

La deuxième partie du chapitre 5 est consacrée au théorème maximal pour les moyennes sphériques sur \mathbb{Z}^d , obtenu plus récemment par Magyar, Stein et Wainger, où l'on voit apparaître des techniques de discrétisation d'opérateurs classiques en analyse harmonique. Des difficultés de nature arithmétique, liées à la décomposition d'entiers en sommes de carrés, nécessitent aussi l'utilisation de méthodes issues de la théorie des nombres.

Les deux derniers chapitres introduisent le lecteur à un ensemble de travaux développés ces quinze dernières années par Nevo et Stein principalement, où sont démontrés des théorèmes ergodiques pour d'importantes classes de groupes non moyennables. Le chapitre 6 étudie l'action d'un groupe libre, cadre de la généralisation la plus naturelle des théorèmes de von Neumann et de Birkhoff (qui concernent le cas du groupe libre \mathbb{Z} à un générateur). Il débute par une observation essentielle, apparaissant déjà dans des travaux d'Arnold et Krylov, à savoir que l'algèbre de convolution $\ell_{Rad}^1(\mathbb{F}_d)$ formée des fonctions radiales sommables sur le groupe libre \mathbb{F}_d à d générateurs (muni de la distance définie par la longueur des mots) est engendrée par la mesure de probabilité uniforme σ_1 sur la sphère unité de rayon 1. Elle est par conséquent commutative et elle contient évidemment la mesure de probabilité uniforme σ_n sur la sphère de rayon n pour tout n . De plus on connaît explicitement ses caractères : ce sont les fonctions sphériques de type positif sur \mathbb{F}_d . Une étude spectrale dans L^2 conduit à des théorèmes ergodiques en moyenne pour certaines suites de mesures de probabilité sur \mathbb{F}_d , comme les moyennes de Cesàro ν_n des σ_n , ainsi qu'à des résultats de convergence presque partout pour des familles denses dans $L^2(X, m)$. À l'aide d'une inégalité maximale sous-additive, nous montrons, en suivant Nevo, que l'opérateur maximal associé aux ν_n est de type fort $(2, 2)$. Ce fait est ensuite utilisé, avec des estimations fines des caractères de $\ell_{Rad}^1(\mathbb{F}_d)$, pour en déduire que l'opérateur maximal relatif aux moyennes sphériques sur \mathbb{F}_d est lui aussi de type fort $(2, 2)$. On compare les moyennes

sur les sphères et les moyennes relatives aux mesures ν_n par une méthode calquée sur la méthode de Littlewood-Paley-Stein utilisée pour les sphères de \mathbb{R}^d .

Cette méthode de Littlewood-Paley, accompagnée de techniques d'interpolation complexe dues à Stein, permet de montrer que les opérateurs maximaux associés à (σ_n) et (ν_n) sont respectivement de type fort (p, p) si $p > 1$ et de type faible $(1, 1)$. On n'a pas de théorème ergodique relatif à la suite σ_n ; néanmoins, on voit par exemple que pour toute action de \mathbb{F}_d sur (X, \mathcal{B}, m) préservant la mesure de probabilité m , pour tout $p > 1$ et tout $f \in L^p(X, m)$, la suite $\sigma_{2^n} \cdot f$ converge presque partout et dans L^p vers l'espérance conditionnelle de f sachant la tribu des invariants sous l'action des mots de longueur paire dans \mathbb{F}_d . En dernière partie de ce chapitre 6, nous donnons une nouvelle démonstration de ce résultat de Nevo-Stein, due à Bufetov. Une observation simple et élégante permet de déduire le théorème de Nevo-Stein du théorème ergodique de Rota pour un opérateur de Markov bien choisi. Il permet même d'étendre le résultat aux fonctions de $L \log L(X, m)$.

L'un des intérêts de la méthode développée par Nevo pour les groupes libres est qu'elle s'adapte à beaucoup d'autres exemples, dont les groupes de Lie connexes semi-simples, sans facteur compact, à centre fini. Dans le chapitre 7 nous exposons le cas du groupe $G = SO^0(d, 1)$ des isométries positives de l'espace hyperbolique \mathbb{H}^d de dimension d . Cet espace \mathbb{H}^d s'identifie à l'espace symétrique G/K où K est le sous-groupe compact maximal $SO(d)$ de G . Pour $r > 0$, la sphère formée des points de G qui s'envoient sur la sphère hyperbolique centrée à l'origine dans G/K , de rayon r , porte une unique mesure de probabilité σ_r bi- K -invariante. Un fait crucial est la commutativité de l'algèbre des mesures finies bi- K -invariantes sur G . L'estimation de ses caractères et des techniques analogues à celles utilisées pour les groupes libres permettent d'établir l'inégalité maximale forte (p, p) et le théorème ergodique pour les moyennes sur les sphères de $SO^0(d, 1)$ dans le cas $d \geq 3$ et $p > d/(d-1)$.

Pour aider le lecteur, nous avons rassemblé dans plusieurs appendices des résultats techniques et des rappels de résultats variés utilisés au cours du texte.

CHAPITRE 1

PRINCIPES DE DÉMONSTRATION D'UN THÉORÈME ERGODIQUE

Afin de dégager certaines des idées présentes dans la démonstration d'un théorème ergodique pour une action de groupe, nous allons nous placer, dans ce premier chapitre, dans le cadre d'une action mesurable du groupe \mathbb{Z} . C'est, conjointement avec celui d'une action du groupe \mathbb{R} , le cadre le plus classique. On se donne un espace mesuré (X, \mathcal{B}, m) et une transformation bijective t sur X , mesurable ainsi que son inverse, préservant la mesure m . L'action de \mathbb{Z} sur X est alors l'application mesurable de $\mathbb{Z} \times X$ dans X donnée par

$$(k, x) \mapsto t^k(x).$$

La mesure m est positive et non nécessairement finie.

Les moyennes les plus étudiées dans ce contexte sont les moyennes de Birkhoff

$$A_n f = \frac{1}{n} \sum_{k=0}^{n-1} f \circ t^k$$

dont on cherche à connaître le comportement asymptotique pour une fonction f appartenant à un espace fonctionnel du type $L^p(X, m)$. Le problème se pose, plus généralement, pour toute transformation t de (X, \mathcal{B}, m) , non nécessairement inversible, qui préserve m . Nous allons rappeler la démonstration des deux théorèmes ergodiques fondamentaux : le théorème ergodique en moyenne dû à von Neumann (1932), donnant pour une fonction f de carré intégrable la convergence en moyenne de la suite $(A_n f)$, et le théorème ergodique ponctuel de Birkhoff (1931) qui, dans le cas d'une fonction f intégrable, assure la convergence presque partout.

Cela nous permettra de dégager plusieurs principes de démonstration qui seront utilisés dans les chapitres suivants pour des actions de groupes plus généraux. Nous verrons ainsi qu'un théorème ergodique en moyenne repose sur une étude spectrale des opérateurs A_n , tandis que la démonstration d'un théorème ergodique ponctuel dans un espace fonctionnel L nécessite la plupart du temps deux ingrédients essentiels :

- une inégalité maximale dans l'espace fonctionnel L , qui garantit que l'ensemble des fonctions pour lesquelles la convergence ponctuelle a lieu est un sous-espace fermé de L ;
- un sous-espace dense de fonctions pour lesquelles on connaît *a priori* la convergence presque partout.

La transformation t de (X, \mathcal{B}, m) préservant m induit sur chaque espace fonctionnel L raisonnable associé à (X, \mathcal{B}, m) un opérateur linéaire T , qui est l'opérateur de composition par t :

$$T : f \mapsto f \circ t.$$

Pour nous, L sera en général un espace normé sur lequel l'opérateur T est une contraction positive. Ces deux propriétés de T sont fondamentales, et suffisent à elles seules à l'obtention de théorèmes ergodiques. Il est en effet remarquable que ceux-ci s'étendent à des opérateurs T qui ne sont plus nécessairement issus d'une transformation t d'un espace (X, \mathcal{B}, m) . Plus précisément, le fait que l'opérateur T soit une contraction suffit à l'obtention de théorèmes ergodiques en moyenne, tandis que la positivité assure la validité des théorèmes ergodiques ponctuels. Nous verrons ainsi dans le prochain paragraphe le théorème de von Neumann, qui est un théorème ergodique en moyenne dans le cadre des contractions d'un espace de Hilbert, puis nous montrerons dans les paragraphes suivants le théorème ergodique ponctuel dans le cadre des contractions $L^1 - L^\infty$ positives.

1.1. Le théorème de von Neumann

On se donne un espace de Hilbert $(H, \langle \cdot, \cdot \rangle)$. Un opérateur *unitaire* T de H est un opérateur inversible dont l'adjoint T^* est égal à l'inverse T^{-1} . L'exemple fondamental pour nous est l'opérateur T de $L^2(X, m)$ induit par une transformation t inversible de (X, \mathcal{B}, m) lorsqu'elle préserve la mesure m . Signalons qu'une transformation t non nécessairement inversible mais préservant m induit un opérateur de $L^2(X, m)$ qui est toujours une *isométrie* de $L^2(X, m)$, mais qui n'est pas nécessairement unitaire.

On appelle *contraction* de H tout endomorphisme T de H tel que, pour tout $h \in H$, on ait $\|Th\| \leq \|h\|$. Les opérateurs unitaires et les isométries sont bien sûr des contractions de H .

Théorème 1.1.1 (von Neumann [vN32]). — Soient H un espace de Hilbert et T une contraction de H . On définit pour tout $n \geq 1$ l'opérateur de moyenne $A_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k$. Alors, pour tout $h \in H$,

$$A_n h \rightarrow E_1 h \quad \text{quand } n \rightarrow +\infty,$$

où E_1 désigne le projecteur orthogonal sur le sous-espace fermé $\text{Ker}(T - \text{Id})$ des vecteurs invariants par T .

La démonstration de ce théorème donnée en 1932 par von Neumann est basée sur la décomposition spectrale des opérateurs unitaires. Nous utiliserons ici un argument plus élémentaire proposé par F. Riesz quelques années plus tard et basé sur le lemme suivant.

Lemme 1.1.2. — *Soient H un espace de Hilbert et T une contraction de H . Alors on a la décomposition en somme directe orthogonale :*

$$H = \text{Ker}(T - Id) \oplus \overline{\text{Im}(T - Id)}.$$

Démonstration. — Comme T est une contraction, on a $|\langle Th, h \rangle| \leq \|h\|^2$ pour tout $h \in H$. Le cas d'égalité dans l'inégalité de Cauchy-Schwarz donne $Th = h$ si et seulement si $\langle Th, h \rangle = \|h\|^2$. On en déduit que

$$\text{Ker}(T - Id) = \text{Ker}(T^* - Id)$$

puisque T^* est aussi une contraction. La décomposition classique

$$H = \text{Ker}(T^* - Id) \oplus \overline{\text{Im}(T - Id)}$$

permet de conclure. □

Démonstration du théorème 1.1.1. — L'opérateur E_1 est l'opérateur de projection sur $\text{Ker}(T - Id)$, parallèlement à $\overline{\text{Im}(T - Id)}$.

Pour tout élément h de $\text{Ker}(T - Id)$ et tout $n \geq 1$, on a

$$\frac{1}{n} \sum_{k=0}^{n-1} T^k h = h = E_1 h,$$

tandis que pour tout $h \in \text{Im}(T - Id)$, il existe $g \in H$ tel que $h = Tg - g$ et l'égalité

$$\frac{1}{n} \sum_{k=0}^{n-1} T^k h = \frac{1}{n} (T^n g - g)$$

assure la convergence de $A_n h$ vers $0 = E_1 h$, puisque la suite $(T^n g)$ est bornée dans H . Les opérateurs de moyenne A_n étant des contractions, l'ensemble des $h \in H$ pour lesquels $A_n h \rightarrow 0$ est fermé. On a donc $A_n h \rightarrow 0$ pour tout $h \in \overline{\text{Im}(T - Id)}$. Par linéarité, on obtient le résultat annoncé. □

Examinons le cas particulier d'une contraction induite par une transformation ponctuelle. Soit T l'isométrie de $H = L^2(X, \mathcal{B}, m)$ associée à une transformation t de l'espace (X, \mathcal{B}, m) préservant la mesure. Notons \mathcal{I} la *tribu des invariants* de t , c'est-à-dire l'ensemble des éléments B de \mathcal{B} tels que $m(t^{-1}B \Delta B) = 0$. Le sous-espace $\text{Ker}(T - Id)$ s'identifie à l'espace $L^2(X, \mathcal{I}, m)$. Dans le cas où la mesure m est finie (et normalisée de façon à être une probabilité), le projecteur E_1 est l'espérance conditionnelle sachant \mathcal{I} . Plus généralement, on a le résultat suivant.

Théorème 1.1.3 (Théorème ergodique en moyenne)

Soient (X, \mathcal{B}, m) un espace probabilisé et t une transformation mesurable préservant la mesure de cet espace. Soit $p \in [1, +\infty[$. Pour tout $f \in L^p(X, \mathcal{B}, m)$, la suite

$$\left(\frac{1}{n} \sum_{k=0}^{n-1} f \circ t^k \right)_{n \geq 1}$$

converge dans l'espace normé $L^p(X, \mathcal{B}, m)$ vers l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$.

Démonstration. — Notons comme précédemment

$$A_n f = \frac{1}{n} \sum_{k=0}^{n-1} f \circ t^k.$$

Chaque A_n est une contraction de l'espace L^p .

Dans le cas $p = 2$, le théorème 1.1.3 est un cas particulier du théorème de von Neumann. On en déduit que, pour toute fonction mesurable bornée f sur X , la suite $(A_n f)$ converge en moyenne d'ordre 2, et donc en probabilité, vers $\mathbb{E}(f|\mathcal{I})$. La suite $(A_n f)$ étant uniformément bornée (par $\|f\|_\infty$), cette convergence en probabilité entraîne la convergence en moyenne d'ordre p .

Ainsi la suite $(A_n f)$ converge vers $\mathbb{E}(f|\mathcal{I})$ dans L^p pour tout f dans une partie dense de L^p . Les opérateurs A_n étant des contractions de L^p , cette convergence s'étend à tout l'espace. \square

Remarque 1.1.4. — La décomposition de H donnée au lemme 1.1.2 est invariante par les opérateurs A_n , et l'on a vu que, sur le sous-espace $\overline{\text{Im}(T - Id)}$, la suite (A_n) tend fortement vers 0. Si, plus généralement, on dispose d'une suite d'opérateurs (A_n) préservant une décomposition de $H = H_0 \oplus H_1$, où A_n est l'identité sur H_1 et tend fortement vers 0 sur H_0 , on aura bien sûr la convergence des $A_n h$ pour tout $h \in H$. Cette situation apparaît lorsque l'on considère une suite (F_n) de sous-ensembles finis et non vides de \mathbb{Z} , et les opérateurs de moyenne

$$A_n = \frac{1}{|F_n|} \sum_{k \in F_n} T^k,$$

où $|F_n|$ désigne le nombre d'éléments de F_n . Ces opérateurs sont des contractions qui préservent la décomposition associée à T par le lemme 1.1.2. En restriction à $H_1 = \text{Ker}(T - Id)$, ils valent l'identité, tandis que pour un élément $h = Tg - g$ de $\text{Im}(T - Id)$,

$$A_n h = \frac{1}{|F_n|} \left(\sum_{k \in 1+F_n} T^k g - \sum_{k \in F_n} T^k g \right).$$

On a convergence de $A_n h$ vers 0 si la suite (F_n) possède la propriété d'invariance asymptotique par translation

$$\frac{|(1 + F_n)\Delta F_n|}{|F_n|} \rightarrow 0.$$

Pour de telles suites (F_n) , on a donc encore convergence forte des opérateurs A_n vers le projecteur E_1 .

Cette propriété d'invariance asymptotique peut être introduite pour des suites (F_n) de parties dans tout groupe localement compact. On verra dans le chapitre suivant comment le théorème 1.1.3 s'étend aux moyennes sur de telles suites, dites suites de Følner.

Remarque 1.1.5. — On pourrait se demander s'il y a convergence en norme d'opérateurs de la suite (A_n) vers le projecteur E_1 . En général cela n'est pas le cas, comme le montre la suite $A_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k$ associée à la rotation r_α du cercle $S^1 = \mathbb{R}/\mathbb{Z}$, d'angle α irrationnel. En effet l'opérateur unitaire T sur $L^2(S^1, dx)$ défini par

$$Tf(x) = f(x + \alpha),$$

admet chaque fonction $f_m(x) = e^{2i\pi mx}$, $m \in \mathbb{Z}$, comme fonction propre pour la valeur propre $e^{2i\pi m\alpha}$. L'espace $H_1 = \text{Ker}(T - Id)$ des fonctions invariantes par T est réduit à la droite engendrée par la fonction constante $\mathbf{1}$ (ce qui signifie que la rotation d'angle α est ergodique). L'orthogonal de H_1 est le sous-espace $H_0 = L_0^2(S^1, dx)$ des fonctions d'intégrale nulle. Sur H_0 , l'opérateur A_n a pour spectre ponctuel

$$\text{spectre}_p(A_n|_{H_0}) = \left\{ \frac{1}{n} \frac{e^{2i\pi n m \alpha} - 1}{e^{2i\pi m \alpha} - 1} ; m \in \mathbb{Z}^* \right\}.$$

Or, par densité de la suite $(m\alpha \pmod{1})_m$ dans l'intervalle $[0, 1]$, on a

$$\sup_{m \neq 0} \frac{1}{n} \frac{|e^{2i\pi n m \alpha} - 1|}{|e^{2i\pi m \alpha} - 1|} = \sup_{x \in]0, 1[} \frac{1}{n} \frac{\sin(\pi n x)}{\sin(\pi x)} = 1.$$

On a donc, pour tout $n \geq 1$,

$$\|A_n - E_1\| = \|A_n|_{H_0}\| = 1.$$

Le phénomène que nous venons de décrire pour la rotation irrationnelle est tout à fait général pour les actions de \mathbb{Z} . Soit t une transformation préservant la mesure d'un espace probabilisé (X, \mathcal{B}, m) . En dehors du cas trivial où il existe un entier ℓ tel que $t^\ell(x) = x$ pour m -presque tout x , on a toujours $\|A_n - E_1\| = 1$. En effet, pour tout entier p , d'après le lemme de Kakutani-Rohlin (voir par exemple [Hal51, Lemma 2, page 71]), il existe alors $B \in \mathcal{B}$, avec $m(B) > 0$, tel que les ensembles $t^k(B)$, $1 \leq k \leq 2p$, soient deux à deux disjoints. Notons f_p la fonction nulle en dehors de $\cup_{k=1}^{2p} t^k(B)$, valant 1 sur $t^k(B)$ pour $1 \leq k \leq p$ et -1 sur $t^k(B)$ pour $p < k \leq 2p$. On a $E_1(f_p) = 0$ et $\lim_{p \rightarrow +\infty} \|A_n(f_p)\| / \|f_p\| = 1$ pour tout n . Nous reviendrons sur ce sujet dans la remarque 2.2.3 du chapitre suivant.

1.2. Inégalités maximales et principe de Banach

On va s'intéresser maintenant à la convergence ponctuelle de la suite des moyennes $A_n f$, lorsque f varie dans un espace fonctionnel convenable. Nous avons vu, dans le cadre d'une convergence L^2 , qu'il était utile de savoir *a priori* que l'ensemble des h pour lesquels la suite $(A_n h)$ converge (dans L^2) est fermé. Les *inégalités maximales* jouent un rôle analogue pour la convergence ponctuelle : elles donnent un critère sur une suite d'opérateurs (A_n) d'un espace fonctionnel L pour que l'ensemble des $f \in L$ tels que $A_n f$ converge *presque partout* soit un ensemble fermé.

Nous nous plaçons dans le cadre général d'applications linéaires entre un espace de Banach E et un espace L de fonctions mesurables sur un espace mesuré (X, \mathcal{B}, m) . Rappelons qu'un opérateur $A : E \rightarrow L$ est dit *continu en mesure* si pour toute suite $(f_k)_{k \in \mathbb{N}}$ d'éléments de E convergeant vers f et pour tout $a > 0$,

$$m(\{|A f_k - A f| > a\}) \rightarrow 0.$$

Théorème 1.2.1 (Principe de Banach). — Soient $(E, \|\cdot\|)$ un espace de Banach et L un espace vectoriel de fonctions mesurables définies sur un espace mesuré (X, \mathcal{B}, m) . Soit $(A_n)_{n \geq 1}$ une suite d'applications linéaires de E dans L . On définit l'opérateur maximal

$$A^* : f \in E \mapsto A^* f = \sup_{n \geq 1} |A_n f|.$$

- (i) S'il existe une fonction $c : \mathbb{R}_+^* \rightarrow \mathbb{R}_+$ tendant vers 0 à l'infini telle que, pour tout $f \in E$ et tout $\alpha > 0$,

$$(1.2.1) \quad m(\{A^* f > \alpha \|f\|\}) \leq c(\alpha),$$

alors A_n est continu en mesure pour tout $n \geq 1$ et l'ensemble des $f \in E$ pour lesquels la suite $(A_n f)_{n \geq 1}$ converge m -presque partout est fermé dans E .

- (ii) Réciproquement, si chaque opérateur A_n est continu en mesure et si $A^* f$ est fini m -presque partout pour tout $f \in E$ (en particulier si $(A_n f)$ converge presque partout), alors il existe une fonction $c : \mathbb{R}_+^* \rightarrow \mathbb{R}_+$ décroissante et tendant vers 0 à l'infini pour laquelle l'inégalité maximale (1.2.1) est satisfaite.

Nous ne donnons pas ici la démonstration du point (ii) de ce théorème, qui affirme la nécessité de l'inégalité maximale, et que nous n'utiliserons pas dans la suite. Nous renvoyons pour cela le lecteur à [Gar70, §1.1] ou à [Kre85, §1.7].

Démonstration de (i). — La continuité en mesure de chaque opérateur A_n découle simplement de l'inégalité (1.2.1). En effet, si la suite (f_k) converge dans E vers f on a, pour tout $a > 0$,

$$m(\{|A_n f_k - A_n f| > a\}) \leq m(\{A^*(f_k - f) > a\}) \leq c\left(\frac{a}{\|f_k - f\|}\right),$$

et par conséquent $m(\{|A_n f_k - A_n f| > a\})$ tend vers 0 lorsque k tend vers l'infini.

Notons maintenant \mathcal{C} l'ensemble des $f \in E$ tels que $A_n f$ converge m -presque partout et posons, pour $f \in E$ et $x \in X$,

$$\Delta f(x) = \lim_{N \rightarrow +\infty} \sup_{n, m \geq N} |A_n f(x) - A_m f(x)|.$$

Observons que si $\Delta f(x) = 0$, la suite $(A_n f(x))_n$ est une suite de Cauchy dans \mathbb{R} et donc convergente. Par conséquent, un élément f appartient à \mathcal{C} si et seulement si $\Delta f = 0$ m -presque partout. Montrons que \mathcal{C} est un sous-espace fermé de E . Soit $(f_k)_{k \in \mathbb{N}}$ une suite d'éléments de \mathcal{C} , que l'on suppose converger vers $f \in E$. En particulier, $\Delta f_k = 0$ m -presque partout pour tout $k \geq 0$. Nous avons alors

$$\Delta f = \Delta f - \Delta f_k \leq \Delta(f - f_k) \leq 2A^*(f - f_k),$$

et donc pour tout $a > 0$,

$$m(\{\Delta f > a\}) \leq m(\{2A^*(f - f_k) > a\}) \leq c \left(\frac{a}{2\|f - f_k\|} \right).$$

Par suite, $m(\{\Delta f > a\}) = 0$ pour tout $a > 0$ et donc $\Delta f = 0$ m -presque partout. Ceci montre que $f \in \mathcal{C}$ et conclut la démonstration de la première partie du principe de Banach. \square

Ce furent Hardy et Littlewood qui, en 1930, donnèrent le premier exemple d'inégalité maximale pour la famille des opérateurs de moyenne sur des intervalles de \mathbb{R} . On verra plus généralement dans le chapitre 3 que l'opérateur maximal A^* associé aux moyennes sur les boules de \mathbb{R}^d est *de type faible* $(1, 1)$, c'est-à-dire qu'il existe une constante $C > 0$ telle que pour toute fonction $f \in L^1(\mathbb{R}^d, \lambda)$,

$$\lambda(\{A^* f > \alpha \|f\|_1\}) \leq \frac{C}{\alpha}.$$

L'inégalité maximale (1.2.1) est donc vérifiée avec $c(\alpha) = C/\alpha$.

L'opérateur maximal associé à une suite (A_n) d'opérateurs linéaires définis sur $L^p(X, m)$ est un exemple d'opérateur *sous-additif*, c'est-à-dire tel que pour tous f_1, f_2 ,

$$|A(f_1 + f_2)(x)| \leq |A f_1(x)| + |A f_2(x)| \quad m\text{-pp.}$$

Un tel opérateur défini sur $L^p(X, m)$ sera dit de *type faible* (p, p) s'il existe une constante C telle que, pour tout $f \in L^p(X, m)$,

$$(1.2.2) \quad m(\{|A f| > \alpha \|f\|_p\}) \leq \frac{C^p}{\alpha^p},$$

et de *type fort* (p, p) (ou *borné dans* L^p) s'il existe $C > 0$ telle que

$$(1.2.3) \quad \|A f\|_p \leq C \|f\|_p.$$

On parle encore d'*inégalité faible* pour (1.2.2) et d'*inégalité forte* pour (1.2.3).

L'inégalité faible dans L^p apparaît comme un substitut lorsque A n'est pas borné dans L^p . Cependant, par interpolation, en présence d'une inégalité dans L^∞ , elle implique que A est borné dans L^q pour tout $q > p$, comme l'indique la proposition

suivante. C'est un cas particulier du théorème d'interpolation de Marcinkiewicz (voir par exemple [SW71, Theorem V.2.4]). Nous en donnons une démonstration directe.

Proposition 1.2.2. — Soient (X, \mathcal{B}) un espace mesurable muni d'une mesure σ -finie m , et $p \in [1, +\infty[$. Soit A un opérateur sous-additif de $L^p(X, m) + L^\infty(X, m)$ dans $L^p(X, m) + L^\infty(X, m)$. On suppose que la restriction de A à $L^p(X, m)$ satisfait à l'inégalité (1.2.2) et qu'il existe $C' > 0$ tel que $\|Af\|_\infty \leq C'\|f\|_\infty$ pour tout $f \in L^\infty(X, m)$. Alors A est un opérateur de type fort (q, q) pour tout $q > p$.

Démonstration. — Soient $q > p$, $f \in L^q(X, m)$ et $\alpha > 0$. Alors

$$f = f\mathbf{1}_{\{|f|>\alpha/2C'\}} + f\mathbf{1}_{\{|f|\leq\alpha/2C'\}}$$

est une décomposition de f comme somme d'un élément de $L^p(X, m)$ et d'un élément de $L^\infty(X, m)$. En notant g la fonction $g = f\mathbf{1}_{\{|f|>\alpha/2C'\}}$, on obtient $|Af| \leq |Ag| + \alpha/2$ et donc

$$m(\{|Af| > \alpha\}) \leq m(\{2|Ag| > \alpha\}) \leq \frac{(2C')^p}{\alpha^p} \int_{\{|f|>\alpha/2C'\}} |f|^p(x) dm(x).$$

Rappelons que, par application du théorème de Fubini, on a, pour toute fonction mesurable positive F ,

$$\int_X F(x)^q dm(x) = \int_X \int_0^{F(x)} q\alpha^{q-1} d\alpha dm(x) = q \int_0^{+\infty} m(\{F > \alpha\}) \alpha^{q-1} d\alpha.$$

On en déduit que

$$\begin{aligned} \|Af\|_q^q &= q \int_0^{+\infty} m(\{|Af| > \alpha\}) \alpha^{q-1} d\alpha \\ &\leq (2C')^p q \int_0^{+\infty} \alpha^{q-1-p} \left(\int_{\{\alpha < 2C'|f\}} |f|^p(x) dm(x) \right) d\alpha \\ &\leq \frac{q 2^q C^p C'^{q-p}}{q-p} \int_X |f|^q(x) dm(x), \end{aligned}$$

en appliquant à nouveau le théorème de Fubini. C'est l'inégalité forte dans $L^q(X, m)$. \square

Dans la pratique, il est souvent plus facile de démontrer des inégalités maximales pour des fonctions ayant des propriétés supplémentaires (fonctions bornées, continues,...). La proposition suivante montre qu'il suffit de vérifier l'inégalité maximale (1.2.1) sur un ensemble dense \mathcal{D} .

Proposition 1.2.3. — Soient, comme dans le théorème 1.2.1, (A_n) une suite d'opérateurs d'un espace de Banach $(E, \|\cdot\|)$ dans un espace fonctionnel $L(X, \mathcal{B}, m)$, et A^* l'opérateur maximal. Soit $\alpha \mapsto c(\alpha)$ une fonction continue à gauche.

Si les opérateurs A_n sont continus en mesure et s'il existe un sous-espace dense \mathcal{D} de E tel que, pour tout $f \in \mathcal{D}$ et tout $\alpha > 0$,

$$(1.2.1) \quad m(\{A^*f > \alpha\|f\|\}) \leq c(\alpha),$$

alors (1.2.1) est vérifié par tout élément de E .

Si $E = L = L^p(X, m)$ avec $1 \leq p < \infty$, et si les opérateurs A_n sont continus, l'inégalité maximale

$$\|A^*f\|_p \leq C\|f\|_p$$

est vraie pour tout $f \in E$ si elle est vraie pour tout f appartenant à un sous-espace dense.

Démonstration. — Soit $f \in E$, $f \neq 0$. Pour démontrer que (1.2.1) est vrai, il suffit de démontrer que pour tout $N \geq 1$,

$$(1.2.4) \quad m\left(\left\{\max_{1 \leq n \leq N} |A_n f| > \alpha\|f\|\right\}\right) \leq c(\alpha).$$

Soit (f_k) une suite d'éléments de \mathcal{D} convergeant vers f . Donnons-nous $\varepsilon \in]0, \alpha[$. Grâce à la continuité en mesure des opérateurs A_n , $n = 1, \dots, N$, on voit qu'il existe k_0 tel que, pour $k \geq k_0$ et pour $n = 1, \dots, N$, on ait

$$m(\{|A_n f_k - A_n f| \geq \varepsilon\|f\|\}) \leq \frac{\varepsilon}{N}.$$

Posons $\Omega_{n,k} = \{|A_n f_k - A_n f| < \varepsilon\|f\|\}$ et

$$\Omega_k = \left\{\max_{1 \leq n \leq N} |A_n f| - \max_{1 \leq n \leq N} |A_n f_k| < \varepsilon\|f\|\right\}.$$

On vérifie aisément que $\bigcap_{n=1}^N \Omega_{n,k} \subset \Omega_k$, ce qui donne, par passage aux complémentaires,

$$m\left(\left\{\max_{1 \leq n \leq N} |A_n f| - \max_{1 \leq n \leq N} |A_n f_k| \geq \varepsilon\|f\|\right\}\right) \leq \varepsilon,$$

si $k \geq k_0$. On peut aussi supposer que $\|f_k\| \leq (1 + \varepsilon)\|f\|$ pour $k \geq k_0$, d'où dans ce cas,

$$\begin{aligned} m\left(\left\{\max_{1 \leq n \leq N} |A_n f| > \alpha\|f\|\right\}\right) &\leq \varepsilon + m\left(\left\{\max_{1 \leq n \leq N} |A_n f_k| > \frac{\alpha - \varepsilon}{1 + \varepsilon}\|f_k\|\right\}\right) \\ &\leq \varepsilon + c\left(\frac{\alpha - \varepsilon}{1 + \varepsilon}\right). \end{aligned}$$

L'inégalité (1.2.4) en résulte puisque ε est arbitraire.

La démonstration dans le cas d'inégalités fortes est laissée au lecteur. \square

1.3. Les théorèmes ergodiques de Birkhoff et de Hopf-Dunford-Schwartz

Dans cette section, nous allons étudier la convergence ponctuelle de la suite $(A_n f)$, où $A_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k$ est l'opérateur de moyenne associé à une contraction $L^1 - L^\infty$ sur un espace mesuré (X, \mathcal{B}, m) et f une fonction intégrable.

Commençons par préciser nos définitions. Soit (X, \mathcal{B}, m) un espace mesuré, où m est une mesure σ -finie. Une contraction $L^1 - L^\infty$ est un endomorphisme T de l'espace vectoriel $L^1(X, m)$ tel que, pour tout $f \in L^1(X, m)$, on ait $\|Tf\|_1 \leq \|f\|_1$ et, pour tout $f \in L^1(X, m) \cap L^\infty(X, m)$, on ait $\|Tf\|_\infty \leq \|f\|_\infty$. Par ailleurs, un endomorphisme de l'espace $L^1(X, m)$ est dit *positif* s'il préserve le cône des fonctions positives sur X .

L'opérateur de composition par une transformation t de X préservant la mesure m est une contraction $L^1 - L^\infty$ positive. Le théorème ergodique de Birkhoff assure la convergence ponctuelle de la suite $(A_n f)$ lorsque l'opérateur T est associé à une telle transformation et la mesure m est une mesure de probabilité. Le théorème de Birkhoff a été étendu par Hopf au cas d'un *opérateur de Markov* P , c'est-à-dire d'une contraction positive de $L^1(X, m)$ satisfaisant à $P(\mathbf{1}) = \mathbf{1}$. Il est clair qu'un opérateur de composition sur un espace probabilisé est un opérateur de Markov, et que tout opérateur de Markov est une contraction $L^1 - L^\infty$. Nous allons donner une version plus générale du théorème ergodique ponctuel, celle de Dunford et Schwartz, qui concerne toutes les contractions $L^1 - L^\infty$ positives.

Théorème 1.3.1 (Birkhoff, Hopf, Dunford-Schwartz)

Soit T une contraction $L^1 - L^\infty$ positive sur l'espace mesuré (X, \mathcal{B}, m) , où m est une mesure finie. Notons $A_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k$.

Pour tout $f \in L^1(X, m)$, la suite $(A_n f)$ converge m -presque partout vers une fonction limite \tilde{f} qui est T -invariante.

Démonstration. — Nous établissons d'abord que l'ensemble des fonctions f appartenant à $L^1(X, m)$ pour lesquelles la suite $(A_n f)$ converge ponctuellement est dense dans $L^1(X, m)$. Pour cela, nous utilisons la propriété de contraction dans L^2 de T . Ce fait résulte de la version ci-dessous du théorème de convexité de Riesz. Nous renvoyons à [Gar70, Proposition 2.1.4] pour une démonstration directe de cette version. Pour le cas général, on pourra consulter [SW71, Theorem V.1.3].

Proposition 1.3.2. — Soit T une contraction $L^1 - L^\infty$ positive de l'espace mesuré (X, \mathcal{B}, m) , où m est une mesure σ -finie. Pour tout $p \in [1, +\infty[$, la restriction de T à $L^p(X, m) \cap L^1(X, m)$ admet un prolongement continu en une contraction de $L^p(X, m)$.

La propriété de contraction dans L^2 nous permet d'appliquer le lemme 1.1.2. On a ainsi

$$(1.3.1) \quad L^2(X, m) = \text{Ker}(T - Id) \oplus \overline{\text{Im}(T - Id)}.$$

Nous utilisons maintenant la densité de $L^\infty(X, m)$ dans $L^2(X, m)$, qui nécessite l'hypothèse de finitude de la mesure m dans l'énoncé du théorème 1.3.1. La décomposition (1.3.1) nous garantit alors que le sous-espace $\text{Ker}(T - Id) + (T - Id)L^\infty(X, m)$ est dense dans $L^2(X, m)$ et donc dans $L^1(X, m)$. Or, pour toute fonction f dans le sous-espace $\text{Ker}(T - Id) + (T - Id)L^\infty(X, m)$, la suite $(A_n f)$ converge m -presque partout. En effet,

- si $f \in \text{Ker}(T - Id)$, alors $A_n f = f$ converge presque partout vers f ;
- si $f = Tg - g$ avec $g \in L^\infty(X, m)$, alors $A_n f = (T^n g - g)/n$ converge presque partout vers 0.

Pour en déduire que l'ensemble des fonctions de $L^1(X, m)$ pour lesquelles on a convergence ponctuelle coïncide avec tout $L^1(X, m)$, il nous suffit de savoir que cet ensemble est fermé et donc, d'après le principe de Banach, que la suite des opérateurs de moyenne (A_n) vérifie l'inégalité maximale faible dans $L^1(X, m)$. Ceci sera établi dans la proposition 1.3.4 et terminera la démonstration du théorème 1.3.1. \square

Comme nous allons le voir, cette inégalité maximale est vraie si la mesure est seulement supposée σ -finie. Elle est démontrée à l'aide du théorème suivant, connu sous le nom de *théorème ergodique maximal*. Il est dû à Yosida et Kakutani pour les contractions issues de transformations préservant une mesure et à Hopf pour les contractions positives de L^1 . La démonstration très ingénieuse que nous donnons ici est due à Garsia.

Remarquons que si T est une contraction positive de $L^1(X, m)$, on peut étendre son domaine à toutes les fonctions mesurables positives à condition d'accepter les valeurs infinies : si f est une fonction mesurable positive, on pose $Tf = \lim T f_n$ pour une suite croissante (f_n) de fonctions intégrables positives convergeant vers f . Cette définition est correcte car, comme on peut le vérifier, le résultat obtenu ne dépend pas de la suite choisie. On peut ensuite étendre le domaine de T à toutes les fonctions dont la partie positive $f^+ = \max(f, 0)$ est intégrable, en posant

$$Tf = T(f^+) - T(f^-)$$

où $f^- = (-f)^+$. La fonction Tf prend alors ses valeurs dans $[-\infty, +\infty[$ et vérifie $(Tf)^+ \leq T(f^+)$.

Théorème 1.3.3 (Théorème ergodique maximal). — Soient (X, \mathcal{B}, m) un espace mesuré, où m est une mesure σ -finie, et T une contraction positive de $L^1(X, m)$. Notons $A_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k$.

Soit $f : X \rightarrow \mathbb{R}$ une fonction mesurable de partie positive $f^+ = \max(f, 0)$ intégrable. On note E_n l'ensemble $\{\max_{1 \leq k \leq n} A_k f \geq 0\}$. Alors, pour tout $n \geq 1$,

$$\int_{E_n} f \, dm \geq 0.$$

Démonstration. — Posons $S_k f = \sum_{i=0}^{k-1} T^i f = k A_k f$ et $M_n f = \max_{1 \leq k \leq n} S_k f$. Alors, $(M_n f)^+$ est une fonction intégrable et, pour tout $k = 1, \dots, n$,

$$(M_n f)^+ \geq S_k f,$$

d'où

$$T((M_n f)^+) \geq T(S_k f).$$

Puisque $f + T(S_k f) = S_{k+1} f$, on obtient $f \geq S_{k+1} f - T((M_n f)^+)$ pour tout $k = 0, \dots, n$ (le cas $k = 0$ est trivial), ce qui implique que

$$f \geq M_n f - T((M_n f)^+)$$

et

$$\int_{E_n} f \, dm \geq \int_{E_n} M_n f \, dm - \int_{E_n} T((M_n f)^+) \, dm.$$

Or, d'une part, comme $E_n = \{M_n f \geq 0\}$, on a

$$\int_{E_n} M_n f \, dm = \int_X (M_n f)^+ \, dm = \|(M_n f)^+\|_1.$$

D'autre part,

$$\begin{aligned} \int_{E_n} T((M_n f)^+) \, dm &\leq \int_X T((M_n f)^+) \, dm \\ &= \|T((M_n f)^+)\|_1 \leq \|(M_n f)^+\|_1. \end{aligned}$$

puisque T est une contraction positive de L^1 . Le résultat annoncé en découle. \square

Déduisons à présent que, si l'on suppose de plus que T est une contraction L^∞ , la suite (A_n) vérifie l'inégalité maximale faible dans L^1 .

Proposition 1.3.4 (Inégalité maximale faible L^1). — Soit (X, \mathcal{B}, m) un espace mesuré, où m est une mesure σ -finie. Soit T une contraction L^1 - L^∞ positive. Notons $A_n = \frac{1}{n} \sum_{k=0}^{n-1} T^k$ et A^* l'opérateur maximal défini par $A^* f(x) = \sup_{n \geq 1} |A_n f(x)|$.

Pour toute fonction f mesurable dont la partie positive $f^+ = \max(f, 0)$ est intégrable et pour tous $n \geq 1$, $\alpha > 0$, on a

$$\alpha m \left(\left\{ \max_{1 \leq k \leq n} A_k f \geq \alpha \right\} \right) \leq \int_{\{\max_{1 \leq k \leq n} A_k f \geq \alpha\}} f \, dm.$$

De plus, pour tout $f \in L^1(X, m)$, on a

$$\alpha m(\{A^* f > \alpha\}) \leq \|f\|_1.$$

Démonstration. — La première inégalité se montre en appliquant le théorème maximal à la fonction $g = f - \alpha$, dont la partie positive g^+ est intégrable. On obtient

$$\int_{\{\max_{1 \leq k \leq n} A_k g \geq 0\}} (f - \alpha) \, dm \geq 0.$$

Dans le cas où $T\mathbf{1} = \mathbf{1}$, par exemple si T provient d'une transformation de X ou si T est markovien, on a immédiatement l'inégalité souhaitée puisque $A_n g = A_n f - \alpha$.

Dans le cas général d'une contraction L^∞ positive, on a seulement $T\mathbf{1} \leq \mathbf{1}$, et donc $A_n g \geq A_n f - \alpha$, ce qui donne l'inclusion

$$\left\{ \max_{1 \leq k \leq n} A_k f \geq \alpha \right\} \subset \left\{ \max_{1 \leq k \leq n} A_k g \geq 0 \right\}.$$

Notons $E_n \subset F_n$ ces deux sous-ensembles. On écrit alors

$$\int_{E_n} (f - \alpha) \, dm = \int_{F_n} (f - \alpha) \, dm - \int_{F_n \setminus E_n} (f - \alpha) \, dm$$

et le fait que $f - \alpha$ soit négative sur $E_n^c = \{\max_{1 \leq k \leq n} A_k f < \alpha\}$ permet de conclure.

Pour la seconde inégalité, notons que les opérateurs A_n sont positifs et donc que $A^* f \leq A^* |f|$. Par conséquent

$$\{A^* f > \alpha\} \subset \{A^* |f| > \alpha\} = \bigcup_n \left\{ \max_{1 \leq k \leq n} A_k |f| > \alpha \right\}.$$

La première inégalité appliquée à la fonction $|f|$ donne la majoration voulue. \square

Proposition 1.3.5. — *Gardons les hypothèses et notations du théorème 1.3.1. Pour tout $p \in [1, +\infty[$ et tout $f \in L^p(X, m)$, la suite $(A_n f)$ converge dans L^p vers \tilde{f} .*

Démonstration. — La convergence dans L^p de la suite $(A_n f)$ se déduit immédiatement de la convergence dans L^2 donnée par le théorème 1.1.1. Pour cela nous réutilisons un argument déjà vu dans la démonstration du théorème 1.1.3. L'espace $L^\infty(X, m)$ étant contenu dans l'espace $L^2(X, m)$, le théorème de von Neumann assure que, pour toute fonction f bornée, la suite $(A_n f)$ converge en moyenne quadratique, et donc en mesure. Puisque la suite $(A_n f)$ est uniformément bornée, cette convergence est également vraie au sens de la norme de $L^p(X, m)$. Mais la propriété de contraction entraîne que l'ensemble des f tels que la suite $(A_n f)$ converge dans $L^p(X, m)$ est fermé dans $L^p(X, m)$. Comme l'espace $L^\infty(X, m)$ est dense dans l'espace $L^p(X, m)$, on peut conclure que, pour tout $f \in L^p(X, m)$, la suite $(A_n f)$ converge dans $L^p(X, m)$. \square

En revanche, comme nous le verrons dans la remarque 1.3.7, si $f \in L^\infty(X, m)$, la suite $(A_n f)$ ne converge pas nécessairement au sens de la norme dans $L^\infty(X, m)$.

Remarque 1.3.6. — Dans le cas d'un opérateur T issu d'une transformation t préservant la mesure d'un espace probabilisé (X, \mathcal{B}, m) , la limite ponctuelle \tilde{f} de l'énoncé du théorème de Birkhoff coïncide avec l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$ de f sachant la tribu des invariants \mathcal{I} . Dans ce cas, on peut d'ailleurs démontrer directement ce théorème, c'est-à-dire la convergence ponctuelle des moyennes $A_n f = \frac{1}{n} \sum_{k=0}^{n-1} f \circ t^k$ vers $\mathbb{E}(f|\mathcal{I})$ pour une fonction f intégrable, sans utiliser le principe de Banach ni l'existence d'un ensemble dense, mais en se basant uniquement sur la première inégalité de la

proposition 1.3.4. En effet, en l'appliquant à la fonction $f\mathbf{1}_B$, où B est une partie mesurable de X invariante par t de sorte que $A_n(\mathbf{1}_B f) = \mathbf{1}_B A_n f$ et en passant à la limite en n , on obtient

$$\alpha m(B \cap \{\sup_n A_n f \geq \alpha\}) \leq \int_{B \cap \{\sup_n A_n f \geq \alpha\}} f \, dm.$$

Soit B l'ensemble

$$B = \left\{ \limsup_n A_n f \geq \alpha \right\}.$$

C'est un ensemble invariant qui est inclus dans $\{\sup_n A_n f \geq \alpha\}$ et l'inégalité précédente donne donc

$$\alpha m(B) \leq \int_B f \, dm.$$

Comme on peut supposer, quitte à la retrancher à f , que l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$ est nulle, on a

$$\int_B f \, dm = \int_B \mathbb{E}(f|\mathcal{I}) \, dm = 0$$

et donc $m(B) = 0$. Ceci étant vrai pour tout $\alpha > 0$, on en déduit que $\limsup_n A_n f \leq 0$ m -presque partout. Cette inégalité est aussi vraie pour la fonction $-f$ et on trouve finalement, pour toute fonction $f \in L^1(X, m)$ dont l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$ est nulle,

$$\lim_n A_n f = 0 = \mathbb{E}(f|\mathcal{I}) \quad m\text{-p.p.}$$

Remarque 1.3.7. — Considérons toujours le cas le cas où T provient d'une transformation t préservant la mesure d'un espace probabilisé (X, \mathcal{B}, m) . On dit que t est *ergodique* si pour tout élément A de la tribu \mathcal{I} des invariants de t , on a $m(A) = 0$ ou 1. Cela se traduit de façon équivalente par le fait que $L^1(X, m)$ (ou $L^p(X, m)$, $p \in [1, +\infty)$) ne contient pas de fonctions invariantes par T autres que les fonctions constantes. Dans ce cas on a donc $\mathbb{E}(f|\mathcal{I}) = \int_X f \, dm$.

Par exemple, la rotation irrationnelle r_α du cercle $S^1 = \mathbb{R}/\mathbb{Z}$ (voir la remarque 1.1.5) est une transformation ergodique. En effet, on voit facilement, par transformation de Fourier, que seules les fonctions constantes sont invariantes par r_α . Le théorème de Birkhoff affirme donc que, pour toute fonction f intégrable sur S^1 et pour presque tout x relativement à la mesure de Lebesgue sur S^1 , on a

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(x + k\alpha) = \int_{S^1} f(y) \, dy.$$

Lorsque f est une fonction continue, les moyennes de Birkhoff de f ont un sens pour tout x de S^1 , et on peut se demander si la convergence a lieu pour *tout* $x \in S^1$. C'est effectivement le cas, comme l'affirme le théorème d'équidistribution de Weyl. La convergence est même uniforme sur S^1 (voir [Rau76]). On en déduit immédiatement que la mesure de Lebesgue est la seule mesure sur S^1 invariante par r_α .

Notons qu'il existe des fonctions $f \in L^\infty(S^1, dx)$ pour lesquelles la convergence des moyennes de Birkhoff n'est pas réalisée en norme L^∞ . Il suffit par exemple de considérer la fonction caractéristique d'un sous-ensemble borélien A de S^1 , de mesure strictement positive, tel que, pour tout $n > 0$, on ait

$$\frac{1}{n} \sum_{k=0}^{n-1} \mathbf{1}_A(x + k\alpha) = 0$$

pour un ensemble de points x de mesure > 0 . En utilisant une notation d'intervalles pour représenter les arcs de cercle sur $S^1 = [0, 1[$, on peut par exemple choisir l'ensemble A comme le complémentaire dans le cercle de

$$\bigcup_{n \geq 0} [n\alpha, n\alpha + 2^{-n-2}[\pmod{1}.$$

Remarque 1.3.8. — Considérons maintenant le cas général d'un homéomorphisme t d'un espace compact X . L'opérateur associé sur l'espace $C(X)$ des fonctions continues sur X donne, par dualité, une transformation continue affine du convexe $*$ -faiblement compact formé par les mesures de probabilité sur X . En utilisant le théorème du point fixe de Markov-Kakutani [Rud91, Theorem 5.23], nous obtenons l'existence d'une mesure de probabilité m invariante par t . L'homéomorphisme t est dit *uniquement ergodique* en cas d'unicité de la mesure de probabilité invariante. Cette condition entraîne l'ergodicité. Supposons en effet qu'il existe un ensemble borélien A , invariant par t et tel que $0 < m(A) < 1$. Alors la mesure m_A , définie par $m_A(B) = m(B \cap A)/m(A)$ pour tout $B \in \mathcal{B}$, est invariante par t et distincte de m , ce qui contredit l'unique ergodicité.

Un résultat classique énonce que t est uniquement ergodique si et seulement si pour toute fonction continue f sur X , la suite des moyennes $A_n f = \frac{1}{n} \sum_{k=0}^{n-1} f \circ t^k$ converge ponctuellement vers $\int_X f dm$. De plus cette propriété est encore équivalente à la convergence uniforme sur X de $(A_n f)$ pour tout $f \in C(X)$. (Les références sont données dans la remarque 2.2.9.)

La convergence dans l'espace des fonctions continues d'opérateurs de moyenne A_n est un phénomène rare, mais nous en rencontrerons d'autres exemples, comme dans le chapitre 2 où le théorème de Weyl est étendu au cas d'une action par isométries d'un groupe moyennable (corollaire 2.2.8), ou encore dans le chapitre 6 pour certaines actions d'un groupe libre (théorèmes 6.1.2 et 6.2.3).

Remarque 1.3.9. — L'action par translation d'un groupe G sur lui-même fournit un cadre naturel pour lequel il est nécessaire de montrer des théorèmes ergodiques et des inégalités maximales en mesure infinie. Le théorème de Dunford-Schwartz, que nous avons ici montré en mesure finie, reste en partie vrai en mesure infinie :

si $p \in]1, +\infty[$, il y a convergence ponctuelle et dans L^p de la suite $(A_n f)$ pour une fonction $f \in L^p(X, m)$.

Si $p = 1$ on a encore convergence ponctuelle mais cette convergence n'a en général pas lieu dans L^1 . Pour voir ce dernier point, il suffit de considérer la transformation $T : k \mapsto k + 1$ sur $X = \mathbb{Z}$ muni de la mesure de comptage et la fonction $f = \delta_1$. Dans ce cas $A_n f = \frac{1}{n} \sum_{k=0}^{n-1} \delta_{1-k}$. Cette suite tend vers 0 ponctuellement et dans L^p si $p > 1$ mais ne converge pas dans L^1 .

Si $p = +\infty$, on n'a pas en général convergence ponctuelle, lorsque la mesure est infinie. Dans l'exemple précédent, il suffit de considérer une fonction f bornée sur \mathbb{Z} , qui ne converge pas au sens de Cesàro, c'est-à-dire telle que la suite $(A_n f(0))$ ne converge pas.

Notice bibliographique. — Les idées de base exposées dans ce chapitre sont initialement dues aux travaux de Banach [**Ban26**] en 1926, Birkhoff [**Bir31**] en 1931 et von Neumann [**vN32**] en 1932.

Ces articles sont à l'origine de nombreuses études qui ont progressivement élargi le champ des recherches en théorie ergodique. Parmi les nombreuses améliorations très rapidement apportées au théorème de Birkhoff figurent notamment les résultats de Khintchine [**Khi33**] et de Yosida et Kakutani [**YK39**]. L'article de Wiener [**Wie39**], dont nous reparlerons dans la suite, illustre brillamment l'utilisation des inégalités maximales.

Nous avons déjà mentionné dans ce chapitre les contributions de Hopf, qui a généralisé dans [**Hop54**] le travail de Birkhoff, concernant le cas d'une transformation préservant une mesure finie, au cas d'un opérateur de Markov, puis de Dunford et Schwartz, qui ont étendu dans [**DS56**] le résultat de Hopf pour obtenir l'énoncé du théorème 1.3.1. En fait leur résultat est même plus général : en supposant seulement que T est une contraction $L^1 - L^\infty$, mais sans les hypothèses de positivité pour T et de finitude pour m , l'énoncé reste vrai, sauf la convergence dans L^1 . Pour un exposé complet de la démonstration du théorème de Dunford et Schwartz nous renvoyons au livre de Garsia [**Gar70**].

Pour de nombreuses extensions de théorèmes ergodiques, notamment au cas d'opérateurs agissant sur d'autres espaces de Banach que les espaces L^p , on pourra consulter le livre de Krengel [**Kre85**]. Le lecteur y trouvera aussi une abondante bibliographie. Le livre de Aaronson [**Aar97**] fournit de nombreux exemples, dans le cas d'une action de \mathbb{Z} ou de \mathbb{R} , de la variété des résultats que l'on peut obtenir lorsqu'on ne se limite pas à des mesures finies.

CHAPITRE 2

THÉORÈMES GÉNÉRAUX POUR LES GROUPES MOYENNABLES

À partir de ce chapitre, nous nous intéressons aux propriétés ergodiques d'actions de groupes localement compacts. Le cadre général est le suivant : G est un groupe localement compact à base dénombrable d'ouverts qui agit à gauche sur un espace mesuré (X, \mathcal{B}, m) en préservant la mesure σ -finie m . On considère une famille de parties mesurables F_r dans G , avec $r \in \mathbb{N}$ ou $r \in \mathbb{R}_+^*$, de mesure de Haar $|F_r|$ finie et non nulle, et on étudie le comportement, quand r tend vers l'infini, des moyennes

$$A_r f(x) = \frac{1}{|F_r|} \int_{F_r} f(gx) \, dg,$$

où $f : (X, \mathcal{B}) \rightarrow \mathbb{R}$ appartient à l'un des espaces $L^p(X, m)$, $1 \leq p < +\infty$.

Le théorème ergodique classique de Birkhoff, rappelé au chapitre 1, porte sur le cas particulier où $G = \mathbb{Z}$ et $F_n = \{0, \dots, n-1\}$. Cette suite (F_n) de parties de \mathbb{Z} possède la propriété de Følner d'invariance asymptotique par translations : pour tout $k \in \mathbb{Z}$, on a

$$\lim_{n \rightarrow +\infty} \frac{|(k + F_n) \Delta F_n|}{|F_n|} = 0,$$

la mesure de Haar sur \mathbb{Z} étant la mesure de comptage.

Le théorème de Birkhoff pour une action de \mathbb{R} est tout aussi classique. On considère en général la famille $F_r = [0, r]$ indexée par les réels $r > 0$. Cette famille possède évidemment la propriété de Følner. Les théorèmes ergodiques en "temps continu" sont tout à fait similaires à ceux du "temps discret" décrits dans le premier chapitre.

Nous limitons dans un premier temps notre étude au cas des suites de moyennes $(A_n)_{n \geq 1}$ associées à des suites de parties $(F_n)_{n \geq 1}$, le cas continu se traitant de la même façon (voir 2.2.12). Nous montrons que la propriété de Følner de la suite $(F_n)_{n \geq 1}$ suffit encore, dans le cadre général d'une action de G , pour assurer la convergence en norme L^p de $(A_n f)_{n \geq 1}$. Les groupes qui possèdent des suites de Følner, appelés groupes moyennables, ont été beaucoup étudiés. Nous en rappelons quelques propriétés et en donnons des exemples au début de ce chapitre.

Le problème de la convergence presque partout de la suite $(A_n f)_{n \geq 1}$ est plus délicat. Nous verrons dans le chapitre 3 qu'elle n'est pas toujours assurée pour une suite de Følner quelconque. Cependant, on démontre aisément la convergence presque partout pour les fonctions f d'une partie dense de $L^p(X, m)$. Suivant le principe de Banach exposé au chapitre 1, la convergence presque partout a lieu pour tout $f \in L^p(X, m)$ si et seulement si la suite des opérateurs A_n satisfait à une inégalité maximale.

L'obtention d'inégalités maximales est donc cruciale. Nous verrons apparaître un phénomène important pour les actions de groupes moyennables, le principe de transfert : si les inégalités maximales sont établies pour l'action de G sur lui-même, on peut les transférer à toute action de G sur un espace mesuré. Comme on le verra (théorème 2.3.1), on peut transférer des inégalités maximales relatives à des opérateurs de moyenne par rapport à une suite de mesures de probabilité sur G , pourvu que G soit moyennable et que les mesures vérifient une propriété de support compact. Dans le chapitre 5, ceci sera utilisé par exemple pour établir des théorèmes ergodiques relatifs aux moyennes sur les sphères de \mathbb{R}^d .

2.1. Groupes moyennables

La moyennabilité d'un groupe localement compact G peut être introduite de diverses façons. Nous privilégions ici une définition liée à la géométrie du groupe. Pour une étude détaillée nous renvoyons à [Gre69] ou [Pat88].

2.1.1. Suites de Følner. — Nous munissons G de la mesure de Haar à gauche λ . Pour alléger les notations, $|E|$ désignera le plus souvent la mesure de Haar d'une partie mesurable E de G . On écrira indifféremment $\int f(g) d\lambda(g)$ ou $\int f(g) dg$ pour l'intégrale correspondante.

Définition 2.1.1. — Soit G un groupe localement compact.

- (i) On dit qu'une suite (F_n) de parties mesurables de G est de Følner (à gauche) si les F_n sont de mesure strictement positive et finie, et si

$$\forall g \in G, \quad \lim_{n \rightarrow +\infty} \frac{|gF_n \Delta F_n|}{|F_n|} = 0.$$

- (ii) On dit qu'une suite (μ_n) de mesures de probabilité sur G est *asymptotiquement invariante* si

$$\forall g \in G, \quad \lim_{n \rightarrow +\infty} \|\delta_g * \mu_n - \mu_n\|_1 = 0,$$

où $\|\cdot\|_1$ désigne la variation totale et $\delta_g * \mu_n$ le produit de convolution de la mesure de Dirac δ_g en g par μ_n , c'est-à-dire la mesure prenant la valeur $\mu_n(g^{-1}A)$ pour tout ensemble borélien A .

Ces définitions s'étendent de façon évidente au cas des familles à un paramètre réel.

Remarque 2.1.2. — On peut de même introduire la notion de *suite de Følner à droite*. C'est une suite (F_n) telle que pour tout $g \in G$ on ait $\lim_{n \rightarrow +\infty} \frac{\lambda_d(F_n g \Delta F_n)}{\lambda_d(F_n)} = 0$, où λ_d désigne la mesure de Haar à droite définie par $\lambda_d(B) = \lambda(B^{-1})$. Clairement, (F_n) est une suite de Følner à gauche si et seulement si la suite des images F_n^{-1} par passage à l'inverse est une suite de Følner à droite.

Définition 2.1.3. — On dit qu'un groupe localement compact G est *moyennable* s'il possède une suite de Følner.

Remarque 2.1.4. — Toute suite de Følner (F_n) fournit une suite asymptotiquement invariante de mesures de probabilité : on prend pour μ_n la mesure de densité $\frac{\mathbf{1}_{F_n}}{|F_n|}$ par rapport à la mesure de Haar, c'est-à-dire la mesure de probabilité uniforme sur F_n .

La moyennabilité peut aussi être définie en terme de moyenne invariante. Notons $C^b(G)$ l'espace des fonctions continues bornées sur G , et pour $g \in G$, $f \in C^b(G)$, rappelons que $\delta_g * f$ est la fonction translatée $x \mapsto f(g^{-1}x)$.

Définition 2.1.5. — Une *moyenne* sur $C^b(G)$ est une forme linéaire positive M sur $C^b(G)$ telle que $M(\mathbf{1}) = 1$. On dit que M est *G -invariante* si $M(\delta_g * f) = M(f)$ pour tout $g \in G$ et tout $f \in C^b(G)$.

Remarquons que toute moyenne est continue de norme 1 sur $C^b(G)$ muni de la norme de la convergence uniforme.

À partir d'une suite asymptotiquement invariante (μ_n) de mesures de probabilité sur G , on montre l'existence d'une moyenne G -invariante sur $C^b(G)$ comme suit. À chaque μ_n on associe la forme linéaire continue $\varphi_n : f \mapsto \int_G f d\mu_n$ sur $C^b(G)$. Comme la boule unité du dual de $C^b(G)$ est $*$ -faiblement compacte, la suite (φ_n) y admet au moins un point d'accumulation, que nous notons M . Cette forme linéaire M est évidemment une moyenne sur $C^b(G)$. En outre, l'invariance asymptotique de (μ_n) entraîne la G -invariance de M .

Réciproquement, l'existence d'une moyenne G -invariante sur $C^b(G)$ implique la moyennabilité de G (voir [Gre69, Theorem 3.6.2]).

Remarque 2.1.6. — Il existe de nombreuses autres variantes de la notion de suite de Følner. Ainsi, dans tout groupe moyennable G il existe une suite (F_n) de parties mesurables avec $0 < |F_n| < +\infty$ (que l'on peut prendre compactes), telle que pour toute partie compacte K de G on ait $\lim_{n \rightarrow +\infty} \frac{|KF_n|}{|F_n|} = 1$. Ce résultat, établi par exemple dans [Pat88, Corollary 4.14], sera utilisé dans la démonstration du théorème de transfert. De telles suites seront appelées *suites de Følner uniformes* car alors la suite des fonctions $g \mapsto \frac{|gF_n \Delta F_n|}{|F_n|}$ converge vers 0 uniformément sur les parties compactes de G . En effet, soit C une partie compacte de G . Posons $K = C \cup C^{-1} \cup \{e\}$,

où e désigne l'élément neutre. Alors, pour $g \in C$ nous avons

$$\begin{aligned} \frac{|gF_n \Delta F_n|}{|F_n|} &= \frac{|gF_n \setminus F_n|}{|F_n|} + \frac{|g^{-1}F_n \setminus F_n|}{|F_n|} \\ &\leq 2 \frac{|KF_n \setminus F_n|}{|F_n|} = 2 \left(\frac{|KF_n|}{|F_n|} - 1 \right), \end{aligned}$$

d'où résulte notre remarque.

Exemples 2.1.7. — (1) Pour $r > 0$, notons $B(0, r)$ la boule euclidienne de centre 0 et de rayon r dans \mathbb{R}^d . Soit (r_n) une suite de réels > 0 qui tend vers $+\infty$. Alors la suite $(B(0, r_n))$ est une suite de Følner uniforme. En effet si K est une partie compacte contenue dans $B(0, p)$, on a

$$1 \leq \frac{|KB(0, r_n)|}{|B(0, r_n)|} \leq \frac{|B(0, r_n + p)|}{|B(0, r_n)|} = \frac{(r_n + p)^d}{(r_n)^d},$$

d'où $\lim_{n \rightarrow +\infty} \frac{|KB(0, r_n)|}{|B(0, r_n)|} = 1$.

(2) Plus généralement, tout groupe abélien localement compact G est moyennable. En effet, notons C la partie convexe *-faiblement compacte de la boule unité du dual de $C^b(G)$ formée des moyennes. En transposant l'action par translation de G sur $C^b(G)$, on obtient une action de G sur C par applications continues affines, qui commutent. D'après le théorème du point fixe de Markov-Kakutani [Rud91, Theorem 5.23], il existe dans C un point fixe par l'action de G , c'est-à-dire une moyenne G -invariante.

(3) Tout groupe compact est moyennable. Sa mesure de Haar fournit une moyenne invariante et la suite constante $F_n = G$ possède la propriété de Følner.

(4) La classe des groupes localement compacts moyennables est très stable. Tout sous-groupe fermé d'un tel groupe est encore moyennable [Gre69, Theorem 2.3.2]. Toute extension d'un groupe moyennable par un groupe moyennable est un groupe moyennable [Gre69, Theorem 2.3.3]. En particulier les groupes résolubles sont moyennables.

(5) À l'opposé, les groupes localement compacts connexes semi-simples (*i.e.* sans sous-groupe normal connexe résoluble non réduit à l'élément neutre) sont moyennables si et seulement si ils sont compacts [Gre69, Theorem 3.3.2].

(6) Les exemples les plus simples de groupes localement compacts non moyennables sont les groupes libres à d (≥ 2) générateurs. Considérons par exemple le cas du groupe libre \mathbb{F}_2 à deux générateurs a et b et supposons l'existence d'une moyenne invariante M sur $C^b(\mathbb{F}_2) = \ell^\infty(\mathbb{F}_2)$. Elle induit une mesure finiment additive, encore notée M , sur la tribu des parties de \mathbb{F}_2 . Désignons par E_a (resp. $E_b, E_{a^{-1}}, E_{b^{-1}}$) le sous-ensemble de \mathbb{F}_2 formé des mots commençant par la lettre a (resp. b, a^{-1}, b^{-1}). Les trois décompositions $\mathbb{F}_2 = \{e\} \sqcup E_a \sqcup E_b \sqcup E_{a^{-1}} \sqcup E_{b^{-1}}$, $\mathbb{F}_2 = E_a \sqcup aE_{a^{-1}}$,

$\mathbb{F}_2 = E_b \sqcup bE_{b^{-1}}$ en parties disjointes, et l'invariance de M , conduisent aux égalités $1 = M(\mathbb{F}_2) = M(\{e\}) + 2$, d'où la contradiction.

2.1.2. Croissance et moyennabilité. — Soient G un groupe localement compact engendré par un compact, et B un voisinage compact de l'élément neutre e tels que $\bigcup_{n=1}^{\infty} B^n = G$, où B^n désigne l'ensemble des éléments de G qui s'écrivent comme produit de n éléments de B . D'après [Gui73, Théorème I.1], la suite $(|B^n|^{1/n})$ converge vers un nombre $\gamma \geq 1$, qui décrit le *type de croissance* de G . L'égalité $\gamma = 1$ est indépendante du choix de B . Dans ce cas on dit que G est à *croissance sous-exponentielle*. Si $\gamma > 1$, on dit que G est à *croissance exponentielle*. Pour plus de détails sur ce sujet, nous renvoyons à [Gui73] et [Pat88]. Rappelons seulement quelques faits.

Un groupe à croissance sous-exponentielle est moyennable. En effet, fixons une partie compacte K de G et $\varepsilon > 0$, et montrons l'existence d'un entier q tel que $\frac{|KB^q|}{|B^q|} \leq 1 + \varepsilon$. Choisissons p entier tel que $K \subset B^p$; comme

$$\frac{|KB^{pn}|}{|B^{pn}|} \leq \frac{|B^{p(n+1)}|}{|B^{pn}|},$$

il suffit d'observer que

$$1 \leq \liminf_{n \rightarrow +\infty} \frac{|B^{p(n+1)}|}{|B^{pn}|} \leq \liminf_{n \rightarrow +\infty} |B^{pn}|^{1/n} = 1,$$

pour trouver q . En particulier, nous pouvons extraire de la suite (B^n) une sous-suite de Følner uniforme au sens de la remarque 2.1.6.

Un cas particulier important de croissance sous-exponentielle est la croissance polynomiale. On dit que G est à *croissance polynomiale* s'il existe $k > 0$ et $r \in \mathbb{N}$ tels que $|B^n| \leq kn^r$ pour tout n . Cette notion ne dépend pas du choix de B . Les groupes localement compacts nilpotents sont les exemples fondamentaux de groupes à croissance polynomiale [Pat88, Corollary. 6.19].

Rappelons qu'un groupe localement compact G est dit *unimodulaire* si sa fonction modulaire⁽¹⁾ Δ est constante, égale à 1. Tout groupe G non unimodulaire est à croissance exponentielle. En effet soit B un voisinage compact de e tel que $\bigcup_{n=1}^{\infty} B^n = G$. Pour $g \in B$, on a $|B^n| \geq |Bg^{n-1}| = \Delta(g)^{n-1}|B|$, d'où

$$\lim_{n \rightarrow +\infty} |B^n|^{1/n} \geq \sup_{g \in B} \Delta(g).$$

Si le groupe G n'est pas unimodulaire, alors Δ est un homomorphisme non trivial de G dans le groupe \mathbb{R}_+^* et, puisque que $G = \bigcup_{n=1}^{+\infty} B^n$, il existe $g \in B$ tel que $\Delta(g) > 1$; cela démontre notre assertion.

⁽¹⁾Elle est caractérisée par l'égalité $|Ex| = \Delta(x)|E|$ pour tout $x \in G$ et toute partie borélienne E de G .

Ainsi tout groupe résoluble et non unimodulaire (par exemple le groupe S_d des transformations affines de \mathbb{R}^d qui sera étudié au chapitre 4) est à la fois moyennable et à croissance exponentielle.

Les groupes connexes ainsi que les groupes résolubles sont ou bien à croissance polynomiale ou bien à croissance exponentielle [Gui73], mais en revanche, les travaux de Grigorchuk [Gri84] fournissent des exemples de groupes discrets à croissance intermédiaire entre les croissances polynomiale et exponentielle.

Remarque 2.1.8. — Soit G un groupe localement compact moyennable. La recherche de suites de Følner de nature géométrique dans G est un problème intéressant. Il sera examiné dans le chapitre 4 pour le groupe des affinités de \mathbb{R}^d .

En particulier, considérons à nouveau le cas où G est engendré par un compact. Si B est un voisinage compact de l'élément neutre tel que $\bigcup_{n=1}^{\infty} B^n = G$, la question de savoir si la suite (B^n) est de Følner se pose naturellement. Tessera a démontré que la réponse est positive si G est à croissance polynomiale [Tes05]. Récemment, Breuillard a démontré [Bre07] le résultat plus précis suivant : si G est à croissance polynomiale, il existe un entier $r(G)$ et une constante $c(B) > 0$ tels que $\lim_{n \rightarrow \infty} \frac{|B^n|}{n^{r(G)}} = c(B)$. Cette condition entraîne immédiatement que la suite (B^n) est de Følner.

Notons que lorsque la croissance est seulement sous-exponentielle, cette question reste ouverte. Enfin, si G est à croissance exponentielle, la suite (B^n) n'est jamais de Følner. Voyons pourquoi, dans le cas où G est discret. Ici B est un ensemble fini qui engendre G . Si la suite (B^n) est de Følner, on a $\lim_{n \rightarrow \infty} \frac{|B^{n+1} \setminus B^n|}{|B^n|} = 0$, d'où $\lim_{n \rightarrow \infty} \frac{|B^{n+1}|}{|B^n|} = 1$. Ainsi, pour tout $\varepsilon > 0$, il existe n_0 tel que pour tout $k \geq 0$ on ait $|B^{n_0}| \leq |B^{n_0+k}| \leq |B^{n_0}|(1 + \varepsilon)^k$. Par conséquent, la croissance de G est sous-exponentielle.

2.1.3. Trou spectral d'une représentation unitaire. — Nous abordons brièvement ce sujet ici. D'une part, il permet de formuler d'autres caractérisations utiles de la moyennabilité des groupes et, d'autre part, il intervient dans l'estimation de la vitesse de convergence des moyennes ergodiques (voir la postface).

Définition 2.1.9. — Soit π une représentation unitaire continue d'un groupe localement compact G dans un espace hilbertien H . On dit que π possède presque des vecteurs invariants s'il existe une suite (ξ_n) de vecteurs de norme 1 dans H telle que

$$\lim_{n \rightarrow +\infty} \|\pi(g)\xi_n - \xi_n\| = 0$$

uniformément sur tout compact de G .

On dit aussi dans ce cas que la représentation triviale de G est faiblement contenue dans la représentation π .

Soient μ une mesure de probabilité sur G et $\pi(\mu) = \int_G \pi(g) d\mu(g)$ l'opérateur (de norme ≤ 1) associé sur H (voir l'appendice A). On voit facilement que $\|\pi(\mu)\| = 1$ dès lors que π possède presque des vecteurs invariants.

Définition 2.1.10. — On dit qu'une représentation unitaire continue π possède un *trou spectral* s'il existe une mesure de probabilité μ sur G , absolument continue par rapport à la mesure de Haar, symétrique (*i.e.* $\mu = \check{\mu}$, où $\check{\mu}$ désigne la mesure $A \mapsto \mu(A^{-1})$), dont le support engendre un sous-groupe dense dans G , telle que $\|\pi(\mu)\| < 1$.

En fait, la définition ne dépend pas du choix de μ et équivaut à la non existence de suites presque invariantes de vecteurs, comme le montre la proposition suivante (pour sa démonstration, voir par exemple [BdlHV08, Proposition G.4.2]).

Proposition 2.1.11. — Une représentation unitaire continue π de G possède presque des vecteurs invariants si et seulement si il existe une mesure de probabilité μ sur G , absolument continue, symétrique, dont le support engendre un sous-groupe dense dans G , et telle que $\|\pi(\mu)\| = 1$.

Il résulte de travaux de Hulanicki, Reiter, Kesten, Derriennic et Guivarc'h qu'un groupe localement compact G est moyennable si et seulement si la représentation régulière par translations à gauche dans $L^2(G, \lambda)$ ne possède pas de trou spectral (voir [BdlHV08, Appendix G] pour les détails).

À l'opposé, se situe la classe ci-dessous des groupes de Kazhdan qui, bien que n'apparaissant que très peu dans cet ouvrage, joue néanmoins un rôle important en théorie ergodique (voir la postface).

Définition 2.1.12. — On dit qu'un groupe localement compact G a la *propriété (T) de Kazhdan* si toute représentation unitaire continue π de G sans vecteur invariant non nul possède un trou spectral.

Évidemment, un tel groupe ne peut être moyennable que s'il est compact. Parmi la vaste classe des groupes ayant la propriété (T), citons les groupes de Lie semi-simples, connexes, de centre fini, dont chaque facteur est de rang ≥ 2 , par exemple les groupes $SL(n, \mathbb{R})$, $n \geq 3$. Le lecteur intéressé par une étude approfondie pourra consulter le livre [BdlHV08] consacré à ce sujet.

La proposition suivante montre qu'on dispose d'une estimation *uniforme* des normes des opérateurs $\pi(\mu)$ lorsque le groupe a la propriété (T).

Proposition 2.1.13. — Soit G un groupe localement compact ayant la propriété (T) et soit μ une mesure de probabilité sur G , absolument continue, symétrique, dont le support engendre un sous-groupe dense. Il existe une constante $\delta < 1$ (ne dépendant que de μ) telle que pour toute représentation unitaire continue π de G , sans vecteur invariant non nul, on ait $\|\pi(\mu)\| \leq \delta$.

Démonstration. — Supposons au contraire que pour tout entier $n \geq 1$, il existe une représentation unitaire π_n , sans vecteur invariant non nul, telle que $\|\pi_n(\mu)\| \geq 1 - 1/n$. Posons $\pi = \bigoplus_{n \geq 1} \pi_n$. Alors $\|\pi(\mu)\| = 1$ et, puisque G a la propriété (T), la représentation π possède un vecteur invariant $\xi = \bigoplus_{n \geq 1} \xi_n$ non nul. Toute composante ξ_{n_0} non nulle est un vecteur invariant de la représentation π_{n_0} correspondante, d'où la contradiction. \square

Pour d'autres informations concernant la propriété de trou spectral, nous renvoyons par exemple à [BG06].

2.2. Théorèmes ergodiques pour les suites de Følner

Soit G un groupe localement compact agissant à gauche sur un espace mesuré (X, \mathcal{B}, m) . L'action $(g, x) \mapsto gx$ est une application mesurable de $G \times X$ dans X , et nous supposons toujours que la mesure σ -finie m est G -invariante. Nous utilisons les notations introduites dans l'appendice A. En particulier, pour $p \in [1, +\infty[$, nous considérons la représentation continue π de G dans $L^p(X, m)$ par translations à gauche. Si μ appartient à l'algèbre de convolution $\mathcal{M}(G)$ des mesures complexes de variation totale finie sur G et si $f \in L^p(X, m)$, on note $\pi(\mu)f$, ou $\mu * f$, le produit de convolution de μ par f , c'est-à-dire l'élément

$$x \mapsto \int_G f(g^{-1}x) d\mu(g)$$

de $L^p(X)$. Rappelons que $\|\mu * f\|_p \leq \|\mu\|_1 \|f\|_p$.

Nous aurons plus souvent à considérer $T(\mu)f = \pi(\tilde{\mu})f$, aussi noté $\mu \cdot f$:

$$T(\mu)f = \mu \cdot f(x) = \int_G f(gx) d\mu(g).$$

Dans le cas courant où μ est symétrique, on a bien sûr $\mu \cdot f = \mu * f$. Pour $\mu, \nu \in \mathcal{M}(G)$, remarquons que $(\mu * \nu) * f = \mu * (\nu * f)$ alors que $(\mu * \nu) \cdot f = \nu \cdot (\mu \cdot f)$. Étant donné $g \in G$, on pose $T_g = T(\delta_g) = \pi(g^{-1})$. Alors, $(g, f) \mapsto T_g f$ est une action à droite par isométries de G sur $L^p(X, m)$.

2.2.1. Théorèmes ergodiques en moyenne dans L^p . — Nous supposons toujours $p \in [1, +\infty[$. Si $f \in L^p(X, m)$, nous notons C_f l'enveloppe convexe fermée en norme de l'ensemble $\{T_g f : g \in G\}$. Cet ensemble convexe est faiblement fermé et borné en norme. Lorsque $p > 1$, il est donc faiblement compact. C'est un fait crucial pour la démonstration du théorème suivant, où l'on voit en particulier que la moyennabilité du groupe entraîne une propriété de point fixe. Dans la suite, \mathcal{I} désigne la tribu des ensembles mesurables E de X qui sont G -invariants, c'est-à-dire tels que $m(gE \Delta E) = 0$ pour tout $g \in G$.

Théorème 2.2.1. — *Soit G un groupe localement compact moyennable agissant sur (X, \mathcal{B}, m) en laissant la mesure σ -finie m invariante. Soit $p \in [1, +\infty[$.*

- (i) Pour tout $f \in L^p(X, m)$, l'ensemble C_f contient un point fixe et un seul pour l'action de G , noté $E_1(f)$.
- (ii) Soit (μ_n) une suite asymptotiquement invariante de mesures de probabilité sur G . Alors $(\mu_n \cdot f)$ converge vers $E_1(f)$ dans $L^p(X, m)$.
- (iii) $E_1 : f \mapsto E_1(f)$ est un projecteur contractant et son noyau est le sous-espace vectoriel fermé de $L^p(X, m)$ engendré par $\{f - T_g f : f \in L^p(X, m), g \in G\}$.
- (iv) Si m est une mesure de probabilité, alors $E_1(f)$ est l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$.

Nous avons déjà observé dans le chapitre précédent, pour des actions de \mathbb{Z} , que ce théorème ne s'étend pas aux cas $p = 1$ et $p = +\infty$ (voir les remarques 1.3.7 et 1.3.9). Toutefois, lorsque m est une mesure finie, il reste vrai pour $p = 1$, comme nous le verrons dans le théorème 2.2.4.

Démonstration. — Démontrons simultanément (i) et (ii). Comme $(\mu_n * f)$ est une suite d'éléments du convexe faiblement compact C_f , on peut en extraire une sous-suite $(\mu_{n_k} * f)$ qui converge faiblement vers un élément f_0 de C_f . Vérifions que f_0 est G -invariant. Soient $g \in G$ et $h \in L^q(X, m)$, où q désigne l'exposant conjugué de p . On a

$$|\langle \delta_g * \mu_{n_k} * f, h \rangle - \langle \mu_{n_k} * f, h \rangle| \leq \|\delta_g * \mu_{n_k} - \mu_{n_k}\|_1 \|f\|_p \|h\|_q,$$

d'où

$$\langle \delta_g * f_0, h \rangle = \langle f_0, h \rangle$$

par passage à la limite. Il en résulte que $\delta_g * f_0 = f_0$.

Soit $h \in C_f$ un point fixe pour l'action de G . C'est aussi un point fixe pour la convolution par toute mesure de probabilité sur G , d'où $\mu_n \cdot h = h$ pour tout n . Montrons que $\lim_{n \rightarrow +\infty} \mu_n \cdot f = h$, ce qui entraînera à la fois l'unicité du point fixe et l'assertion (ii). Étant donné $\varepsilon > 0$, choisissons une combinaison convexe $v = \sum_{i=1}^k \alpha_i (T_{g_i} f)$ telle que $\|h - v\|_p \leq \varepsilon/2$. On a alors

$$\begin{aligned} \|h - \mu_n \cdot f\|_p &\leq \|h - \mu_n \cdot v\|_p + \sum_{i=1}^k \alpha_i \|\mu_n \cdot (T_{g_i} f) - \mu_n \cdot f\|_p \\ &\leq \|h - v\|_p + \sum_{i=1}^k \alpha_i \|\delta_{g_i} * \mu_n - \mu_n\|_1 \|f\|_p \\ &\leq \varepsilon/2 + \sum_{i=1}^k \alpha_i \|\delta_{g_i} * \mu_n - \mu_n\|_1 \|f\|_p, \end{aligned}$$

d'où $\|h - \mu_n \cdot f\|_p \leq \varepsilon$ si n est assez grand.

Démontrons (iii). Comme les applications $f \mapsto \mu_n \cdot f$ sont contractantes, on a $\|E_1(f)\|_p \leq \|f\|_p$ pour tout $f \in L^p(X, m)$. On vérifie immédiatement que E_1 est un projecteur sur le sous-espace des éléments G -invariants de $L^p(X, m)$. Les éléments de

la forme $h - T_g h$ sont évidemment dans le noyau de E_1 . De plus, si $f \in \text{Ker } E_1$, on a $0 \in C_f$. Par conséquent, étant donné $\varepsilon > 0$, on peut trouver une combinaison convexe $\sum_{i=1}^k \alpha_i (T_{g_i} f)$ de norme majorée par ε , c'est-à-dire que $\|f - \sum_{i=1}^k \alpha_i (f - T_{g_i} f)\|_p \leq \varepsilon$.

Vérifions enfin l'assertion (iv). Étant donné $A \in \mathcal{I}$, on a

$$\langle E_1(f), \mathbf{1}_A \rangle = \lim_n \langle \mu_n \cdot f, \mathbf{1}_A \rangle = \lim_n \langle f, \check{\mu}_n \cdot \mathbf{1}_A \rangle.$$

Or $\check{\mu}_n \cdot \mathbf{1}_A = \mathbf{1}_A$ pour tout n puisque A est G -invariant, d'où $\langle E_1(f), \mathbf{1}_A \rangle = \langle f, \mathbf{1}_A \rangle$. \square

Remarque 2.2.2. — Il résulte de la démonstration du théorème précédent que pour $f \in L^p(X, m)$, la suite $(\mu_n * f)$ n'a pas d'autre valeur d'adhérence que $E_1(f)$ dans C_f muni de la topologie faible. Comme C_f est faiblement compacte et métrisable, cette suite converge faiblement vers $E_1(f)$. En revanche la suite $(\mu_n \cdot f)$ converge vers $E_1(f)$ en norme L^p .

Observons par ailleurs que $L^p(X, m)$ est la somme directe topologique du sous-espace des fonctions G -invariantes et du sous-espace vectoriel fermé engendré par les cobords $f - T_g f$, avec $f \in L^p(X, m)$ et $g \in G$.

Remarque 2.2.3. — Dans le cas $p = 2$, l'espace hilbertien $H = L^2(X, m)$ est la somme directe orthogonale de $H_1 = E_1 H$, sous-espace des fonctions G -invariantes, et de H_0 qui est engendré par les fonctions de la forme $f - T_g f$, avec $f \in L^2(X, m)$ et $g \in G$. L'espace H_1 étant stable par la représentation unitaire π de G dans H , il en est de même pour H_0 . On notera π_0 la restriction de π à H_0 . Si $f \in H_0$, le théorème 2.2.1 montre que $\pi_0(\check{\mu}_n) f = \mu_n \cdot f$ tend vers 0 quand n tend vers l'infini.

Dans l'exemple de l'action de \mathbb{Z} sur le cercle par rotations irrationnelles, et en prenant pour μ_n la probabilité uniforme sur $\{0, 1, \dots, n-1\}$, on a vu au chapitre précédent (remarque 1.1.5) que $\|\pi_0(\check{\mu}_n)\| = 1$ pour tout n . Cette propriété de \mathbb{Z} reflète une propriété générale des groupes moyennables. En effet, soit G un groupe agissant sur un espace (X, \mathcal{B}, m) en préservant m supposée finie, et notons comme ci-dessus π_0 la restriction de π à l'orthogonal H_0 dans $L^2(X, m)$ du sous-espace H_1 des fonctions G -invariantes. Nous supposons que l'action est ergodique, ce qui équivaut à dire que H_1 est réduit aux fonctions constantes. La représentation π_0 ne possède pas de vecteur (non nul) invariant, mais un résultat de K. Schmidt assure que si G est un groupe discret moyennable, π_0 possède presque des vecteurs invariants (voir [Sch81, Theorem 2.4, Proposition 2.3]). Si μ est une mesure de probabilité sur G , on a donc $\|\pi_0(\mu)\| = 1$.

En revanche, si le groupe G a la propriété (T) de Kazhdan, la représentation π_0 possède un trou spectral.

Théorème 2.2.4. — *Dans le cas $p = 1$, les résultats du théorème 2.2.1 restent valables lorsque m est une mesure de probabilité.*

Démonstration. — Le passage du théorème ergodique en moyenne L^2 au théorème ergodique en moyenne L^1 , lorsque la mesure est finie, est classique. Il a déjà été

mentionné dans la démonstration du théorème de Hopf au chapitre 1. Il utilise le fait que $L^2(X, m)$ s'injecte continûment et densément dans $L^1(X, m)$ et que les applications $f \mapsto \mu_n \cdot f$ sont des contractions de $L^1(X, m)$. On voit alors que l'ensemble des $f \in L^1(X, m)$ tels que $(\mu_n \cdot f)$ converge dans $L^1(X)$ est fermé et contient $L^2(X, m)$. C'est donc $L^1(X, m)$. De plus, on vérifie aisément que pour tout $f \in L^1(X, m)$, la limite $E_1(f)$ de $(\mu_n \cdot f)$ est égale à $\mathbb{E}(f|\mathcal{I})$.

Assurons-nous, dans ce cas encore, que $\text{Ker } E_1$ est le sous-espace vectoriel fermé engendré par les $f - T_g f$. Ces fonctions appartiennent évidemment à $\text{Ker } E_1$. Soit $h \in L^\infty(X, m)$ telle que

$$\int_X (f - T_g f) h \, dm = 0$$

pour tout $g \in G$ et tout $f \in L^1(X, m)$. Il s'agit de montrer que $\int_X k h \, dm = 0$ pour toute fonction k dans le noyau de E_1 . Comme

$$\int_X (f - T_g f) h \, dm = \int_X f (h - T_{g^{-1}} h) \, dm,$$

on voit que h est G -invariante. On a donc

$$\int_X h k \, dm = \int_X h E_1(k) \, dm = 0$$

si $E_1(k) = 0$, ce qu'on voulait démontrer. \square

Remarque 2.2.5. — Plaçons-nous toujours dans le cas $p = 1$ avec $m(X) < +\infty$. Soit $f \in L^1(X)$. L'enveloppe convexe fermée en norme C_f de l'ensemble $\{T_g f : g \in G\}$ possède la propriété d'équi-intégrabilité suivante : pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que si $m(E) \leq \delta$ on a $\int_E |h| \, dm \leq \varepsilon$ pour tout $h \in C_f$. Comme de plus C_f est bornée, c'est une partie faiblement compacte (voir [DS88, Th. 9, p. 292]). C'est la propriété cruciale de C_f utilisée dans la démonstration du théorème 2.2.1. On peut donc aussi reproduire celle-ci pour établir directement le théorème 2.2.4.

Remarque 2.2.6. — On dit que l'action d'un groupe localement compact G sur un espace mesuré (X, \mathcal{B}, m) est ergodique si tout élément de \mathcal{B} invariant par l'action de G est soit de mesure nulle, soit de complémentaire de mesure nulle. Dans le cas où m est finie, cela se traduit de façon équivalente par le fait que $L^1(X, m)$ (ou n'importe quel espace $L^p(X, m)$, $1 \leq p < +\infty$) ne contient pas d'autres fonctions G -invariantes que les fonctions constantes. Si de plus m est une mesure de probabilité, pour tout $f \in L^1(X, m)$, l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$ qui apparaît dans les théorèmes ergodiques n'est autre que la moyenne $\int_X f \, dm$.

Le principe de démonstration du théorème 2.2.1 peut être utilisé pour établir le résultat plus général ci-dessous. Précisons d'abord quelques notations. Soit B un espace de Banach et, pour tout $g \in G$, soit T_g un opérateur borné de B . Nous supposons que $g \mapsto T_g$ est une antireprésentation de G , continue et uniformément bornée, c'est-à-dire que $T_{g_1 g_2} = T_{g_2} \circ T_{g_1}$ pour tous $g_1, g_2 \in G$, que pour tout $\xi \in B$

la fonction $g \mapsto T_g \xi$ est continue de G dans B , et enfin que $\sup_{g \in G} \|T_g\| < +\infty$. Pour $\xi \in B$, nous notons C_ξ l'enveloppe convexe fermée de $\{T_g \xi : g \in G\}$.

Théorème 2.2.7. — *Supposons que G est moyennable et que les parties convexes C_ξ , $\xi \in B$ sont faiblement compactes (ce qui est réalisé par exemple lorsque B est réflexif). Alors :*

- (i) $B = B_1 \oplus B_0$ (somme directe topologique) où B_1 est le sous-espace des vecteurs fixes de l'action et B_0 est le sous-espace vectoriel fermé engendré par $\{\xi - T_g \xi : \xi \in B, g \in G\}$.
- (ii) Soit (μ_n) une suite asymptotiquement invariante de mesures de probabilité sur G . Pour tout $\xi \in B$, la suite des moyennes $\int_G T_g \xi d\mu_n(g)$ converge vers la projection de ξ sur B_1 parallèlement à B_0 . Cette projection est l'unique point fixe de l'action de G appartenant à C_ξ .

Donnons un exemple d'application du théorème précédent dans le cas où l'espace de Banach B n'est pas réflexif.

Si G est un groupe topologique agissant sur un espace topologique X , nous disons que l'action est continue si l'application $(g, x) \mapsto gx$ est continue de $G \times X$ dans X .

Corollaire 2.2.8. — *Soit G un groupe localement compact moyennable agissant continûment à gauche par isométries sur un espace métrique compact X . Soit (μ_n) une suite asymptotiquement invariante de mesures de probabilité sur G . Pour toute fonction continue f sur X , la suite $(\mu_n \cdot f)$ converge uniformément sur X vers une fonction invariante sous l'action de G .*

Démonstration. — L'action à gauche de G sur X , puisqu'elle est continue, induit une action à droite de G sur l'espace de Banach $C(X)$ des fonctions continues sur X , à valeurs complexes, muni de la norme de la convergence uniforme : pour $g \in G$, $f \in C(X)$ et $x \in X$, on a $T_g f(x) = f(gx)$. Puisque G agit par isométries, pour chaque $f \in C(X)$, l'ensemble $\{T_g f : g \in G\}$ est équicontinu. Il résulte alors du théorème d'Ascoli que l'enveloppe convexe fermée C_f de $\{T_g f : g \in G\}$ est compacte dans $C(X)$, et donc aussi faiblement compacte. Le corollaire se déduit alors directement du théorème 2.2.7. \square

Remarque 2.2.9. — Une autre situation où l'on observe la convergence uniforme de la suite $(\mu_n \cdot f)$ (pour les fonctions continues f) est celle de l'unique ergodicité.

Soit G un groupe localement compact agissant continûment sur un espace métrique compact X , muni de sa structure borélienne. Supposons que l'action est *uniquement ergodique*, c'est-à-dire qu'il existe une unique mesure de probabilité sur X invariante sous l'action de G . Notons m cette mesure invariante.

Supposons de plus que G est moyennable et que (μ_n) est une suite asymptotiquement invariante de mesures de probabilité sur G . Pour chaque entier n , la fonction

$\mu_n \cdot f : x \mapsto \int_G f(gx) d\mu_n(g)$ est bien définie et continue sur X . Le théorème de Kryloff-Bogoliouboff affirme que la suite $(\mu_n \cdot f)$ converge uniformément vers $\int_X f dm$. Ce théorème classique se trouve expliqué, pour une action de \mathbb{R} , dans l'article original [KB37] et, pour une action de \mathbb{Z} , des références très accessibles sont [Oxt52], [Gla03, chap. 4] ou [CFS82, chap.1, §8] (le cas de l'action d'un groupe moyennable se traite comme celui d'une action de \mathbb{Z}).

Remarque 2.2.10. — L'action de G sur X est dite *topologiquement transitive* s'il existe un point de X d'orbite dense. Sous les hypothèses du corollaire 2.2.8, c'est-à-dire quand le groupe G agit par isométries sur l'espace X , la propriété d'unique ergodicité est satisfaite dès que l'action de G sur X est topologiquement transitive. En effet, notons $\text{Iso}(X)$ le groupe des isométries de X . C'est un groupe compact pour la topologie de la convergence uniforme sur X . À l'action de G est associé un homomorphisme ϕ continu de G dans $\text{Iso}(X)$. L'adhérence $\overline{\phi(G)}$ de $\phi(G)$ dans $\text{Iso}(X)$ agit transitivement sur X puisque ses orbites sont fermées et que l'une d'elles est dense. Ainsi, X apparaît comme un espace homogène du groupe compact $\overline{\phi(G)}$. Il possède en particulier une seule mesure de probabilité $\phi(G)$ -invariante.

2.2.2. Un sous-espace dense de convergence ponctuelle pour des suites de Følner. — Revenons au cas d'une action sur un espace mesuré (X, \mathcal{B}, m) et intéressons-nous maintenant à la convergence ponctuelle. Un exemple de suite de Følner dans \mathbb{Z} pour lequel le théorème ergodique ponctuel est faux sera donné en 3.3.5. La convergence ponctuelle est toutefois assurée sur un sous-espace dense :

Proposition 2.2.11. — *Soit G un groupe localement compact moyennable agissant sur (X, \mathcal{B}, m) en laissant invariante la mesure σ -finie m . Soit $p \in [1, +\infty[$, avec la convention que si $p = 1$, la mesure m est finie. Soit (μ_n) une suite asymptotiquement invariante de mesures de probabilité sur G . Il existe un sous-espace vectoriel \mathcal{D} dense dans $L^p(X, m)$ tel que, pour tout f appartenant à \mathcal{D} , la suite $(\mu_n \cdot f)$ converge presque partout vers une fonction G -invariante.*

Démonstration. — Si f est G -invariante on a $\mu_n \cdot f = f$ pour tout n . Prenons maintenant f de la forme $h - T_g h$ avec $h \in L^\infty(X) \cap L^p(X)$. On a

$$\begin{aligned} |(\mu_n \cdot f)(x)| &= \left| \int (h(g'x) - h(gg'x)) d\mu_n(g') \right| \\ &= \left| \int h(g'x) d(\mu_n - \delta_g * \mu_n)(g') \right| \\ &\leq \|h\|_\infty \|\mu_n - \delta_g * \mu_n\|_1, \end{aligned}$$

d'où il résulte que la suite $(\mu_n \cdot f)$ converge vers 0 presque partout. Pour conclure, on utilise les théorèmes 2.2.1 et 2.2.4 : ils impliquent que l'espace vectoriel \mathcal{D} engendré par les fonctions G -invariantes et les fonctions de la forme $h - T_g h$, avec $h \in L^\infty(X, m) \cap L^p(X, m)$ est dense dans $L^p(X, m)$. \square

Remarque 2.2.12. — Le plus souvent, les résultats précédents seront appliqués au cas où la suite (μ_n) provient d'une suite de Følner (F_n) , comme dans la remarque 2.1.4.

Nous aurons aussi à considérer des familles asymptotiquement invariantes $(\mu_r)_{r>0}$ de mesures de probabilité sur G , indexées par \mathbb{R}_+^* . Tous les résultats de la section 2.2.1 s'étendent à ce cadre. En ce qui concerne la proposition 2.2.11, le seul problème est que les fonctions $\mu_r \cdot f$ ne sont définies que presque partout. Il n'existe donc pas nécessairement d'ensemble négligeable en dehors duquel toutes ces fonctions soient définies. Malgré tout, l'assertion s'étend à la famille $(\mu_r \cdot f)_{r \in \mathbb{Q}^+}$ par exemple. La question de la définition presque partout de la famille $(\mu_r \cdot f)_{r \in \mathbb{R}_+^*}$ est discutée dans l'appendice A.

2.3. Le principe de transfert

L'intérêt des inégalités maximales, par l'intermédiaire du principe de Banach, a été mis en évidence au chapitre 1. Elles permettent d'étendre la convergence presque partout pour une partie dense de $L^p(X, m)$, comme cela a été obtenu dans la proposition 2.2.11, à la convergence presque partout pour tout élément de $L^p(X, m)$. Dans le cas de l'action d'un groupe moyennable, nous allons montrer que de telles inégalités peuvent être obtenues à partir d'inégalités maximales pour l'action du groupe sur lui-même. Cette propriété est connue sous le nom de principe de transfert de Calderón.

Soit $(\mu_r)_{r>0}$ une famille continue à un paramètre de mesures de probabilité sur G , c'est-à-dire telle que pour toute fonction φ continue bornée sur G , la fonction $r \mapsto \mu_r(\varphi)$ est continue. En particulier, pour tout $g \in G$, la fonction $r \mapsto \mu_r \cdot \varphi(g)$ est continue. La fonction maximale φ_μ^* est définie par

$$\varphi_\mu^*(g_0) = \sup_{r>0} |\mu_r \cdot \varphi(g_0)| = \sup_{r>0} \left| \int_G \varphi(gg_0) d\mu_r(g) \right|.$$

C'est une fonction mesurable puisque, grâce à l'hypothèse de continuité, on a $\varphi_\mu^*(g_0) = \sup_{r \in \mathbb{Q}^+} |\mu_r \cdot \varphi(g_0)|$. La définition de la fonction maximale associée à une fonction de L^p soulève des difficultés qui sont examinées dans l'appendice A.

Théorème 2.3.1 (Principe de transfert). — Soit $(\mu_r)_{r>0}$ une famille continue de mesures de probabilité sur un groupe localement compact moyennable G . Soit $p \in [1, \infty[$. Nous faisons les deux hypothèses suivantes :

- (a) pour tout $R > 0$, il existe une partie compacte K_R de G telle que si $r \leq R$, le support de μ_r est contenu dans K_R ;
- (b) il existe une constante $C > 0$ telle que, pour toute fonction φ continue à support compact sur G , on a $\|\varphi_\mu^*\|_{L^p(G)} \leq C \|\varphi\|_{L^p(G)}$.

Soit (X, G, m) un G -espace mesuré, où m est finie, et soit $f \in L^p(X, m)$. Alors, il existe un ensemble $E \subset X$, négligeable tel que

- (i) pour $x \notin E$ et $r > 0$, la fonction $g \mapsto f(gx)$ est μ_r -intégrable ;
(ii) pour $x \notin E$, la fonction $r \mapsto \mu_r \cdot f(x)$ est continue sur $]0, +\infty[$.
- De plus, en posant $f_\mu^*(x) = \sup_{r>0} |\mu_r \cdot f(x)|$ si $x \notin E$, on a

$$\|f_\mu^*\|_{L^p(X)} \leq C \|f\|_{L^p(X)}.$$

Démonstration. — Comme c'est expliqué dans l'appendice A, on peut supposer que X est compact et que l'action de G sur X est continue.

Soit f une fonction continue sur X . On fixe $R > 0$ ainsi qu'une partie compacte K de G contenant le support de μ_r pour tout $r \leq R$. On fixe également un compact F de G et un voisinage compact V de e . Choisissons une fonction continue positive ψ sur G , à support compact, telle que

$$\mathbf{1}_{KF} \leq \psi \leq \mathbf{1}_{KFV}.$$

Pour $x \in X$, posons $\varphi_x(g) = \psi(g)f(gx)$. La fonction φ_x est continue à support compact. De plus, pour tout $h \in F$ et tout $r \leq R$, on a

$$\int_G f(ghx) d\mu_r(g) = \int_G \varphi_x(gh) d\mu_r(g),$$

c'est-à-dire $\mu_r \cdot f(hx) = \mu_r \cdot \varphi_x(h)$. En utilisant l'hypothèse (b), on en déduit que

$$\begin{aligned} \int_F \sup_{0 < r \leq R} |\mu_r \cdot f(hx)|^p dh &\leq \int_G \sup_{0 < r \leq R} |\mu_r \cdot \varphi_x(h)|^p dh \\ &\leq C^p \int_G |\varphi_x(h)|^p dh \leq C^p \int_{KFV} |f(hx)|^p dh. \end{aligned}$$

On intègre ensuite sur X et, grâce à l'invariance de m sous l'action de G , on obtient

$$|F| \int_X \sup_{0 < r \leq R} |\mu_r \cdot f(x)|^p dm(x) \leq C^p |KFV| \int_X |f(x)|^p dm(x).$$

Puisque G est moyennable, pour tout $\varepsilon > 0$, nous pouvons choisir F tel que $\frac{|KF|}{|F|} \leq 1 + \varepsilon$, d'après la remarque 2.1.6. Ensuite on choisit V tel que $|KFV| \leq |KF|(1 + \varepsilon)$. Il en résulte que

$$\int_X \sup_{0 < r \leq R} |\mu_r \cdot f(x)|^p dm(x) \leq (1 + \varepsilon)^2 C^p \int_X |f(x)|^p dm(x).$$

L'inégalité maximale s'en déduit en faisant tendre R vers $+\infty$ et ε vers 0. Pour traiter le cas de $f \in L^p(X, m)$, on fait ensuite appel à la proposition A.2.2. \square

Remarque 2.3.2. — Si l'hypothèse (b) est remplacée par une inégalité maximale de type faible, on obtient de même (i) et (ii), ainsi que la conclusion finale du théorème précédent sous la forme d'une inégalité maximale de type faible.

On déduit alors des résultats obtenus dans ce chapitre le théorème ergodique suivant.

Théorème 2.3.3. — Soit G un groupe localement compact moyennable agissant sur (X, \mathcal{B}, m) en laissant invariante la mesure de probabilité m . Soit $(\mu_r)_{r>0}$ une famille continue à un paramètre de mesures de probabilité sur G satisfaisant aux conditions (a) et (b) du théorème précédent. On la suppose asymptotiquement invariante, c'est-à-dire telle que $\lim_{r \rightarrow +\infty} \|\delta_g * \mu_r - \mu_r\|_1 = 0$ pour tout $g \in G$. Soit $p \in [1, +\infty[$. Alors, pour tout $f \in L^p(X, m)$, on a $\lim_{r \rightarrow +\infty} \mu_r \cdot f = \mathbb{E}(f|\mathcal{I})$ presque partout et dans L^p , où $\mathbb{E}(f|\mathcal{I})$ désigne l'espérance conditionnelle de f sachant la tribu \mathcal{I} des invariants.

Démonstration. — En utilisant la proposition 2.2.11, les théorèmes 2.2.1, 2.2.4 et 2.3.1, ainsi que le principe de Banach, on voit que

$$\lim_{r \rightarrow +\infty, r \in \mathbb{Q}^+} \mu_r \cdot f = \mathbb{E}(f|\mathcal{I}),$$

presque partout et dans L^p . De plus, d'après le théorème 2.3.1, on sait que pour presque tout $x \in X$, la fonction $r \mapsto \mu_r \cdot f(x)$ est continue. La conclusion du théorème en résulte immédiatement. \square

Remarque 2.3.4. — Le théorème 2.3.1 admet plusieurs variantes.

Tout d'abord, il arrive souvent qu'on n'ait pas besoin de passer par l'intermédiaire d'une action continue sur un espace compact pour justifier les propriétés (i) et (ii) du théorème 2.3.1. C'est le cas par exemple si $G = \mathbb{R}^d$, lorsque, pour tout $r > 0$, μ_r est la mesure de probabilité uniforme sur la boule de centre 0 et de rayon r . Alors le théorème de transfert reste vrai sous les hypothèses (a) et (b) sans supposer que la mesure m est finie (cette hypothèse étant seulement utilisée pour avoir (i) et (ii) comme cela est expliqué dans l'appendice A).

Cette remarque s'applique aussi lorsqu'on travaille avec une suite $(\mu_n)_{n \in \mathbb{N}}$ de mesures de probabilité à support compact au lieu de $(\mu_r)_{r>0}$.

Dans de tels cas, les conclusions du théorème 2.3.3 restent vraies avec m σ -finie, excepté l'affirmation relative à la convergence de $\mu_r \cdot f$ dans L^1 .

On peut également transférer des inégalités mettant en jeu une suite (T_n) d'opérateurs semi-locaux sur $L^1_{loc}(G)$ (espace des fonctions localement intégrables sur G) à valeurs dans l'espace des fonctions continues sur G , au lieu de la suite des applications $\varphi \mapsto \mu_n \cdot \varphi$. Il suffit que chaque T_n soit sous-additif, commute avec les translations à droite, et soit *semi-local*, c'est-à-dire qu'il existe un compact K_n de G tel que si $\varphi \in L^1_{loc}(G)$ est nulle en dehors d'un sous-ensemble E , alors $T_n \varphi$ est nulle en dehors de $K_n E$ (voir [Eme74, Theorem 2]).

Notice bibliographique. — Comme on l'a vu, la démonstration du théorème ergodique en moyenne 2.2.1 se fait en établissant un théorème d'existence et d'unicité de point fixe. Cette méthode, qui apparaît notamment dans les travaux de F. Riesz [Rie38], de Yosida et Kakutani [YK41], et de Dunford [Dun39], a été mise au point dans un cadre abstrait par Eberlein [Ebe49]. Sa présentation est ici inspirée de l'article [Gre73] de Greenleaf.

Les théorèmes ergodiques ponctuels, pour des moyennes relatives à des suites de Følner, ont fait l'objet de nombreuses études. Les premiers cas traités concernent les moyennes sur les boules de \mathbb{Z}^d et \mathbb{R}^d , avec notamment les travaux fondateurs de Birkhoff (cas du groupe \mathbb{Z} [Bir31]) et de Wiener [Wie39]. Calderón fut le premier à considérer des actions de groupes non abéliens [Cal53], en prenant des moyennes sur des suites de Følner particulières, possédant entre autres une propriété de doublement de volume. Cette propriété sera considérée dans le chapitre 3.

Dès le départ, l'importance des inégalités maximales en théorie ergodique a été mise en évidence par Birkhoff [Bir31]. À la même époque, elles ont été établies par Hardy et Littlewood [HL30], en vue d'applications à l'analyse complexe, pour des moyennes sur des intervalles de \mathbb{R} . Le principe de réduction des inégalités maximales au cas de l'action naturelle de G sur lui-même apparaît déjà dans l'article de Wiener [Wie39]. Il a été formalisé par Calderón [Cal68], puis généralisé par divers auteurs dont Coifman et Weiss [CW76], Emerson [Eme74], Herz [Her70] et Tempel'man [Tem67].

CHAPITRE 3

INÉGALITÉS MAXIMALES ET LEMMES DE RECouvreMENT

Soit G un groupe localement compact et soit (F_n) une suite de parties mesurables de G avec $0 < |F_n| < +\infty$. Considérons les opérateurs de moyenne donnés, pour $\varphi \in L^p(G)$, par

$$A_n\varphi(x) = \frac{1}{|F_n|} \int_{F_n} \varphi(gx) \, dg = \frac{1}{|F_nx|} \int_{F_nx} \varphi(g) \, dg,$$

et l'opérateur maximal correspondant donné par

$$A^*\varphi = \sup_{n \geq 1} |A_n\varphi|.$$

Sous l'hypothèse que (F_n) est une suite de Følner, nous avons établi dans le chapitre précédent le théorème ergodique ponctuel dans L^p pour les moyennes A_n , lorsque A^* est de type faible (p, p) .

Que la suite (F_n) soit ou non de Følner, l'étude de l'opérateur A^* est utile en théorie ergodique et en analyse classique. Dans ce chapitre, nous examinons diverses techniques de recouvrement qui conduisent à des théorèmes maximaux. Nous verrons comment ces résultats sont liés aux propriétés géométriques de la suite (F_n) .

La première section est consacrée au cas le plus simple, le lemme de recouvrement de Vitali pour des boules de \mathbb{R}^d , dont résulte le théorème de Wiener : l'opérateur maximal associé aux boules centrées est de type faible $(1, 1)$. De nombreuses questions naturelles se posent dans ce contexte. En particulier, qu'obtient-on lorsque la mesure de Lebesgue sur \mathbb{R}^d est remplacée par une mesure de Radon quelconque ou lorsque les moyennes sont effectuées sur les boules non centrées ? Ces questions sont étudiées dans la section 2. Si la mesure conserve la même propriété de doublement de volume que la mesure de Lebesgue, à savoir que pour tout x , les boules de centre x et de rayon r et $2r$ ont des mesures comparables, l'opérateur maximal est encore de type faible $(1, 1)$, que les boules soient centrées ou non. Sans cette hypothèse de doublement de volume, nous montrons que l'opérateur maximal relatif aux boules centrées reste de type faible $(1, 1)$, à l'aide du lemme de recouvrement de Besicovitch. Dans le cas

des boules non centrées, cette propriété subsiste en dimension 1 mais est fautive par exemple pour la mesure gaussienne sur \mathbb{R}^2 . Nous examinons enfin un dernier exemple, l'opérateur maximal relatif à certains réseaux dyadiques de \mathbb{R}^d , dont l'étude se déduit du lemme maximal de Doob en théorie des martingales.

Dans la section 3, nous revenons au cas de l'opérateur A^* associé à une suite (F_n) de parties d'un groupe localement compact G . Lorsque la suite (F_n) est croissante et vérifie la condition de Tempel'man, analogue à la propriété de doublement de volume (plus précisément, $\sup_n |F_n^{-1}F_n|/|F_n| < +\infty$), l'opérateur A^* conserve la propriété d'être de type faible $(1, 1)$.

Un exemple dû à E. Lindenstrauss montre que certains groupes moyennables (à croissance exponentielle) ne possèdent pas de suites de Følner satisfaisant de surcroît à la condition de Tempel'man. En revanche, il existe toujours des suites de Følner qui ont la propriété plus faible d'être des suites tempérées (définition 3.4.1). L'opérateur maximal associé à toute suite tempérée est lui aussi de type faible $(1, 1)$. Nous terminons ce chapitre par deux démonstrations de ce résultat de Lindenstrauss. Encore une fois, elles sont basées sur des lemmes de recouvrement. La première démonstration, plus élémentaire, est due à Ornstein et Weiss. Elle permet de traiter le cas où G est discret, et la sélection des ensembles du recouvrement est effectuée de façon déterministe. La deuxième démonstration est celle de Lindenstrauss. Elle traite le cas général avec des arguments de nature probabiliste pour le choix de la sélection.

Nous n'avons parlé que de l'inégalité de type faible $(1, 1)$. Rappelons que, pour les opérateurs de moyenne que nous considérons, tout opérateur maximal de type faible $(1, 1)$ est de type fort (p, p) quel que soit $p > 1$ (voir la proposition 1.2.2).

3.1. Moyennes sur les boules de \mathbb{R}^d .

Dans cette section, nous nous plaçons dans \mathbb{R}^d muni de la mesure de Lebesgue. Nous nous intéressons à l'opérateur maximal relatif aux moyennes sur les boules euclidiennes centrées.

Nous notons $B(x, r)$ la boule euclidienne de centre x et de rayon r et β_r la mesure de Lebesgue normalisée sur la boule $B(0, r)$. Pour $\varphi \in L^1_{loc}(\mathbb{R}^d)$, rappelons que la fonction $\beta_r \cdot \varphi$ est définie par

$$\beta_r \cdot \varphi(x) = \frac{1}{|B(0, r)|} \int_{B(0, r)} \varphi(x + y) \, dy = \frac{1}{|B(x, r)|} \int_{B(x, r)} \varphi(y) \, dy.$$

Cette fonction est aussi le produit de convolution $\beta_r * \varphi$ de β_r par φ . Elle est continue ainsi que, pour tout $x \in \mathbb{R}^d$, la fonction $r \mapsto \beta_r \cdot \varphi(x)$.

Enfin nous introduisons la fonction maximale de Hardy-Littlewood

$$\varphi^*_\beta(x) = \sup_{r>0} |\beta_r \cdot \varphi(x)|,$$

qui est semi-continue inférieurement sur \mathbb{R}^d .

Les inégalités correspondant à cette fonction maximale ont été étudiées par Hardy et Littlewood en 1930 dans le cas $d = 1$ et par Wiener en 1939 dans le cas général. Nous rappelons ci-dessous le lemme de recouvrement de Vitali, et son utilisation pour établir l'inégalité maximale de type faible (1, 1).

Lemme 3.1.1 (Vitali). — Soit E un sous-ensemble d'un espace métrique recouvert par une famille finie $(B(x_i, r_i))_{1 \leq i \leq N}$, de boules de centre x_i et de rayon r_i . On peut extraire de cette famille une sous-famille $(B(x_{i_k}, r_{i_k}))$, $1 \leq k \leq n$, de boules deux à deux disjointes telle que

$$E \subset \bigcup_{k=1}^n B(x_{i_k}, 3r_{i_k}).$$

Démonstration. — Posons $B_i = B(x_i, r_i)$ et supposons les boules numérotées de façon que $r_1 \leq r_2 \leq \dots \leq r_N$. Commençons avec $i_1 = N$, puis, par récurrence, et tant que c'est possible, choisissons pour i_{k+1} le plus grand entier strictement inférieur à i_k tel que $B_{i_{k+1}}$ ne rencontre pas les boules B_{i_1}, \dots, B_{i_k} déjà sélectionnées. Notons n l'étape où s'arrête le processus. Alors toute boule $B = B(x, r)$ de la famille initiale rencontre une boule B_{i_k} de la sous-famille choisie, avec $r \leq r_{i_k}$, d'où $B \subset B(x_{i_k}, 3r_{i_k})$.

Il en résulte que $E \subset \bigcup_{k=1}^n B(x_{i_k}, 3r_{i_k})$. \square

Nous pouvons maintenant démontrer l'inégalité maximale de type faible (1, 1).

Théorème 3.1.2. — Pour toute fonction $\varphi \in L^1(\mathbb{R}^d)$ et tout $\alpha > 0$, on a

$$(3.1.1) \quad |\{\varphi_\beta^* > \alpha\}| \leq 3^d \frac{\|\varphi\|_1}{\alpha}.$$

Ainsi, l'opérateur maximal centré relatif aux boules euclidiennes est de type faible (1, 1).

Démonstration. — Puisque la mesure de Lebesgue est régulière, il suffit de montrer que pour toute partie compacte K contenue dans l'ensemble $\{\varphi_\beta^* > \alpha\}$, on a $|K| \leq 3^d \frac{\|\varphi\|_1}{\alpha}$. Pour tout $x \in K$, il existe un rayon $r_x > 0$ tel que

$$|B(x, r_x)| \leq \frac{1}{\alpha} \left| \int_{B(x, r_x)} \varphi(y) dy \right|.$$

Puisque K est compact, nous pouvons le recouvrir par un nombre fini de ces boules $B(x, r_x)$ et, grâce au lemme de Vitali, il existe une sous-famille B_1, \dots, B_n formée de boules deux à deux disjointes et telle que $|K| \leq 3^d \sum_{k=1}^n |B_k|$. Il en résulte que

$$|K| \leq 3^d \sum_{k=1}^n \frac{1}{\alpha} \int_{B_k} |\varphi(y)| dy \leq 3^d \frac{\|\varphi\|_1}{\alpha}.$$

\square

Le théorème de Wiener est alors une conséquence des résultats obtenus dans le chapitre 2, puisque la famille $(\beta_r)_{r>0}$ est continue et asymptotiquement invariante (voir le théorème 2.3.3 et les remarques 2.3.2 et 2.3.4).

Théorème 3.1.3 (Wiener). — *Supposons que le groupe \mathbb{R}^d agisse sur un espace mesuré (X, \mathcal{B}, m) en laissant invariante la mesure σ -finie m . Soient $p \in [1, +\infty[$, et $f \in L^p(X, m)$.*

- (i) *Pour $p > 1$, la famille $\beta_r \cdot f$ tend presque partout et dans L^p , lorsque r tend vers l'infini, vers une fonction \mathbb{R}^d -invariante.*
- (ii) *Lorsque $p = 1$, la famille $\beta_r \cdot f$ converge presque partout vers une fonction \mathbb{R}^d -invariante. Si de plus la mesure m est finie, la convergence a lieu aussi dans L^1 . Lorsque m est une mesure de probabilité, la limite est l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$ de f sachant la tribu \mathcal{I} des invariants.*

Remarque 3.1.4. — Pour toute fonction φ continue sur \mathbb{R}^d et tout $x \in \mathbb{R}^d$, on a

$$(3.1.2) \quad \varphi(x) = \lim_{r \rightarrow 0} \frac{1}{|B(x, r)|} \int_{B(x, r)} \varphi(y) \, dy.$$

En utilisant le principe de Banach (voir théorème 1.2.1) et l'inégalité maximale faible relative aux moyennes sur les boules, on obtient alors le théorème de différentiation de Lebesgue : pour tout $\varphi \in L^1(\mathbb{R}^d)$ et pour presque tout $x \in \mathbb{R}^d$, la propriété (3.1.2) reste vraie.

Remarque 3.1.5. — On peut aussi considérer la fonction maximale non centrée définie par

$$\varphi_{\beta}^{**}(x) = \sup_{B \ni x} \frac{1}{|B|} \left| \int_B \varphi(y) \, dy \right|,$$

où B varie dans l'ensemble des boules contenant x . La comparaison des volumes des boules de même centre permet de montrer que

$$\varphi_{\beta}^* \leq \varphi_{\beta}^{**} \leq 2^d \varphi_{\beta}^*$$

si $\varphi \geq 0$. L'opérateur maximal non centré relatif aux boules euclidiennes est donc, lui aussi, de type faible (1, 1).

Remarque 3.1.6. — La recherche des meilleures constantes pour les inégalités maximales est l'objet de nombreux travaux.

Dans le cas centré, et si $d = 1$, la meilleure constante C pour l'inégalité maximale faible (1, 1) a été obtenue par A. D. Melas [Mel03]. C'est la plus grande racine de l'équation $12x^2 - 22x + 5$, c'est-à-dire $C = \frac{11 + \sqrt{61}}{12} = 1,5675208\dots$. Le problème est non résolu pour $d > 1$.

Pour l'inégalité forte de type (p, p) avec $p > 1$, la détermination de la meilleure constante $C_{p,d}$ est un problème ouvert en toute dimension $d \geq 1$ dans le cas centré. Dans cette direction, on sait toutefois qu'il existe une constante C_p (indépendante de d) telle que $\|\varphi_{\beta}^*\|_p \leq C_p \|\varphi\|_p$ pour tout $\varphi \in L^p(\mathbb{R}^d)$ (voir [SS83]).

Dans le cas non centré, si $p > 1$ et $d \geq 1$, la meilleure constante $C_{p,d}$ telle que $\|\varphi_{\beta}^{**}\|_p \leq C_{p,d}\|\varphi\|_p$ pour tout $\varphi \in L^p(\mathbb{R}^d)$ a été déterminée dans [GMS97]. Contrairement au cas centré, elle croît exponentiellement quand d tend vers l'infini.

3.2. Mesures doublantes et non doublantes

Dans cette section nous continuons l'étude de l'opérateur maximal associé à des moyennes sur des boules. Nous nous plaçons dans le cadre général d'un espace métrique (X, d) localement compact, muni d'une mesure de Radon positive m sur les boréliens de X , c'est-à-dire une mesure borélienne finie sur les ensembles compacts et intérieurement régulière, ce qui signifie que la mesure de tout borélien E est égale au supremum des mesures des sous-ensembles compacts de E .

Nous verrons que si m possède, comme la mesure de Lebesgue, la propriété de *doublement du volume*, alors les résultats obtenus en 3.1 se transposent grâce au lemme de Vitali.

On étudiera ensuite le cas d'une mesure de Radon positive m arbitraire sur $X = \mathbb{R}^d$. Le lemme de recouvrement de Besicovitch permet d'établir une inégalité faible pour l'opérateur maximal centré associé à m mais, en dimension $d \geq 2$, même si la mesure m est absolument continue par rapport à la mesure de Lebesgue, l'opérateur maximal non centré n'est plus forcément de type faible $(1, 1)$.

On désigne par $B(x, r)$ la boule ouverte de centre x et de rayon r . L'opérateur maximal (centré) associé à m est défini, pour x appartenant au support $\text{Supp}(m)$ de m , par

$$A^* f(x) = \sup_{r>0} \frac{1}{m(B(x, r))} \left| \int_{B(x, r)} f \, dm \right|$$

où f est une fonction borélienne intégrable sur X , ou de manière plus générale intégrable sur toute boule. Les valeurs de $A^* f(x)$ en dehors du support de m ne sont pas significatives car on n'étudie en général cette fonction que m -presque partout.

On considérera également la fonction maximale non centrée définie sur le support de m par

$$A^{**} f(x) = \sup_{B \ni x} \frac{1}{m(B)} \left| \int_B f \, dm \right|,$$

où B varie dans l'ensemble des boules ouvertes contenant le point x .

La fonction $A^{**} f$ est semi-continue inférieurement. D'autre part la fonction maximale centrée est mesurable. En effet, la propriété de continuité des mesures implique que pour tout $x \in \text{Supp}(m)$, la fonction $r \mapsto \frac{1}{m(B(x, r))} \left| \int_{B(x, r)} f \, dm \right|$ est continue à gauche sur $]0, \infty[$. Le supremum dans la définition de $A^* f(x)$ peut donc être restreint aux rationnels r . Il en découle que $A^* f$ est mesurable.

3.2.1. Mesures doublantes sur un espace métrique. —

Définition 3.2.1. — On dit que la mesure de Radon positive m possède la *propriété de doublement* s'il existe $C > 0$ tel que pour tout x et tout r ,

$$(3.2.1) \quad m(B(x, 2r)) \leq Cm(B(x, r)).$$

Cela signifie que les boules $B(x, r)$ et $B(x, 2r)$ ont une mesure comparable. Cette condition entraîne immédiatement que pour tout $\gamma > 1$,

$$m(B(x, \gamma r)) \leq C(\gamma)m(B(x, r)).$$

Plus généralement deux boules de rayons comparables et d'intersection non vide ont une mesure comparable. Ceci implique qu'une mesure doublante ne charge aucune hypersurface (voir [Ste93]). Ceci entraîne aussi que les opérateurs maximaux centrés et non centrés sont comparables : si $f \geq 0$,

$$A^*f \leq A^{**}f \leq CA^*f.$$

On est ici dans une situation un peu plus générale que $X = \mathbb{R}^d$ muni de la mesure de Lebesgue m , pour laquelle $m(B(x, 2r)) = 2^d m(B(x, r))$.

De nombreux résultats établis initialement pour des opérateurs associés à la mesure de Lebesgue ont été ensuite généralisés à des mesures doublantes. L'une des raisons est que le lemme de recouvrement de Vitali est très adapté à l'étude de ce genre de mesure. L'étude des opérateurs maximaux peut se faire dans ce cadre. Pour l'étude d'autres opérateurs classiques associés à des mesures doublantes, voir [Ste93], [Ste70b].

Théorème 3.2.2. — *Si la mesure de Radon m a la propriété de doublement, les opérateurs maximaux centrés et non centrés sont de type faible (1, 1).*

Démonstration. — La démonstration est identique à celle du théorème 3.1.2, qui concerne le cas particulier où $X = \mathbb{R}^d$ et m est la mesure de Lebesgue. \square

La propriété de doublement est utilisée deux fois dans cette démonstration : pour comparer les deux fonctions maximales, et pour appliquer le lemme de recouvrement de Vitali. Lorsque la mesure ne possède plus cette propriété, nous allons voir que seul l'opérateur centré reste borné.

3.2.2. Mesures non doublantes sur \mathbb{R}^d . — Ici la mesure m est une mesure de Radon positive sur $X = \mathbb{R}^d$. On ne suppose plus que la propriété de doublement est satisfaite.

3.2.2.1. L'opérateur maximal centré. — Le lemme de recouvrement de Besicovitch est très adapté à ce type de mesure. Pour des variantes de ce résultat, voir [Bes45, Bes46, Bes47] et [Mor47]. Nous énonçons et démontrons ce lemme.

Théorème 3.2.3 (Lemme de recouvrement de Besicovitch)

Soit E un ensemble borné de \mathbb{R}^d . On suppose donnée, pour tout $x \in E$, une boule $B(x) = B(x, r_x)$ centrée en x et de rayon $r_x > 0$. Alors on peut trouver une partie dénombrable D dans E telle que

- (i) la famille de boules $(B(x))_{x \in D}$ recouvre E ;
- (ii) tout point de \mathbb{R}^d appartient à moins de $d \times 10^d$ boules de la famille $(B(x))_{x \in D}$.

Remarque 3.2.4. — A la différence du lemme de Vitali, on n'impose pas aux boules extraites d'être disjointes. Par contre, la multiplicité du recouvrement est inférieure à $d \times 10^d$, et cette constante est indépendante du diamètre de E . Ce lemme est adapté aux mesures non doublantes car on ne change pas le rayon des boules.

Ce lemme est énoncé pour des boules euclidiennes. On peut prendre aussi des cubes, ou des boules pour d'autres normes. On peut même prendre des boules associées à une famille infinie de normes, à condition que les excentricités soient contrôlées. Ici les boules sont choisies ouvertes, bien que ce ne soit pas nécessaire.

Avant de démontrer le lemme de recouvrement, commençons par l'appliquer à l'étude de la fonction maximale.

Corollaire 3.2.5. — Soit m une mesure de Radon sur \mathbb{R}^d . Alors l'opérateur maximal centré associé à m est de type faible $(1, 1)$.

Démonstration. — Nous reprenons la démonstration du théorème 3.1.2, dont nous conservons les notations. Au lieu d'utiliser le lemme de recouvrement de Vitali, nous utilisons celui de Besicovitch. On extrait une sous-famille de boules $B(x_i, r_i)$ recouvrant l'ensemble K et de multiplicité inférieure à $d \times 10^d$. Le calcul suivant permet de conclure :

$$\begin{aligned} m(K) &\leq \sum_i m(B(x_i, r_i)) \leq \sum_i \frac{1}{\alpha} \int_{B(x_i, r_i)} |f| dm \\ &= \frac{1}{\alpha} \int_{\mathbb{R}^d} \left(\sum_i \mathbf{1}_{B(x_i, r_i)} \right) |f| dm \\ &\leq \frac{d \times 10^d}{\alpha} \int_{\mathbb{R}^d} |f| dm. \end{aligned}$$

On a donc l'inégalité faible. □

La partie la plus difficile est la démonstration du lemme de recouvrement.

Démonstration. — Soit M le diamètre de E . S'il existe un point $x \in E$ tel que $r_x > M$, on conclut directement puisque l'unique boule correspondante recouvre E . On suppose donc dans la suite que les rayons r_x sont bornés par M .

Nous allons choisir l'ensemble D comme une réunion dénombrable d'ensembles finis D_k construits par récurrence, de telle sorte que D_k soit inclus dans l'ensemble E_k des

$x \in E$ pour lesquels

$$(3.2.2) \quad \frac{M}{2^k} < r_x \leq \frac{M}{2^{k-1}}.$$

Étape 1 : On choisit pour D_1 une famille maximale $x_1^1, \dots, x_{N_1}^1$ dans E_1 telle que chaque centre x_j^1 n'appartienne à aucune des boules $B(x_i^1)$ pour $i < j$. Cet ensemble D_1 est fini car il est borné et ses éléments sont deux à deux distants d'au moins $M/2$.

Étape k : Supposant ainsi choisis D_1, \dots, D_{k-1} , on sélectionne une partie maximale $D_k = \{x_1^k, \dots, x_{N_k}^k\}$ dans E_k telle que chaque centre x_j^k n'appartient ni aux boules $B(x_i^k)$ pour $i < j$, ni aux boules $B(x_i^l)$ des étapes $l < k$. Comme dans l'étape 1, on s'aperçoit que l'ensemble D_k est fini.

On pose $D = \cup_{k \geq 1} D_k$. Par construction, la famille de boules $(B(x))_{x \in D}$ a la propriété suivante : si B est une boule sélectionnée avant une boule B' alors le centre de B' n'appartient pas à B .

Montrons que l'ensemble dénombrable D possède les propriétés requises. Vérifions tout d'abord que la famille de boules $(B(x))_{x \in D}$ recouvre E . Soit x un élément de E_k qui n'appartient pas à la réunion de ces boules. On dispose donc d'une boule $B(x)$ de l'étape k dont le centre n'est dans aucune des boules sélectionnées aux étapes 1 à k . On aurait donc pu rajouter une boule à l'étape k , ce qui contredit la maximalité de D_k .

Il nous reste à montrer que la multiplicité du recouvrement est inférieure à $d \times 10^d$. Soit $y \in \mathbb{R}^d$. Pour tout entier $k \geq 1$, soit $J_k = J_k(y)$ l'ensemble des $j \in \{1, \dots, N_k\}$ tels que $y \in B(x_j^k)$. Observons tout d'abord que le cardinal de J_k est majoré par 5^d . C'est une conséquence directe du fait que les boules $B(x_j^k, \frac{M}{2^{k+1}})$, pour $j \in J_k$, sont deux à deux disjointes et contenues dans la boule $B(y, \frac{M}{2^{k+1}} + \frac{M}{2^{k-1}})$.

Pour conclure, il nous suffit de démontrer que l'ensemble K des entiers $k \geq 1$ tels que J_k est non vide contient au plus $d \times 2^d$ éléments. Considérons deux boules $B(x_i^k)$ et $B(x_j^l)$, sélectionnées aux étapes $k < l$, telles que y appartienne à leur intersection. On a ainsi $k, l \in K$, $i \in J_k$ et $j \in J_l$. Le rayon r de la boule $B(x_i^k)$ est strictement plus grand que celui de la boule $B(x_j^l)$. En posant $A = y$, $B = x_i^k$ et $C = x_j^l$, on voit que dans le triangle ABC , on a $AB < r$ et $AC < r$. D'autre part, on a $BC > r$ puisque la boule de centre C a été sélectionnée après celle de centre B . Ces relations entraînent que dans le triangle ABC , la mesure de l'angle de sommet A est supérieure à $\frac{\pi}{3}$.

Pour avoir une majoration grossière du cardinal de K , il suffit de compter le nombre maximal de calottes sphériques disjointes, d'angle au centre $\frac{\pi}{3}$, qu'on peut trouver sur la sphère unité de \mathbb{R}^d . Celui-ci est inférieur au rapport entre la surface de la sphère et la surface d'une telle calotte sphérique. La majoration indiquée est obtenue par une minoration grossière de la surface de la calotte par la surface de sa projetée sur un hyperplan perpendiculaire à son axe de symétrie. La surface de la calotte est donc minorée par le volume en dimension $d - 1$ d'une boule de rayon $1/2$. D'autre part la

surface de la sphère unité de \mathbb{R}^d est majorée par $2d$ fois le volume de la boule unité en dimension $d - 1$: en effet, la boule unité en dimension d est contenue dans un cylindre de révolution de hauteur 2 ayant pour base la boule unité en dimension $d - 1$. On trouve finalement que le cardinal de K est majoré par $d \times 2^d$. \square

3.2.2.2. L'opérateur maximal non centré. — Il s'avère que lorsqu'on décentre les boules, l'opérateur maximal n'est plus borné en général, même pour des mesures très régulières. Un contre-exemple a été donné par Sjögren en 1983 [Sjö83] en dimension $d \geq 2$ pour la mesure gaussienne standard. Dans ce cas l'opérateur maximal n'est pas de type faible $(1, 1)$.

Théorème 3.2.6. — *Soit la mesure $dm(x, y) = e^{-(x^2+y^2)} dx dy$ sur \mathbb{R}^2 . L'opérateur maximal non centré associé à m n'est pas de type faible $(1, 1)$.*

Démonstration. — On rappelle que l'opérateur maximal non centré est défini par

$$A^{**}f(x, y) = \sup_{B \ni (x, y)} \frac{\int_B f dm}{m(B)}.$$

Soit f la fonction indicatrice de la boule euclidienne centrée en $(0, a + 1)$, avec a grand, et de rayon $\delta < 1/a$. Cette fonction approche la masse de Dirac en $(0, a + 1)$. Pour $|s| \leq 1$, soit B_s la boule fermée de rayon 1 centrée au point $(s, a + 1)$. Si $(x, y) \in B_s$, on a $y \geq a + \frac{(x-s)^2}{2}$. On en déduit une estimation de la mesure de B_s :

$$\begin{aligned} m(B_s) &\leq \int_{-1+s}^{1+s} \int_{y \geq a + \frac{(x-s)^2}{2}} e^{-y^2} dx dy = \int_{-1}^1 \int_{y \geq a + \frac{x^2}{2}} e^{-y^2} dx dy \\ &\leq \int_{-1}^1 \int_{y \geq a + \frac{x^2}{2}} \frac{y}{a} e^{-y^2} dx dy = \frac{1}{2} \int_{-1}^1 \frac{1}{a} e^{-(a + \frac{1}{2}x^2)^2} dx \\ &\leq \frac{1}{2} \frac{e^{-a^2}}{a} \int_{-1}^1 e^{-ax^2} dx \leq \frac{\sqrt{\pi}}{2} \frac{e^{-a^2}}{a\sqrt{a}} \leq \frac{e^{-a^2}}{a\sqrt{a}}. \end{aligned}$$

Pour tout $s \in [-1, 1]$, l'intersection de la boule B_s avec le support de f a une mesure de Lebesgue comparable à δ^2 . De plus, sur le support de f la quantité $e^{-x^2-y^2}$ est comparable à $e^{-(a+1)^2}$ car $\delta < a^{-1}$. Il existe donc une constante assez petite C_1 , indépendante de s et a , telle que $\int_{B_s} f dm > C_1 \delta^2 e^{-(a+1)^2}$ et une constante C_2 , indépendante de a , telle que $\int_{\mathbb{R}^2} f dm \leq C_2 \delta^2 e^{-(a+1)^2}$.

Pour tout $(x, y) \in [-1, 1] \times [a, a + 2]$, on a

$$(3.2.3) \quad A^{**}f(x, y) > C_1 a \sqrt{a} e^{a^2} \delta^2 e^{-(a+1)^2}.$$

Par ailleurs, on peut trouver une constante C_3 telle que

$$\begin{aligned} m([-1, 1] \times [a, a+2]) &= \int_{-1}^1 \int_a^{a+2} e^{-(x^2+y^2)} dx dy \\ &\geq \frac{2}{e} \int_a^{a+2} \frac{y}{a+2} e^{-y^2} dy \geq \frac{C_3}{a} e^{-a^2}. \end{aligned}$$

Posons $\alpha = C_1 a \sqrt{a} e^{a^2} \delta^2 e^{-(a+1)^2}$ et $E_\alpha = \{A^{**} f > \alpha\}$.

On voit d'une part que

$$\frac{\int_{\mathbb{R}^2} f dm}{\alpha} \leq C_2 / C_1 \frac{1}{a \sqrt{a}} e^{-a^2},$$

et d'autre part que

$$m(E_\alpha) \geq m([-1, 1] \times [a, a+2]) \geq \frac{C_2}{a} e^{-a^2}.$$

L'opérateur A^{**} n'est donc pas de type faible (1,1). \square

L'opérateur maximal non centré est toujours de type faible (1,1) en dimension 1. Nous reproduisons ici la démonstration de [Sjö83].

Théorème 3.2.7. — *Sur \mathbb{R} , l'opérateur maximal non centré associé à une mesure de Radon est de type faible (1,1).*

Démonstration. — Soit $f \in L^1(\mathbb{R}, m)$. La fonction maximale non centrée est

$$A^{**} f(x) = \sup_{I \ni x} \frac{1}{m(I)} \left| \int_I f dm \right|.$$

Ici I est un intervalle. Par convention $\frac{1}{m(I)} \int_I f dm = 0$ si $m(I) = 0$. On peut donc en fait restreindre le supremum aux intervalles de mesure positive. Fixons $\alpha > 0$ et soit \mathcal{S} la famille des intervalles I tels que

$$m(I) < \frac{1}{\alpha} \left| \int_I f dm \right|.$$

Il est clair que

$$E_\alpha = \{A^{**} f > \alpha\} = \bigcup_{I \in \mathcal{S}} I.$$

Soit $K \subset E_\alpha$, avec K compact. Il est recouvert par une sous-famille finie $\mathcal{S}' \subset \mathcal{S}$. Nous allons extraire de \mathcal{S}' une sous-famille d'intervalles disjoints de la façon suivante. Le premier intervalle I_1 est choisi de mesure maximale. Ensuite par récurrence on choisit I_k de mesure maximale parmi les intervalles de \mathcal{S}' disjoints de I_1, \dots, I_{k-1} . Le processus s'arrête, donc tout intervalle $I =]a, b[$ de \mathcal{S}' rencontre l'un des I_i .

Nous allons avoir besoin d'agrandir chacun des $I_i =]a_i, b_i[$. Pour tout i , on note $I_i^* =]a_i^*, b_i^*[$ le plus grand intervalle ouvert tel que $m(]a_i^*, a_i]) \leq m(I_i)$ et $m(]b_i, b_i^*]) \leq m(I_i)$. Soit maintenant $I =]a, b[$ un intervalle de \mathcal{S}' et soit i_0 le plus petit indice i

tel que $I \cap I_i \neq \emptyset$. Alors $m(]a, a_{i_0}]) \leq m(I) \leq m(I_{i_0})$, ce qui entraîne que $a \geq a_{i_0}^*$. Par convention, $]a, a_{i_0}] = \emptyset$ si $a \geq a_{i_0}$. De la même manière, $b \leq b_{i_0}^*$. Finalement on obtient $I \subset I_{i_0}^*$. Les I_i^* recouvrent donc l'ensemble K . Par définition de I_i^* , on a $m(I_i^*) \leq 3m(I_i)$. Il s'ensuit que

$$m(K) \leq \sum_i 3m(I_i) \leq \frac{3}{\alpha} \int |f| dm.$$

□

Remarque 3.2.8. — Dans \mathbb{R}^d , l'opérateur maximal non centré associé à des cubes parallèles aux axes est de type fort (p, p) pour $p > 1$, lorsque la mesure est un produit tensoriel, comme la mesure gaussienne par exemple. Ceci découle aisément du théorème 3.2.7 et du lemme d'interpolation de Marcinkiewicz.

Nous allons donner à présent un exemple de mesure m pour lequel l'opérateur maximal non centré n'est de type fort (p, p) pour aucun $p > 1$. On peut construire dans \mathbb{R}^d , $d \geq 2$, une suite de boules euclidiennes ouvertes B_i , $i \geq 0$, contenant le point 0, et telles que pour tout $i \geq 1$ on ait un point $x_i \in B_i$ qui n'appartienne à aucune des boules B_j , $j \neq i$. Si $\lim_{i \rightarrow +\infty} \|x_i\| = +\infty$, la mesure $m = \delta_0 + \sum_{i \geq 1} \delta_{x_i}$ est une mesure de Radon. Soit $f = \mathbf{1}_{B_0}$. Pour tout $i \geq 1$ on a

$$A^{**}f(x_i) \geq \frac{m(B_0 \cap B_i)}{m(B_i)} = \frac{1}{2}.$$

Il s'ensuit que $\|A^{**}f\|_{L^p(\mathbb{R}^d, m)} = +\infty$. Or $\|f\|_{L^p(\mathbb{R}^d, m)} = 1 < \infty$ et donc A^{**} n'est pas de type fort (p, p) .

3.2.3. Fonction maximale dyadique et martingales. — Le dernier type de fonction maximale auquel nous allons nous intéresser est la fonction maximale dyadique. Nous allons voir que l'opérateur maximal dyadique est toujours de type faible $(1, 1)$.

Soit $(\mathcal{D}_n)_{n \in \mathbb{Z}}$ une suite de partitions de \mathbb{R}^d . On dit que c'est une *suite de partitions dyadiques* si pour tout n , la famille \mathcal{D}_n est formée de cubes dyadiques d'arêtes 2^{-n} et si chaque cube de \mathcal{D}_n (encore appelé *cube de la n -ième génération*) est obtenu à partir d'un cube Q de la génération précédente lorsqu'on découpe celui-ci en 2^d cubes dyadiques d'arête moitié. Ces 2^d cubes dyadiques de la n -ième génération sont alors appelés les *fils du cube Q* .

Il est clair que la famille \mathcal{F}_n constituée par les réunions d'éléments de \mathcal{D}_n est une tribu. De plus $\{\mathcal{F}_n\}_{n \in \mathbb{Z}}$ est une filtration croissante de tribus, engendrant la tribu des boréliens.

Quitte à effectuer une translation et une dilatation, on peut supposer que les cubes de \mathcal{D}_0 sont les translatés par \mathbb{Z}^d de $[0, 1]^d$ (si on choisit qu'ils soient "ouverts à droite"). Les cubes de la génération $n > 0$ sont alors déterminés de manière unique : ceux de \mathcal{D}_1 sont les fils de ceux d'ordre 0, c'est-à-dire les cubes obtenus en divisant chaque cube

de \mathcal{D}_0 en 2^d cubes de longueur $\frac{1}{2}$. La famille \mathcal{D}_1 est donc formée par les translatés par $\frac{1}{2}\mathbb{Z}^d$ de $[0, \frac{1}{2}]^d$. On définit de manière similaire les générations \mathcal{D}_n pour $n \geq 0$.

A l'opposé, il n'y a pas de manière canonique de définir les cubes des générations négatives. On commence par choisir un cube Q de longueur 2 contenant $[0, 1]^d$, de sorte que $[0, 1]^d$ soit l'un de ses fils. On a 2^d choix pour Q . Par exemple $Q = [-1, 1]^d$ ou $Q = [0, 2]^d$ conviennent. Les cubes de \mathcal{D}_{-1} seront les translatés par $2\mathbb{Z}^d$ de Q . On construit de la même façon les cubes d'ordre -2 à partir de ceux d'ordre -1 . En général, toutefois, on choisit \mathcal{D}_n de sorte que l'un des cubes de cette génération ait pour sommet 0.

On note \mathcal{D} la réunion des partitions \mathcal{D}_n ; on l'appelle *réseau dyadique*. La propriété fondamentale d'un réseau dyadique est que si deux de ses cubes Q, R ne sont pas disjoints, alors l'un est inclus dans l'autre. Cette particularité est déterminante dans les applications, par exemple dans l'étude de la fonction maximale dyadique.

Soit m une mesure de Radon positive sur \mathbb{R}^d . La fonction maximale dyadique est définie par

$$A_{dy}^* f(x) = \sup_{Q \ni x, Q \in \mathcal{D}} \frac{1}{m(Q)} \left| \int_Q f \, dm \right|.$$

Le supremum est pris sur les cubes $Q \in \mathcal{D}$ tels que $x \in Q$.

Théorème 3.2.9. — *L'opérateur maximal dyadique est de type faible (1, 1).*

Ce résultat est bien connu car c'est un cas particulier du lemme maximal de Doob dans la théorie des martingales (voir [Nev72]).

Démonstration. — Supposons tout d'abord que m est une mesure finie. Quitte à la normaliser, on peut supposer que c'est une mesure de probabilité. Pour établir l'inégalité, il suffit de considérer une fonction $f \in L^1(\mathbb{R}^d, m)$ positive.

Nous changeons de notations en nous plaçant maintenant dans un cadre probabiliste. La mesure m définit une probabilité \mathbb{P} . La fonction $X = f$ est une variable aléatoire, et la suite

$$X_n = \mathbb{E}(X | \mathcal{F}_n)$$

est une martingale positive.

Rappelons le lemme maximal de Doob :

Lemme 3.2.10. — *Soient $(X_n)_{n \in \mathbb{Z}}$ une martingale positive et $N \in \mathbb{Z}$. Alors, pour $\alpha > 0$,*

$$\mathbb{P} \left(\sup_{k \leq N} X_k > \alpha \right) \leq \frac{1}{\alpha} \int_{\{\sup_{k \leq N} X_k > \alpha\}} X_N \, d\mathbb{P}$$

Démonstration. — Pour démontrer cela, on introduit le temps d'arrêt

$$\tau = \inf \{k \in \mathbb{Z}; X_k > \alpha\} \in [-\infty, +\infty],$$

de sorte que $\{\tau \leq N\} = \{\sup_{k \leq N} X_k > \alpha\}$. Le lemme résulte alors du calcul suivant :

$$\begin{aligned}
\mathbb{E}(X_N \mathbf{1}_{\tau \leq N}) &= \sum_{k \leq N} \mathbb{E}(X_N \mathbf{1}_{\tau=k}) \\
&= \sum_{k \leq N} \mathbb{E}(\mathbb{E}(X_N \mathbf{1}_{\tau=k} | \mathcal{F}_k)) \\
&= \sum_{k \leq N} \mathbb{E}(\mathbb{E}(X_N | \mathcal{F}_k) \mathbf{1}_{\tau=k}) \\
&= \sum_{k \leq N} \mathbb{E}(X_k \mathbf{1}_{\tau=k}) \geq \sum_{k \leq N} \mathbb{E}(\alpha \mathbf{1}_{\tau=k}) \\
&= \alpha \mathbb{P}(\tau \leq N).
\end{aligned}$$

□

Terminons maintenant la démonstration du théorème 3.2.9. On a

$$\int_{\{\sup_{k \leq N} X_k > \alpha\}} X_N \, d\mathbb{P} \leq \mathbb{E}(X_N) = \mathbb{E}(X),$$

et on déduit du lemme que

$$(3.2.4) \quad \mathbb{P}(\sup_{n \in \mathbb{Z}} X_n > \alpha) \leq \frac{1}{\alpha} \mathbb{E}(X).$$

Pour notre choix particulier de tribus, l'espérance conditionnelle est

$$\mathbb{E}(Z | \mathcal{F}_n)(x) = \sum_{l(Q)=2^{-n}} \mathbf{1}_Q(x) \frac{1}{m(Q)} \int_Q Z \, dm,$$

où $l(Q)$ est la longueur du cube dyadique $Q \in \mathcal{D}$. La fonction maximale dyadique est donc égale au supremum des X_n . L'inégalité (3.2.4) est donc l'inégalité faible cherchée :

$$m(A_{dy}^* f > \alpha) \leq \frac{1}{\alpha} \int f \, dm.$$

La constante étant indépendante de la mesure finie m , le cas d'une mesure non finie se traite par troncature de m . □

Remarquons que le théorème de convergence des martingales régulières nous assure que la suite (X_n) définie ci-dessus converge presque sûrement et dans L^1 vers $X = f$ (voir [Nev72, prop. IV-2-3]).

Remarquons également qu'il existe un lien direct entre les inégalités maximales pour les moyennes sur les boules de \mathbb{R}^d ou \mathbb{Z}^d et les inégalités maximales pour les moyennes sur les cubes dyadiques. On peut montrer, mais nous ne le ferons pas ici, que les théorèmes 3.2.9 et 3.2.2 se déduisent naturellement l'un de l'autre. Ce lien entre théorie ergodique et théorie des martingales a été brillamment utilisé par Bourgain dans son étude des théorèmes ergodiques le long des suites arithmétiques [Bou89]. On trouvera une description claire de ce principe et des applications intéressantes pour la théorie ergodique des actions de \mathbb{Z}^d dans [JRW03].

3.3. Suites doublantes d'un groupe

Nous nous plaçons ici dans le cadre général d'un groupe localement compact G à base dénombrable d'ouverts, muni de sa mesure de Haar à gauche. La suite des boules euclidiennes de centre 0 et de rayon n sur \mathbb{R}^d sera alors remplacée par une suite (F_n) de parties boréliennes relativement compactes de G , de mesure de Haar $|F_n|$ non nulle. Nous considérons la suite (A_n) des opérateurs de moyenne sur F_n . Si φ est une fonction localement intégrable sur G , nous avons donc

$$A_n\varphi(x) = \frac{1}{|F_n|} \int_{F_n} \varphi(gx) \, dg.$$

Notons Δ la fonction modulaire du groupe G . Comme

$$\int_{F_n} \varphi(gx) \, dg = \frac{1}{\Delta(x)} \int_{F_n x} \varphi(g) \, dg,$$

on voit que

$$(3.3.1) \quad A_n\varphi(x) = \frac{1}{|F_n x|} \int_{F_n x} \varphi(g) \, dg.$$

L'opérateur maximal est défini par

$$A^*\varphi(x) = \sup_n |A_n\varphi(x)|.$$

Nous allons voir que l'inégalité de type faible (1, 1) obtenue dans le théorème 3.1.2 pour les moyennes sur les boules centrées euclidiennes de \mathbb{R}^d reste valable dans notre situation, à condition que la suite (F_n) soit croissante et conserve la propriété de "doublement de volume" suivante, introduite par Tempel'man [Tem67] :

Définition 3.3.1. — On dit que la suite (F_n) de parties boréliennes relativement compactes de G satisfait à la *condition de Tempel'man* s'il existe une constante $c > 0$ telle qu'on ait :

$$\forall n \geq 1, \quad |F_n^{-1}F_n| \leq c|F_n|. \quad (T)$$

Cette condition est réalisée pour la suite (B_n) des boules euclidiennes de centre 0 et de rayon n dans \mathbb{R}^d puisque $B_n^{-1}B_n = B_{2n}$ et $|B_{2n}| = 2^d|B_n|$. Le théorème ci-dessous appliqué à cette situation donne donc l'inégalité (3.1.1) avec la constante $c = 2^d$ au lieu de 3^d .

Théorème 3.3.2. — Soit (F_n) une suite croissante de parties boréliennes relativement compactes de G satisfaisant à la condition (T). On suppose que les F_n sont de mesure de Haar non nulle. Soit $\varphi \in L^1(G)$. Alors, pour tout $\alpha > 0$, on a

$$|\{A^*\varphi > \alpha\}| \leq \frac{c}{\alpha} \|\varphi\|_1.$$

Démonstration. — Puisque $A^*\varphi \leq A^*|\varphi|$, on peut supposer que φ est positive ou nulle.

Pour tout N entier, posons $A_N^*\varphi(x) = \max_{1 \leq k \leq N} A_k\varphi(x)$. On commence par remarquer qu'il suffit d'établir l'inégalité

$$|\{A_N^*\varphi > \alpha\}| \leq \frac{c}{\alpha} \int_G \varphi(g) dg$$

pour tout entier $N \geq 1$. De plus, compte tenu de la régularité de la mesure de Haar, il suffit de démontrer que

$$(3.3.2) \quad |K| \leq \frac{c}{\alpha} \int_G \varphi(g) dg$$

pour toute partie compacte K contenue dans $\{A_N^*\varphi > \alpha\}$.

Fixons un tel K . Pour tout $x \in K$, il existe un entier $i(x)$ avec $1 \leq i(x) \leq N$ et

$$(3.3.3) \quad \alpha|F_{i(x)}x| \leq \int_{F_{i(x)}x} \varphi(g) dg.$$

Écrivons $K = \bigcup_{i=1}^N K_i$ où les K_i sont choisis deux à deux disjoints et tels que

$$\alpha|F_i x| \leq \int_{F_i x} |f(g)| dg \text{ pour } x \in K_i \text{ et } 1 \leq i \leq N.$$

La stratégie pour établir (3.3.2) à l'aide des inégalités (3.3.3) consiste à sélectionner, par un procédé de type Vitali, un nombre fini de translatés à droite de chaque F_i par des éléments de K_i , $1 \leq i \leq N$. On commence par choisir les translatés de la plus grande partie, F_N , et on procède dans l'ordre décroissant, de façon que les translatés des F_i choisis soient deux à deux disjoints et "recouvrent assez" K pour permettre d'évaluer sa mesure. Dégageons ce résultat sous forme de lemme.

Lemme 3.3.3. — *Soient F_1, \dots, F_N et K_1, \dots, K_N des parties relativement compactes de G . On suppose que la suite $(F_i)_{1 \leq i \leq N}$ est croissante et que $|F_1| > 0$. Alors il existe des parties finies D_N, D_{N-1}, \dots, D_1 , respectivement contenues dans K_N, K_{N-1}, \dots, K_1 telles que*

- (i) *les translatés $F_j x$, $x \in D_j$, $j = 1, \dots, N$ sont deux à deux disjoints ;*
- (ii) $\bigcup_{i=1}^N K_i \subset \bigcup_{j=1}^N F_j^{-1} F_j D_j$.

Démonstration. — Nous commençons donc par sélectionner les translatés de F_N en considérant une partie maximale $D_N \subset K_N$ telle que $F_N x \cap F_N y = \emptyset$ si x et y sont deux points distincts de D_N . Notons que l'ensemble D_N est fini puisque les $F_N x$, $x \in D_N$, ont la même mesure de Haar à droite, sont deux à deux disjoints, et sont contenus dans la partie $F_N K_N$, relativement compacte et donc de mesure finie.

Nous choisissons maintenant D_{N-1} maximale dans K_{N-1} telle que les translatés $F_{N-1} x$ avec $x \in D_{N-1}$ soient deux à deux disjoints et ne rencontrent pas $F_N D_N =$

$\bigcup_{y \in D_N} F_N y$. Si $D_N, D_{N-1}, \dots, D_{k+1}$ sont choisis de cette façon, nous sélectionnons de même une partie (nécessairement finie) maximale D_k de K_k telle que les $F_k x, x \in D_k$, soient deux à deux disjoints et ne rencontrent pas $\bigcup_{j>k} F_j D_j$. Nous procédons ainsi jusqu'à $k = 1$.

Ainsi les $F_j x, x \in D_j, j = 1, \dots, N$ sont deux à deux disjoints. Pour tout i , montrons que $K_i \subset \bigcup_{j=1}^N F_j^{-1} F_j D_j$. Soit $x \in K_i$. Il existe $j \geq i$ et $y \in D_j$ tel que $F_i x \cap F_j y \neq \emptyset$. En effet, ou bien $x \in D_i$ et on prend $j = i$ et $y = x$, ou bien $x \notin D_i$, et l'assertion résulte de la maximalité de D_i . Donc x appartient à $F_i^{-1} F_j D_j$ avec $i \leq j$, et on conclut en utilisant l'inclusion $F_i \subset F_j$. \square

Pour terminer la démonstration du théorème 3.3.2, observons que si la suite (F_i) vérifie la condition (T), alors

$$|F_i^{-1} F_i x| = \Delta(x) |F_i^{-1} F_i| \leq c \Delta(x) |F_i| = c |F_i x|$$

pour tout x . Nous avons alors, en prenant les D_i donnés par le lemme précédent,

$$\begin{aligned} |K| &\leq \sum_{i=1}^N \sum_{x \in D_i} |F_i^{-1} F_i x| \leq c \sum_{i=1}^N \sum_{x \in D_i} |F_i x| \\ &\leq \frac{c}{\alpha} \sum_{i=1}^N \sum_{x \in D_i} \int_{F_i x} \varphi(g) \, dg \leq \frac{c}{\alpha} \|\varphi\|_1. \end{aligned}$$

\square

Remarque 3.3.4. — Dans le cas où la suite (F_n) n'est pas supposée croissante, la démonstration précédente s'adapte aisément, lorsque la condition (T) est remplacée par la condition plus forte suivante introduite par Shulman : il existe une constante $c > 0$ telle

$$\forall n \geq 1, \quad \left| \bigcup_{i=1}^n F_i^{-1} F_n \right| \leq c |F_n|. \quad (T')$$

Si (F_n) est une suite satisfaisant à la condition de Tempel'man (ou à la condition (T')) et si de plus c'est une suite de Følner, les théorèmes ergodiques du chapitre 2 s'appliquent aux opérateurs de moyenne sur les parties F_n . Observons cependant que les propriétés de Tempel'man et de Følner sont indépendantes. Pour une discussion du fait que la condition de Tempel'man n'entraîne pas la propriété de Følner nous renvoyons par exemple à [Eme74]. Dans l'autre direction, voici un exemple simple d'une suite de Følner sur le groupe \mathbb{Z} pour laquelle le théorème ergodique ponctuel de Birkhoff est faux.

Exemple 3.3.5. — Soit $X = \{0, 1\}^{\mathbb{Z}}$ muni du produit m des mesures $\frac{1}{2}(\delta_0 + \delta_1)$ sur chaque composante, et soit t le décalage défini par $(tx)_n = x_{n+1}$ pour tout $x = (x_n)_{n \in \mathbb{Z}} \in X$. Cette transformation t est bijective et elle préserve m . On considère l'action correspondante de \mathbb{Z} sur X . Afin de construire une suite de Følner (F_n)

sur le groupe \mathbb{Z} , mettant en défaut le théorème ergodique ponctuel de Birkhoff, on se donne une suite (ℓ_n) de nombres entiers strictement positifs qui tend vers l'infini suffisamment lentement pour que $\sum_{n \geq 0} \frac{1}{2^{\ell_n}} = +\infty$. On pose $L_n = \sum_{i=0}^{n-1} \ell_i$ et

$$F_n = \{1 + L_n, 2 + L_n, \dots, \ell_n + L_n\}.$$

La suite (F_n) est de Følner, mais si f désigne la fonction caractéristique de l'ensemble des suites $(x_n)_{n \in \mathbb{Z}}$ telles que $x_0 = 1$, on voit facilement que la suite des $A_n f = \frac{1}{\ell_n} \sum_{k \in F_n} f \circ t^k$ diverge presque sûrement. En effet, puisque les variables aléatoires $A_n f$ sont indépendantes, on déduit du lemme de Borel-Cantelli que, pour presque tout $x \in X$, la suite $(A_n f(x))$ prend une infinité de fois la valeur 0 et une infinité de fois la valeur 1. Il n'y a donc pas d'inégalité maximale pour cette suite de Følner (F_n) .

En fait, on peut même construire des exemples de suites de Følner *croissantes* pour lesquelles le théorème ergodique est faux. On pourra consulter l'article [BL85] sur ce sujet.

Remarque 3.3.6. — La question de l'existence, sur tout groupe moyennable, de suites de Følner croissantes satisfaisant en même temps à la condition de Tempel'man est donc naturelle. Lindenstrauss a donné un exemple de groupe résoluble (et donc moyennable) qui ne possède pas de telles suites [Lin01]. Il s'agit du produit semi-direct $\mathbb{Z} \ltimes \left(\bigoplus_{i \in \mathbb{Z}} \mathbb{Z}/2\mathbb{Z} \right)$ où \mathbb{Z} agit par décalage sur le groupe $\bigoplus_{i \in \mathbb{Z}} \mathbb{Z}/2\mathbb{Z}$ des applications à support fini de \mathbb{Z} dans $\mathbb{Z}/2\mathbb{Z}$.

Cet exemple concerne un groupe à deux générateurs, à croissance exponentielle. On ne sait pas s'il s'agit d'un fait vérifié plus généralement par tous les groupes discrets moyennables, avec un nombre fini de générateurs et à croissance exponentielle.

En revanche, pour tout groupe localement compact G à génération compacte et à croissance polynomiale, la situation est très claire. Soit B un voisinage compact de l'élément neutre tel que $\bigcup_{n=1}^{+\infty} B^n = G$. En combinant des idées de Guivarc'h, Jenkins, Gromov, Losert, Mostow et Wang, on montre (voir [Tes05]) que G admet une croissance polynomiale stricte : il existe un entier $r \geq 0$, indépendant du choix de B , et $C \geq 1$ tels que, pour tout n , on ait

$$C^{-1}n^r \leq |B^n| \leq Cn^r.$$

En particulier, la suite (B^n) possède la propriété de doublement : $|B^{2n}| \leq C^2 2^r |B^n|$. Donc, si $B = B^{-1}$, la suite (B^n) possède la propriété de Tempel'man. Par ailleurs, R. Tessera a démontré que (B^n) est une suite de Følner⁽¹⁾.

Ainsi, grâce au théorème 3.3.2 et aux résultats du chapitre 2, le théorème ergodique de Wiener 3.1.3 s'étend au cas d'un tel groupe G lorsque la boule $B(0, 1)$ dans \mathbb{R}^d est

⁽¹⁾Tout cela se déduit aussi du résultat récent de Breuillard, établissant que la croissance polynomiale est même exacte : il existe une constante $C(B) > 0$ telle que $\lim_{n \rightarrow \infty} |B^n|/n^r = C(B)$.

remplacée par n'importe quel voisinage symétrique compact de e dans G qui engendre G . Pour plus de détails nous renvoyons à [Tes05] ou [Bre07].

Nous allons maintenant considérer le cas des suites (F_n) sur un groupe localement compact, possédant une propriété plus faible que la condition de Shulman, et également introduite par cet auteur. Cette notion de suite, dite tempérée, est intéressante pour les raisons suivantes : l'inégalité maximale de type faible $(1, 1)$ reste encore valable pour ces suites, et de plus tout groupe moyennable possède des suites de Følner tempérées. Nous consacrons la section suivante à l'étude de cet important résultat de Lindenstrauss.

3.4. Suites tempérées

Définition 3.4.1. — Une suite (F_n) de parties relativement compactes d'un groupe localement compact G est dite *tempérée* s'il existe une constante $c > 0$ telle qu'on ait :

$$\forall n \geq 1, \quad \left| \bigcup_{i=1}^n F_i^{-1} F_{n+1} \right| \leq c |F_{n+1}|. \quad (S)$$

Cette condition (S) ressemble à la condition (T') . Elle est toutefois plus faible et beaucoup plus facile à réaliser.

Lemme 3.4.2. — *Tout groupe moyennable G possède des suites de Følner tempérées.*

Démonstration. — Soit (F_n) une suite de parties relativement compactes dans G de mesure strictement positive, telle que pour toute partie compacte K , on ait

$$\lim_{n \rightarrow +\infty} \frac{|KF_n|}{|F_n|} = 1$$

(voir 2.1.6). Il suffit d'en extraire une sous-suite (F_{n_k}) telle que pour tout k on ait

$$\left| \left(\bigcup_{i=1}^k F_{n_i}^{-1} \right) F_{n_{k+1}} \right| \leq 2 |F_{n_{k+1}}|.$$

□

Énonçons maintenant le théorème maximal relatif aux moyennes sur les suites tempérées. Les notations sont celles de la section 3.3.

Théorème 3.4.3 ([Lin01]). — *Soit (F_n) une suite tempérée dans un groupe localement compact G , formée de parties boréliennes relativement compactes, de mesure strictement positive. Alors il existe une constante c' ne dépendant que de la constante*

c apparaissant dans la condition (S), telle que, pour tout $\varphi \in L^1(G)$ et tout $\alpha > 0$, on ait

$$|\{A^*\varphi > \alpha\}| \leq \frac{c'}{\alpha} \|\varphi\|_1.$$

La démonstration de Lindenstrauss, qui utilise un lemme de recouvrement de nature probabiliste, sera exposée dans le paragraphe 3.4.2. Pour le lecteur non probabiliste, nous allons commencer par donner la démonstration proposée par Ornstein et Weiss dans le cas où G est discret [Wei03].

3.4.1. La démonstration de Ornstein et Weiss. — Elle est basée sur le lemme de recouvrement 3.4.4 ci-dessous.

Démonstration du théorème 3.4.3 ; cas où G est discret. — Soit φ une fonction positive dans $L^1(G)$, et soit $\alpha > 0$. On choisit $N \geq 1$ et une partie finie $K \subset \{A_N^*\varphi > \alpha\}$, partitionnée sous la forme $K = \bigcup_{i=1}^N K_i$, comme au début de la démonstration de la proposition 3.3.2. Pour $i = 1, \dots, N$, on a donc

$$(3.4.1) \quad \forall x \in K_i, \quad \alpha |F_i x| \leq \int_{F_i x} \varphi(g) \, dg.$$

Nous allons montrer que $|K| \leq \frac{8(c+1)}{\alpha} \int_G \varphi(g) \, dg$, ce qui suffit pour obtenir l'inégalité maximale, comme on l'a déjà remarqué dans la démonstration de la proposition 3.3.2. Comme dans cette démonstration, nous allons sélectionner successivement des translatés à droite $F_i x$, avec $x \in K_i$, dans l'ordre décroissant des indices. On va procéder ici de façon que les translatés choisis pour un même F_i ne se recoupent pas trop, que leur réunion soit assez grosse, et que les translatés de parties F_i et F_j avec $i \neq j$ aient une intersection vide. Chaque étape s'appuie sur le lemme suivant, que l'on appliquera avec $E = G$, en prenant pour E_1, \dots, E_l les translatés à droite d'une même partie finie de G , et pour ρ la mesure de densité φ/α par rapport à la mesure de comptage sur G .

Lemme 3.4.4. — Soient E un ensemble, et E_1, \dots, E_l des parties finies de E . Soient $v \in \mathbb{N}^*$ et ρ une mesure positive sur E tels que

- (i) $v \leq \rho(E_i)$ pour $1 \leq i \leq l$;
- (ii) $\sum_{i=1}^l \mathbf{1}_{E_i}(x) \leq v$ pour tout $x \in E$.

Alors il existe une partie D de $\{1, \dots, l\}$ telle que

$$(3.4.2) \quad l \leq 4\rho\left(\bigcup_{i \in D} E_i\right) \quad \text{et} \quad |D|v \leq 4\rho\left(\bigcup_{i \in D} E_i\right)$$

La première inégalité de (3.4.2) signifie que $\bigcup_{i \in D} E_i$ est assez gros, et la deuxième indique que les $E_i, i \in D$, ne se recoupent pas trop.

Démonstration. — On peut supposer $v \geq 2$, car si $v = 1$ la condition (ii) montre que les E_i sont disjoints et le résultat est évident. L'ensemble D est défini par récurrence. On part de $i(1) = 1$. On suppose choisis $i(j)$ pour $j \leq k$, et on choisit pour $i(k+1)$ le plus petit entier $\leq l$, strictement supérieur à $i(k)$, tel que

$$\rho(E_{i(k+1)} \setminus \bigcup_{j=1}^k E_{i(j)}) \geq \frac{1}{2} \rho(E_{i(k+1)}),$$

si un tel entier existe. Sinon, on arrête le processus et on pose $D = \{i(1), \dots, i(k)\}$. Observons que

$$\rho\left(\bigcup_{i \in D} E_i\right) \geq \frac{1}{2} \sum_{i \in D} \rho(E_i) \geq \frac{|D|v}{2}.$$

Si de plus $|D| \geq \frac{l}{v}$, on a aussi $\rho\left(\bigcup_{i \in D} E_i\right) \geq \frac{l}{2}$ et les conditions (3.4.2) sont réalisées.

Il reste à examiner le cas où $|D| < l/v$. Posons $U = \bigcup_{i \in D} E_i$, et notons \bar{D} l'ensemble des entiers i , $1 \leq i \leq l$, qui n'appartiennent pas à D . Si $i \in \bar{D}$, on a $\rho(E_i \cap U) \geq \frac{1}{2} \rho(E_i)$, autrement dit

$$\int_E \mathbf{1}_{E_i}(x) \mathbf{1}_U(x) d\rho(x) \geq \frac{1}{2} \rho(E_i).$$

En utilisant l'hypothèse (ii) et en sommant sur $i \in \bar{D}$, on obtient

$$v\rho(U) \geq \frac{1}{2} \sum_{i \in \bar{D}} \rho(E_i),$$

d'où, grâce à (i),

$$\rho(U) \geq \frac{|\bar{D}|}{2} = \frac{l - |D|}{2} \geq \frac{l}{2} \left(1 - \frac{1}{v}\right) \geq \frac{l}{4}.$$

□

Revenons-en à la démonstration du théorème 3.4.3 dans le cas discret. Nous commençons par sélectionner les translatés de F_N en appliquant le lemme 3.4.4 aux ensembles $F_N x$ pour $x \in K_N$ et à la mesure ρ de densité φ/α , avec $v = |F_N|$. D'après (3.4.1), la condition (i) est réalisée. De plus, le nombre de translatés $F_N x$ qui contiennent un élément donné $g \in G$ est majoré par le cardinal de $g^{-1}F_N$, c'est-à-dire par $|F_N|$, ce qui assure (ii). Grâce à ce lemme, on peut trouver une partie D_N de K_N telle que

$$|K_N| \leq 4\rho(F_N D_N) \quad \text{et} \quad |D_N| |F_N| \leq 4\rho(F_N D_N).$$

Notons \tilde{K}_{N-1} l'ensemble des $x \in K_{N-1}$ tels que $F_{N-1}x \cap F_N D_N = \emptyset$. Pour le choix des translatés de F_{N-1} , on distingue deux cas :

– ou bien $|\tilde{K}_{N-1}| < \frac{1}{2}|K_{N-1}|$ et on ne sélectionne pas de translaté de F_{N-1} (ainsi $D_{N-1} = \emptyset$) ;

– ou bien $|\tilde{K}_{N-1}| \geq \frac{1}{2}|K_{N-1}|$, et alors on applique le lemme 3.4.4 aux translatés $F_{N-1}x$ avec $x \in \tilde{K}_{N-1}$: on choisit $D_{N-1} \subset \tilde{K}_{N-1}$ tel que

$$|K_{N-1}| \leq 8\rho(F_{N-1}D_{N-1}) \quad \text{et} \quad |D_{N-1}||F_{N-1}| \leq 4\rho(F_{N-1}D_{N-1}).$$

Observons que $F_{N-1}D_{N-1} \cap F_N D_N = \emptyset$.

Après la k -ième étape, on dispose d'un ensemble $J_k \subset \{N-k+1, \dots, N\}$, formé des j pour lesquels on a pu sélectionner des translatés de F_j , et pour chacun de ces j , d'une partie $D_j \subset K_j$ telle que les $F_j D_j$ soient deux à deux disjoints avec

$$(3.4.3) \quad |K_j| \leq 8\rho(F_j D_j) \quad \text{et} \quad |D_j||F_j| \leq 4\rho(F_j D_j).$$

On note alors \tilde{K}_{N-k} l'ensemble des $x \in K_{N-k}$ tels que

$$F_{N-k}x \cap \left(\bigcup_{j \in J_k} F_j D_j \right) = \emptyset.$$

Si $|\tilde{K}_{N-k}| < \frac{1}{2}|K_{N-k}|$, on passe à l'étape suivante, sinon, on choisit $D_{N-k} \subset \tilde{K}_{N-k}$ tel que

$$|K_{N-k}| \leq 8\rho(F_{N-k}D_{N-k}) \quad \text{et} \quad |D_{N-k}||F_{N-k}| \leq 4\rho(F_{N-k}D_{N-k}).$$

La construction s'arrête après la N -ième étape. On pose $J = J_N$ et on désigne par \bar{J} l'ensemble des $j \in \{1, \dots, N\}$ n'appartenant pas à J . D'après (3.4.3), on a

$$(3.4.4) \quad \sum_{j \in J} |K_j| \leq 8\rho\left(\bigcup_{j \in J} F_j D_j\right).$$

Par ailleurs, si $i \in \bar{J}$, au moins la moitié des $g \in K_i$ sont dans

$$\bigcup_{j \in J, j > i} F_i^{-1} F_j D_j.$$

Il en résulte qu'au moins la moitié de $\bigcup_{i \in \bar{J}} K_i$ est contenue dans

$$\bigcup_{j \in J, j > 1} \left(\bigcup_{1 \leq i < j} F_i^{-1} \right) F_j D_j.$$

En utilisant le fait que (F_n) est tempérée, puis la deuxième inégalité de (3.4.3), on obtient

$$\begin{aligned} \sum_{i \in \bar{J}} |K_i| &\leq 2 \left| \bigcup_{j \in J, j > 1} \left(\bigcup_{1 \leq i < j} F_i^{-1} \right) F_j D_j \right| \\ &\leq 2 \sum_{j \in J, j > 1} \left| \left(\bigcup_{1 \leq i < j} F_i^{-1} \right) F_j D_j \right| \\ &\leq 2 \sum_{j \in J, j > 1} \left| \left(\bigcup_{1 \leq i < j} F_i^{-1} \right) F_j \right| |D_j| \\ &\leq 2c \sum_{j \in J} |F_j| |D_j| \leq 8c \sum_{j \in J} \rho(F_j D_j). \end{aligned}$$

Finalement, en utilisant (3.4.4) on trouve

$$|K| \leq 8(c+1)\rho(G) = \frac{8(c+1)}{\alpha} \int_G \varphi(g) dg.$$

□

3.4.2. La démonstration de Lindenstrauss. — Dans cette section G est à nouveau un groupe localement compact à base dénombrable d'ouverts, muni de la mesure de Haar à gauche λ . Pour un borélien B , on notera également $|B|$ sa mesure de Haar à gauche. On utilisera aussi la mesure de Haar à droite λ_d sur G définie par $\lambda_d(B) = \lambda(B^{-1})$. Si G est discret on prendra pour λ la mesure de comptage et alors $\lambda_d = \lambda$.

La démonstration de Lindenstrauss repose sur un lemme de sélection probabiliste jouant le rôle du lemme de recouvrement de Vitali. L'idée de la construction mise en œuvre dans ce lemme est la suivante.

Un compact K , réunion disjointe de boréliens K_i , $1 \leq i \leq N$, est donné dans G ainsi que la famille $(F_i, 1 \leq i \leq N)$ de compacts. Pour chaque indice i , $1 \leq i \leq N$, la famille $(F_i h, h \in K_i)$ est une famille de translatés de F_i . On sélectionne alors successivement, de manière aléatoire, pour $i = N, N-1, \dots, 1$, une partie finie de K_i , donc une famille finie de translatés de F_i . Ces sélections sont réalisées à l'aide de mesures ponctuelles aléatoires dont les supports sont des sous-ensembles discrets de G qui sont, en un certain sens, équirépartis dans G . Si les densités de ces ensembles aléatoires sont convenablement choisies, on obtient que, en moyenne, les translatés des ensembles F_i ainsi sélectionnés ne se recoupent pas trop et recouvrent un ensemble de mesure assez grande relativement à la mesure de K .

Avant d'énoncer ce lemme, précisons une notation. Soient $(\Omega, \mathcal{A}, \mathbb{P})$ un espace probabilisé, Y une variable aléatoire sur Ω et A un événement de \mathcal{A} . Si $\mathbb{P}(A) \neq 0$, on note $\mathbb{E}(Y|A) = \frac{\mathbb{E}(Y\mathbf{1}_A)}{\mathbb{P}(A)}$ l'espérance de Y pour la probabilité conditionnelle sachant A . Si $\mathbb{P}(A) = 0$, on pose $\mathbb{E}(Y|A) = 0$.

Lemme 3.4.5. — Soient un réel $c > 0$, un entier N et des parties compactes F_1, \dots, F_N de G telles que, pour $1 \leq k < N$, on ait

$$\left| \bigcup_{1 \leq i \leq k} F_i^{-1} F_{k+1} \right| \leq c |F_{k+1}|.$$

Soit un réel $\delta > 0$ vérifiant, dans le cas où G est discret, $\delta/|F_i| \leq 1$ pour tout $1 \leq i \leq N$. Soient, d'autre part, des boréliens relativement compacts K_1, \dots, K_N de G .

Il existe un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et, pour tout élément ω de Ω , des parties finies $D_N(\omega), D_{N-1}(\omega), \dots, D_1(\omega)$ respectivement contenues dans les ensembles

K_N, K_{N-1}, \dots, K_1 , telles que, si on considère sur $\Omega \times G$ la fonction de comptage

$$\Lambda(\omega, g) = \sum_{i=1}^N \sum_{h \in D_i(\omega)} \mathbf{1}_{F_i h}(g),$$

alors

- (i) la fonction Λ est mesurable sur $\Omega \times G$;
- (ii) pour tout $g \in G$, on a $\mathbb{E}(\Lambda(g) | \Lambda(g) \geq 1) \leq 1 + \delta$;
- (iii) $\mathbb{E} \left(\int_G \Lambda(g) d\lambda(g) \right) \geq \frac{\delta}{1 + c\delta} \left| \bigcup_{i=1}^N K_i \right|$.

Comme on le voit, la condition (ii) exprime bien que les translatés des F_i qui ont été sélectionnés ne se recoupent pas trop et la condition (iii) qu'ils recouvrent un ensemble de mesure assez grande relativement à la mesure de $\bigcup_{i=1}^N K_i$.

Donnons-nous d'abord quelques outils de nature probabiliste.

Définition 3.4.6. — On dit qu'une mesure de Radon μ sur G est une *mesure ponctuelle* si elle est de la forme $\mu = \sum_i \delta_{g_i}$ où (g_i) est une famille localement finie de points de G . On dit que c'est une mesure ponctuelle simple si les points g_i sont distincts. L'ensemble \mathcal{M}_p des mesures ponctuelles sur G est muni de la plus petite tribu rendant mesurables les applications $\mu \mapsto \mu(B)$, pour tous les boréliens B de G .

Une *mesure ponctuelle aléatoire* définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ et à valeurs dans G est une application mesurable π de (Ω, \mathcal{A}) dans \mathcal{M}_p .

Une mesure ponctuelle aléatoire π à valeurs dans G définit donc, pour tout borélien B , une application mesurable $\omega \mapsto \pi(\omega, B)$ à valeurs dans $\mathbb{N} \cup \{+\infty\}$. Si π est simple, $\pi(\omega, B)$ est le cardinal de $B \cap \text{supp}(\pi(\omega))$.

Définition 3.4.7. — Nous dirons qu'une mesure ponctuelle aléatoire π à valeurs dans G est une *mesure ponctuelle aléatoire homogène* (m.p.h.) de paramètre θ , ($0 \leq \theta \leq 1$), si elle vérifie les conditions suivantes :

- (i) Pour toute suite (B_1, \dots, B_m) de boréliens disjoints de G , les variables aléatoires $\pi(B_1), \dots, \pi(B_m)$ sont indépendantes.
- (ii) Pour tout borélien B de G , de mesure de Haar finie,
 1. si G est un groupe discret, $\pi(B)$ est une variable aléatoire binomiale de paramètres $(\lambda(B), \theta)$;
 2. si G n'est pas discret, $\pi(B)$ est une variable aléatoire de Poisson de paramètre $\theta \lambda_d(B)$.

Dans le cas où G n'est pas discret, on ne fait que redonner la définition d'un processus ponctuel de Poisson sur G d'intensité $\theta \lambda_d$. Dans le cas où G est discret une m.p.h. de paramètre θ s'obtient en considérant une famille $(X_g)_{g \in G}$ de variables aléatoires de Bernoulli indépendantes de paramètre θ et en posant $\pi(\omega) = \sum_{g \in G} X_g(\omega) \delta_g$. Dans

les deux cas, presque sûrement, la mesure π est simple et son support est un ensemble localement fini de points de G . En moyenne, on a $\mathbb{E}(\pi(B)) = \theta \lambda_d(B)$. Plus généralement, si B est un borélien de G et ψ une fonction mesurable positive sur G , on pourra écrire

$$\mathbb{E} \left(\sum_{h \in B \cap \text{supp}(\pi)} \psi(h) \right) = \mathbb{E} \left(\int_B \psi(g) d\pi(g) \right) = \theta \int_B \psi(g) d\lambda_d(g).$$

Comme on le voit, le support d'une m.p.h. sur G constitue, que G soit discret ou non, un ensemble aléatoire de points de G qui sont, d'une certaine manière, équirépartis dans G .

Nous aurons aussi besoin des majorations suivantes concernant les lois binomiales et de Poisson.

Lemme 3.4.8. — *Soit Y une variable aléatoire réelle qui suit la loi de Poisson de paramètre α ou bien la loi binomiale de paramètres (n, p) . On a, dans les deux cas,*

$$\mathbb{E}(Y|Y \geq 1) \leq 1 + \mathbb{E}(Y).$$

Démonstration. — Si on explicite l'expression $\mathbb{E}(Y|Y \geq 1)$, on est ramené à vérifier selon le cas l'inégalité $\frac{\alpha}{1 - e^{-\alpha}} \leq 1 + \alpha$ ou bien l'inégalité $\frac{np}{1 - (1-p)^n} \leq 1 + np$. Ces vérifications sont élémentaires. \square

On a, plus généralement,

Lemme 3.4.9. — *Étant donné $\alpha > 0$, soit Y une variable aléatoire définie sur un espace probabilisé produit*

$$(\Omega, \mathcal{A}, \mathbb{P}) = (\Omega_1, \mathcal{A}_1, \mathbb{P}_1) \times (\Omega_2, \mathcal{A}_2, \mathbb{P}_2)$$

telle que, pour tout $\omega_2 \in \Omega_2$, la variable aléatoire $\omega_1 \mapsto Y(\omega_1, \omega_2)$ suit une loi de Poisson de moyenne $\alpha(\omega_2)$. Si $\alpha(\omega_2) \leq \alpha$ pour tout $\omega_2 \in \Omega_2$, alors $\mathbb{E}(Y|Y \geq 1) \leq 1 + \alpha$.

De même, si pour tout $\omega_2 \in \Omega_2$, la variable aléatoire $\omega_1 \mapsto Y(\omega_1, \omega_2)$ suit une loi binomiale de moyenne majorée par α , on a encore $\mathbb{E}(Y|Y \geq 1) \leq 1 + \alpha$.

Démonstration. — Dans le cas de variables de Poisson on a, si $\mathbb{P}(Y \geq 1) > 0$,

$$\begin{aligned} \mathbb{E}(Y|Y \geq 1) &= \frac{\mathbb{E}(Y \mathbf{1}_{Y \geq 1})}{\mathbb{P}(Y \geq 1)} = \frac{\mathbb{E}(Y)}{\mathbb{P}(Y \geq 1)} \\ &= \frac{\int Y(\omega_1, \omega_2) d\mathbb{P}_1(\omega_1) d\mathbb{P}_2(\omega_2)}{\int \mathbf{1}_{\{Y \geq 1\}}(\omega_1, \omega_2) d\mathbb{P}_1(\omega_1) d\mathbb{P}_2(\omega_2)} \\ &= \frac{\int \alpha(\omega_2) d\mathbb{P}_2(\omega_2)}{\int (1 - e^{-\alpha(\omega_2)}) d\mathbb{P}_2(\omega_2)}. \end{aligned}$$

Mais, comme on l'a vu dans le lemme précédent,

$$\alpha(\omega_2) \leq (1 - e^{-\alpha(\omega_2)})(1 + \alpha(\omega_2)) \leq (1 - e^{-\alpha(\omega_2)})(1 + \alpha).$$

Et donc, en intégrant par rapport à \mathbb{P}_2 ,

$$\frac{\int \alpha(\omega_2) d\mathbb{P}_2(\omega_2)}{\int (1 - e^{-\alpha(\omega_2)}) d\mathbb{P}_2(\omega_2)} \leq 1 + \alpha.$$

Le cas des variables binomiales se traite de même. \square

Démonstration du lemme 3.4.5. — Venons-en maintenant à la démonstration proprement dite du lemme 3.4.5.

Soit une suite $(\pi_i; 1 \leq i \leq N)$ de mesures ponctuelles aléatoires homogènes sur G définies sur des espaces probabilisés $(\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$ et respectivement de paramètre $\theta_i = \delta/|F_i|$. Ces mesures aléatoires définissent sur l'espace probabilisé produit

$$(\Omega, \mathcal{A}, \mathbb{P}) = \prod_{1 \leq i \leq N} (\Omega_i, \mathcal{A}_i, \mathbb{P}_i)$$

des m.p.h. indépendantes si on pose, pour $\omega = (\omega_1, \dots, \omega_N)$, $\pi_j(\omega) = \pi_j(\omega_j)$.

Fixons $\omega \in \Omega$ et décrivons un algorithme conduisant à la définition de la suite $D_N(\omega), \dots, D_1(\omega)$ à partir des données $(F_j)_{1 \leq j \leq N}$ et $(K_j)_{1 \leq j \leq N}$.

Comme on le verra, à chaque étape i de cet algorithme, on définit l'ensemble $D_i(\omega)$ en effectuant une sélection aléatoire parmi les translatés de F_i qui ont été conservés à ce stade, c'est-à-dire ceux qui sont disjoints des translatés $F_j h$, $h \in D_j(\omega)$, des ensembles F_j d'indice $j > i$, sélectionnés aux stades antérieurs.

Pour cela, on part de l'indice $j = N$ pour lequel on pose $\tilde{K}_j(\omega) = K_j$ et $\tilde{K}_{i|j}(\omega) = K_i$ si $1 \leq i < j$.

On effectue alors les opérations suivantes.

I. On pose $D_j(\omega) = \tilde{K}_j(\omega) \cap \text{supp}(\pi_j(\omega))$.

II. Pour $i \leq j - 1$ on pose

$$\tilde{K}_{i|j-1}(\omega) = \{h \in \tilde{K}_{i|j}(\omega); F_i h \cap F_j D_j(\omega) = \emptyset\}$$

$$\text{et } \tilde{K}_{j-1}(\omega) = \tilde{K}_{j-1|j-1}(\omega).$$

III. On remplace j par $j - 1$ et on retourne à l'étape I jusqu'à ce que $j = 1$.

On remarquera que, pour $1 < j \leq N$ et $1 \leq i \leq j - 1$, les ensembles $\tilde{K}_{i|j-1}(\omega)$ et $\tilde{K}_{j-1}(\omega)$ ne dépendent que des coordonnées $\omega_N, \dots, \omega_j$, c'est-à-dire des mesures aléatoires π_N, \dots, π_j .

Posons, pour tout $1 \leq j \leq N$,

$$(3.4.5) \quad \Lambda_j(\omega, g) = \sum_{h \in D_j(\omega)} \mathbf{1}_{F_j h}(g) = \sum_{h \in F_j^{-1}g \cap D_j(\omega)} 1$$

et

$$\Lambda(\omega, g) = \sum_{j=1}^N \Lambda_j(\omega, g).$$

Nous admettrons, sans entrer dans les détails, que cette fonction Λ est bien mesurable (condition (i) du lemme).

Par construction, pour $j \neq j'$, les éléments de la famille $\{F_j h ; h \in D_j(\omega)\}$ sont disjoints de ceux de la famille $\{F_{j'} h ; h \in D_{j'}(\omega)\}$. Il en résulte que, pour tout ω , les fonctions $g \mapsto \Lambda_j(\omega, g)$ ont, pour des indices distincts, des supports disjoints.

Vérifions alors la condition (ii) du lemme 3.4.5.

Pour $j = N$ et pour tout $g \in G$, $\Lambda_N(g) = \pi_N(F_N^{-1}g \cap K_N)$ est une variable aléatoire binomiale ou de Poisson d'espérance

$$\begin{aligned} \mathbb{E}(\Lambda_N(g)) &= \mathbb{E}(\pi_N(F_N^{-1}g \cap K_N)) = \frac{\delta}{|F_N|} \lambda_d(F_N^{-1}g \cap K_N) \\ &\leq \frac{\delta}{|F_N|} \lambda_d(F_N^{-1}g) = \frac{\delta}{|F_N|} \lambda_d(F_N^{-1}) = \delta. \end{aligned}$$

Par application du lemme 3.4.8, on obtient

$$\mathbb{E}(\Lambda_N(g) | \Lambda_N(g) \geq 1) \leq 1 + \delta.$$

Pour $j < N$ on peut écrire de même

$$\Lambda_j(\omega, g) = \pi_j(\omega, F_j^{-1}g \cap \tilde{K}_j(\omega))$$

ou, plus précisément,

$$\Lambda_j(\omega, g) = \pi_j(\omega_j, F_j^{-1}g \cap \tilde{K}_j(\omega_N, \dots, \omega_{j+1})).$$

Pour chaque $(\omega_N, \dots, \omega_{j+1})$ fixé, la variable aléatoire $\omega_j \mapsto \Lambda_j(\omega, g)$, est binomiale ou de Poisson, de moyenne $\frac{\delta}{|F_j|} \lambda_d(F_j^{-1}g \cap \tilde{K}_j(\omega_N, \dots, \omega_{j+1})) \leq \frac{\delta}{|F_j|} \lambda_d(F_j^{-1}g) = \delta$.

Donc, d'après le lemme 3.4.9,

$$\mathbb{E}(\Lambda_j(g) | \Lambda_j(g) \geq 1) \leq 1 + \delta.$$

Comme les événements $\{\Lambda_j \geq 1\}$ et $\{\Lambda_{j'} \geq 1\}$ sont incompatibles si $j \neq j'$, on a alors

$$\begin{aligned} \mathbb{E}(\Lambda(g) | \Lambda(g) \geq 1) &= \sum_{j=1}^N \mathbb{E}(\Lambda_j(g) | \Lambda_j(g) \geq 1) \mathbb{P}(\Lambda_j(g) \geq 1 | \Lambda(g) \geq 1) \\ &\leq 1 + \delta, \end{aligned}$$

ce qui montre la condition (ii) du lemme 3.4.5.

Vérifions maintenant la condition (iii). Le raisonnement se fera par récurrence sur l'entier N .

Soit Δ la fonction modulaire sur G . On voit d'abord que

$$\begin{aligned} (3.4.6) \quad \mathbb{E} \left(\int_G \Lambda_N(g) d\lambda(g) \right) &= \mathbb{E} \left(\sum_{h \in D_N} |F_N h| \right) = |F_N| \mathbb{E} \left(\sum_{h \in D_N} \Delta(h) \right) \\ &= |F_N| \mathbb{E} \left(\int_{K_N} \Delta(g) d\pi_N(g) \right). \end{aligned}$$

Par définition de la m.p.h. π_N , cette dernière expression vaut

$$|F_N| \int_{K_N} \Delta(g) \frac{\delta}{|F_N|} d\lambda_d(g) = \delta \int_{K_N} d\lambda(g).$$

Finalement

$$(3.4.7) \quad \mathbb{E} \left(\int_G \Lambda_N(g) d\lambda(g) \right) = \delta |K_N|,$$

ce qui prouve (iii) dans le cas où $N = 1$.

Supposons alors, par hypothèse de récurrence, que l'inégalité (iii) est vérifiée pour l'entier $N - 1$.

Soit \mathcal{B}_N la sous-tribu de \mathcal{A} des événements qui ne dépendent que de la coordonnée ω_N de ω . Pour une variable aléatoire positive Y sur Ω , l'espérance conditionnelle sachant \mathcal{B}_N est alors simplement donnée par

$$\mathbb{E}(Y|\mathcal{B}_N)(\omega_N) = \int_{\Omega_1 \times \dots \times \Omega_{N-1}} Y(\omega_1, \dots, \omega_N) d\mathbb{P}_1(\omega_1) \dots d\mathbb{P}_{N-1}(\omega_{N-1}).$$

Comme la variable $\Lambda_N(g)$ est \mathcal{B}_N -mesurable, on a

$$\begin{aligned} \mathbb{E} \left(\int_G \Lambda(g) d\lambda(g) | \mathcal{B}_N \right) &= \\ &= \mathbb{E} \left(\int_G \Lambda_N(g) d\lambda(g) | \mathcal{B}_N \right) + \mathbb{E} \left(\int_G \left(\sum_{j=1}^{N-1} \Lambda_j(g) \right) d\lambda(g) | \mathcal{B}_N \right) \\ &= \int_G \Lambda_N(g) d\lambda(g) + \mathbb{E} \left(\int_G \left(\sum_{j=1}^{N-1} \Lambda_j(g) \right) d\lambda(g) | \mathcal{B}_N \right). \end{aligned}$$

Sachant \mathcal{B}_N , c'est-à-dire pour ω_N fixé, nous pouvons appliquer l'hypothèse de récurrence à $(F_j)_{1 \leq j \leq N-1}$ et aux ensembles $(\tilde{K}_{j|N-1}(\omega))_{1 \leq j \leq N-1}$ qui, comme on l'a remarqué, ne dépendent en fait que de ω_N . On en déduit que

$$(3.4.8) \quad \mathbb{E} \left(\int_G \Lambda(g) d\lambda(g) | \mathcal{B}_N \right) (\omega) \geq \int_G \Lambda_N(\omega, g) d\lambda(g) + \frac{\delta}{1 + c\delta} \left| \bigcup_{j=1}^{N-1} \tilde{K}_{j|N-1}(\omega) \right|.$$

D'autre part, l'ensemble $\tilde{K}_{j|N-1}(\omega)$ est, par construction, l'ensemble des éléments h de K_j tels que $F_j h$ ne coupe pas $F_N D_N(\omega)$. Autrement dit

$$\tilde{K}_{j|N-1}(\omega) = K_j \setminus F_j^{-1} F_N D_N(\omega),$$

d'où

$$\bigcup_{j=1}^{N-1} \tilde{K}_{j|N-1}(\omega) \supset \left(\bigcup_{j=1}^{N-1} K_j \right) \setminus \left(\bigcup_{j=1}^{N-1} F_j^{-1} F_N \right) D_N(\omega).$$

On utilise maintenant le fait que la suite $(F_j)_{1 \leq j \leq N}$ est tempérée. On obtient ainsi

$$\begin{aligned} \left| \left(\bigcup_{j=1}^{N-1} F_j^{-1} F_N \right) D_N(\omega) \right| &\leq \left| \bigcup_{j=1}^{N-1} F_j^{-1} F_N \right| \sum_{h \in D_N(\omega)} \Delta(h) \\ &\leq c |F_N| \sum_{h \in D_N(\omega)} \Delta(h). \end{aligned}$$

À partir de l'inégalité (3.4.8), on déduit alors la relation

$$\begin{aligned} &\mathbb{E} \left(\int_{g \in G} \Lambda(g) d\lambda(g) | \mathcal{B}_N \right) (\omega) \\ &\geq \int_G \Lambda_N(\omega, g) d\lambda(g) + \frac{\delta}{1 + c\delta} \left(\left| \bigcup_{j=1}^{N-1} K_j \right| - c |F_N| \sum_{h \in D_N(\omega)} \Delta(h) \right). \end{aligned}$$

En prenant l'espérance des deux membres de cette inégalité et en utilisant le calcul de $\mathbb{E}(\sum_{h \in D_N} \Delta(h))$ fait en (3.4.6) ainsi que l'égalité (3.4.7), il vient

$$\begin{aligned} \mathbb{E} \left(\int_G \Lambda(g) d\lambda(g) \right) &\geq \delta |K_N| + \frac{\delta}{1 + c\delta} \left(\left| \bigcup_{j=1}^{N-1} K_j \right| - c\delta |K_N| \right) \\ &\geq \frac{\delta}{1 + c\delta} \left| \bigcup_{j=1}^N K_j \right|. \end{aligned}$$

Ceci termine la démonstration du lemme 3.4.5. □

Démonstration du théorème 3.4.3. — Voyons, pour finir, comment on déduit le théorème 3.4.3 du lemme de Lindenstrauss 3.4.5.

Soit φ une fonction positive dans $L^1(G)$ et soit $\alpha > 0$. Reprenons le début de la démonstration du théorème 3.3.2, avec les mêmes notations. Nous avons une famille K_i , $1 \leq i \leq N$, de parties de G deux à deux disjointes, de réunion K et vérifiant, comme en (3.3.3),

$$\forall x \in K_i, \quad \alpha |F_i x| \leq \int_{F_i x} \varphi(g) dg.$$

Appliquons le lemme 3.4.5 avec $\delta = 1$. Grâce au théorème de Fubini et à la propriété (ii), on a

$$\mathbb{E} \left(\int_G \Lambda(g) \varphi(g) dg \right) = \int_G \varphi(g) \mathbb{E}(\Lambda(g)) dg \leq 2 \int_G \varphi(g) dg.$$

D'autre part, en utilisant (3.3.3), on obtient

$$\begin{aligned} \mathbb{E} \left(\int_G \Lambda(g) \varphi(g) \, dg \right) &= \mathbb{E} \left(\sum_{1 \leq i \leq N} \sum_{h \in D_i} \int_{F_i h} \varphi(g) \, dg \right) \\ &\geq \alpha \mathbb{E} \left(\sum_{1 \leq i \leq N} \sum_{h \in D_i} |F_i h| \right) = \alpha \mathbb{E} \left(\int_G \Lambda(g) \, dg \right). \end{aligned}$$

D'où, avec la minoration (iii) du lemme 3.4.5

$$\mathbb{E} \left(\int_G \Lambda(g) \varphi(g) \, dg \right) \geq \alpha \frac{1}{1+c} \left| \bigcup_{i=1}^N K_i \right| = \alpha \frac{1}{1+c} |K|.$$

On en déduit finalement que

$$|K| \leq 2 \frac{1+c}{\alpha} \int_G \varphi(g) \, dg$$

et on termine, comme en 3.3.2, la démonstration de l'inégalité maximale. \square

Notice bibliographique. — Les premières inégalités maximales pour des suites (F_n) de parties d'un groupe G remontent à Hardy et Littlewood [HL30] pour des suites d'intervalles de \mathbb{R} et à Wiener [Wie39] pour des suites de boules de \mathbb{R}^d . La propriété de doublement de volume sur un groupe non abélien fut d'abord considérée par Calderón [Cal53], puis les moyennes relatives à des suites (F_n) de plus en plus générales ont fait l'objet de nombreuses études. En particulier, à partir de la fin des années 1960, beaucoup de travaux sont consacrés à l'obtention d'inégalités maximales à l'aide de lemmes de recouvrement. Le théorème 3.3.2 est dû à Tempel'man [Tem67] dans le cas où le groupe est unimodulaire et à Emerson [Eme74] dans le cas général. Le résultat le plus récent et le plus frappant sur ce sujet est celui de Lindenstrauss [Lin01] présenté dans la dernière partie de ce chapitre.

Les inégalités maximales ont été beaucoup développées dans le cadre de la différentiation des intégrales, vue maintenant comme généralisation du théorème de différentiation de Lebesgue, et aussi comme outil fondamental de la théorie de Calderón-Zygmund, les opérateurs maximaux "contrôlant" d'une certaine façon les opérateurs d'intégrale singulière. Le lecteur pourra consulter les ouvrages de Stein [Ste70a] et [Ste93], et les travaux mis en référence dans ce dernier. Il faut mentionner en particulier les inégalités à poids, qui jouent un rôle fondamental dans la théorie de Calderón-Zygmund depuis la découverte, faite par Muckenhoupt, qu'il était possible de caractériser tous les poids w pour lesquels l'opérateur maximal construit à l'aide des boules et de la mesure de Lebesgue dans \mathbb{R}^d s'étend en un opérateur continu dans $L^p(\mathbb{R}^d, w \, dx)$. Si on remplace les boules de \mathbb{R}^d par des rectangles de côtés parallèles aux axes dans la définition de la fonction maximale, il n'y a plus d'inégalité faible (1,1). Córdoba et Fefferman ont toutefois développé des lemmes de recouvrements adaptés [CF75] et [Cór79]. S'il n'y a pas de lemme de Vitali ou de Besicovitch pour

les rectangles, on peut citer le lemme de Journé [**Jou88, Jou86**], qui joue un rôle fondamental dans la théorie de Calderón-Zygmund à plusieurs paramètres, actuellement en plein développement (voir [**Lac07, MP11**] pour ne citer que des publications récentes). Lorsqu'on considère la fonction maximale relative aux rectangles dans toutes les directions, il n'y a plus d'inégalité maximale (p, p) pour $p < \infty$: les contre-exemples, qu'on peut déjà trouver dans les oeuvres de Busemann et Feller [**BF34**], reposent sur la construction par Besicovitch d'un ensemble de Kakeya, c'est-à-dire d'un ensemble de mesure nulle contenant des segments de droite de longueur fixée dans toutes les directions (voir [**Ste93**]). La fonction maximale de Kakeya est relative aux tubes dans \mathbb{R}^d dont le rapport entre le diamètre et la longueur est supérieur à δ . La conjecture dite *conjecture sur la fonction maximale de Kakeya* porte sur le comportement de la norme de l'opérateur maximal associé lorsque δ tend vers 0. Il y a eu regain d'intérêt pour ce problème du fait de ses relations avec le problème de Kakeya, qui porte sur la dimension de Hausdorff d'un ensemble de Kakeya (de dimension 2 dans le plan, mais le problème analogue multidimensionnel est ouvert). On peut consulter l'article de synthèse de Tao [**Tao01**] sur les ensembles de Kakeya.

Une mesure non doublante est typiquement une mesure supportée dans une sous variété de \mathbb{R}^d . De telles mesures sont beaucoup moins faciles à étudier que la mesure de Lebesgue. C'est un fait remarquable que l'opérateur maximal centré reste borné. Depuis une vingtaine d'années on s'est aperçu que d'autres opérateurs que l'on considérait habituellement uniquement pour la mesure de Lebesgue restaient bornés dans les espaces L^p lorsque la mesure n'a plus la propriété de doublement. Au départ on a établi de tels résultats pour la transformée de Hilbert, puis cela a été généralisé aux opérateurs de Calderón-Zygmund. La propriété de doublement est alors remplacée par une condition de Ahlfors : $\mu(B(x, r)) \leq r^k$, $k \leq d$. À la base de ces résultats, il y a l'étude des opérateurs maximaux associés à des mesures non doublantes, et le lemme de recouvrement adapté est celui de Besicovitch. L'étude de la transformée de Cauchy dans ce cadre a permis, par exemple, de résoudre la conjecture de Vitushkin, et celle de la semi-additivité de la capacité analytique. Voir [**Tol01a, Tol01b, NTV97, MMNO00**] pour un aperçu de l'utilisation des mesures non doublantes, ainsi que les références présentes dans ces articles.

On trouve récemment beaucoup d'utilisations des réseaux dyadiques et de la fonction maximale associée, notamment en vue de l'étude des opérateurs de Calderón-Zygmund. Nous renvoyons à [**NTV02**] pour plus de détails. On peut trouver des généralisations des réseaux dyadiques à d'autres ensembles que \mathbb{R}^d dans [**DM00**].

CHAPITRE 4

UN EXEMPLE : LE GROUPE DES AFFINITÉS

Nous étudions dans ce chapitre les propriétés ergodiques de l'action du groupe S_d des affinités de \mathbb{R}^d , appelé aussi groupe affine. Nous ne présentons pas une théorie achevée, mais nous exposons plutôt un certain nombre d'exemples et quelques problèmes ouverts.

Comme tous les groupes résolubles, le groupe S_d est moyennable. Nous poursuivons ici deux objectifs :

- 1) déterminer des suites de Følner de nature géométrique ;
- 2) étudier l'opérateur maximal pour certaines familles naturelles de sous-ensembles de S_d .

Le groupe S_d n'est pas unimodulaire. Par conséquent sa croissance est exponentielle (voir 2.1.7), ce qui signifie que pour tout voisinage compact B de l'origine on a $\lim |B^n|^{\frac{1}{n}} > 1$. Dans ce cas, (B^n) n'est pas une suite de Følner (voir la remarque 2.1.8). Observons par ailleurs que S_d s'identifie de manière canonique à l'espace hyperbolique (voir l'appendice D). Les boules géodésiques constituent donc une famille naturelle d'ensembles à étudier. Là aussi, nous verrons ce n'est pas une famille de Følner.

Une construction générale de familles de Følner, valable dans tout groupe localement compact moyennable, a été proposée par Greenleaf [Gre73] mais, dans le cas du groupe des affinités considéré ici, il est facile de trouver des suites de Følner de nature géométrique simple (Proposition 4.2.1).

En ce qui concerne la question des inégalités maximales, nous ne considérons que le cas de l'action de S_d sur lui-même (opérateurs de convolution). On voit apparaître un phénomène nouveau dû au fait que le groupe S_d n'est pas unimodulaire. Si \mathcal{E} est une famille de parties d'un groupe localement compact G , on peut considérer (comme nous l'avons fait jusqu'ici) l'opérateur maximal A^* relatif aux moyennes sur les translatés à droite Eg des éléments E de \mathcal{E} , mais aussi l'opérateur M^* relatif aux moyennes sur les translatés à gauche gE . Lorsque le groupe G est unimodulaire et que les parties E

sont symétriques (*i.e.* $E = E^{-1}$) les deux opérateurs ont le même comportement. Il n'en est plus de même si G n'est pas unimodulaire. Ainsi les théorèmes de Strömberg et Clerc-Stein montrent que l'opérateur maximal M^* pour les moyennes à gauche sur les boules hyperboliques est de type faible $(1, 1)$ et (donc) borné sur L^p , $p > 1$. En revanche on verra que l'opérateur maximal A^* pour les moyennes à droite sur ces mêmes boules n'est de type faible (p, p) pour aucun $p \geq 1$.

Observons que les translatés à gauche de ces boules sont encore des boules hyperboliques puisque la métrique est invariante à gauche ; il est par conséquent très naturel d'étudier pour eux l'opérateur M^* . Toutefois, le théorème de transfert ne s'applique qu'aux translatés à droite.

Ce chapitre présente un panorama des résultats connus (avec un résultat nouveau en 4.4.5) sur les inégalités maximales pour les opérateurs de moyenne associés à des familles de sous-ensembles de S_d . Il est divisé en cinq parties. La première décrit rapidement la structure du groupe des affinités S_d , ses mesures de Haar et son lien avec l'espace hyperbolique. La deuxième partie montre que les boules géodésiques ne constituent pas une famille de Følner et donne une construction de suite de Følner dans le groupe S_d . Dans la troisième, on étudie l'opérateur maximal pour les moyennes associées aux translatés à gauche des boules hyperboliques. Dans la quatrième partie, on s'intéresse aux inégalités maximales pour d'autres familles de sous-ensembles du groupe des affinités, dont certaines vérifient aussi la propriété de Følner. Cela nous permettra d'énoncer, dans la cinquième partie, un théorème ergodique pour les actions de S_d , relatif à une suite de Følner non tempérée.

4.1. Structure du groupe des affinités

Nous commençons tout d'abord par quelques rappels sur le *groupe des affinités*. Soit d un entier strictement positif. Le groupe S_d est l'ensemble des applications de \mathbb{R}^d dans lui-même de la forme $x \mapsto ax + b$ (avec $a > 0$ et $b \in \mathbb{R}^d$), muni de la loi de composition des applications. Cet ensemble est identifié à $\mathbb{R}^d \times \mathbb{R}_+^*$, en notant (b, a) l'application précédente. Ainsi le produit est donné par

$$(b, a)(b', a') = (b + ab', aa').$$

On distinguera deux sous-groupes de S_d , le sous-groupe N des translations,

$$N = \{n_b = (b, 1); b \in \mathbb{R}^d\} \quad (\text{i.e. } n_b : x \mapsto x + b),$$

isomorphe au groupe additif \mathbb{R}^d , et le sous-groupe A des homothéties,

$$A = \{(0, a); a \in \mathbb{R}_+^*\} \quad (\text{i.e. } (0, a) : x \mapsto ax),$$

isomorphe au groupe multiplicatif \mathbb{R}_+^* . Le plus souvent nous écrirons $n_b a$ au lieu de (b, a) . La relation

$$(b, a)n_c(b, a)^{-1} = n_{ac}$$

montre en particulier que N est un sous-groupe normal de S_d . Le groupe S_d apparaît comme le produit semi-direct $N \rtimes A$ de deux groupes abéliens. C'est donc un groupe résoluble.

L'action de S_d sur \mathbb{R}^d se prolonge en une action à gauche sur le demi-espace supérieur

$$\mathbb{H}^{d+1} = \{z = (x, y); x \in \mathbb{R}^d, y > 0\} = \mathbb{R}^d \times \mathbb{R}_+^*.$$

Pour $g = (b, a) \in S_d$ et $z = (x, y) \in \mathbb{H}^{d+1}$, on pose $gz = (ax + b, ay) = az + (b, 0)$. Cette action de S_d sur \mathbb{H}^{d+1} est simplement transitive. Ainsi S_d et \mathbb{H}^{d+1} s'identifient par l'application $g \mapsto go$ où $o = (0, 1) \in \mathbb{H}^{d+1}$. Notons que A agit par homothéties et que N agit par translations horizontales (cf. fig. 1).

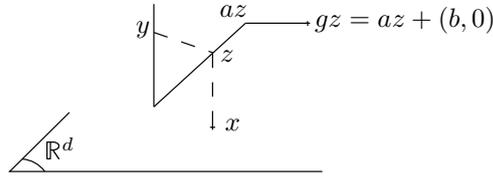


FIGURE 1. L'action du groupe affine sur \mathbb{H}^{d+1} .

Cette action est *isométrique* sur \mathbb{H}^{d+1} muni de la métrique hyperbolique $ds^2 = \frac{dx^2 + dy^2}{y^2}$. Cela définit une métrique invariante à gauche sur S_d . En particulier, la mesure de volume $dz = dx dy / y^{d+1}$ donne une mesure de Haar λ invariante à gauche sur S_d . Si on écrit dn et da/a pour les mesures de Haar sur $N \simeq \mathbb{R}^d$ et $A \simeq \mathbb{R}_+^*$ respectivement, la mesure de Haar à gauche sur S_d est donnée par

$$d\lambda(na) = a^{-(d+1)} dn da.$$

La mesure de Haar invariante à droite est alors donnée par

$$d\lambda_d(na) = a^{-1} dn da$$

et la fonction modulaire Δ de S_d par

$$\Delta(na) = \frac{d\lambda(na)}{d\lambda_d(na)} = a^{-d}.$$

En particulier, S_d n'est pas unimodulaire. On notera que si on utilise la décomposition $S_d = AN$, dans laquelle l'élément $g = an_b$ correspond au point (ab, a) de \mathbb{H}^{d+1} , les mesures de Haar sont données par

$$d\lambda(an) = a^{-1} da dn \quad \text{et} \quad d\lambda_d(an) = a^{d-1} da dn.$$

Pour $t \in \mathbb{R}$, on pose $a_t = (0, e^t)$. On utilisera, suivant les cas, les deux paramétrisations de S_d données par

$$(x, t) \in \mathbb{R}^d \times \mathbb{R} \mapsto n_x a_t = (x, e^t)$$

et

$$(x', t) \in \mathbb{R}^d \times \mathbb{R} \mapsto a_t n_{x'} = (e^t x', e^t),$$

pour lesquelles on a respectivement

$$d\lambda(n_x a_t) = e^{-dt} dx dt, \quad d\lambda_d(n_x a_t) = dx dt$$

et

$$d\lambda(a_t n_{x'}) = dt dx', \quad d\lambda_d(a_t n_{x'}) = e^{dt} dt dx'.$$

4.2. Suites de Følner dans le groupe des affinités

4.2.1. Étude de la famille des boules hyperboliques. — Nous commençons par démontrer que les boules B_r de S_d , centrées à l'élément neutre et de rayon r pour la métrique hyperbolique, ne peuvent pas former une famille (à un paramètre) de Følner à gauche sur S_d ⁽¹⁾. En effet, B_r s'identifie à la boule hyperbolique $B(o, r) = \{z \in \mathbb{H}^{d+1}; d(o, z) < r\}$. Cette boule est la boule euclidienne de centre $z_r = (0, \operatorname{ch} r)$ et de rayon $\operatorname{sh} r$ (voir la formule (D.3.1) de l'appendice D).

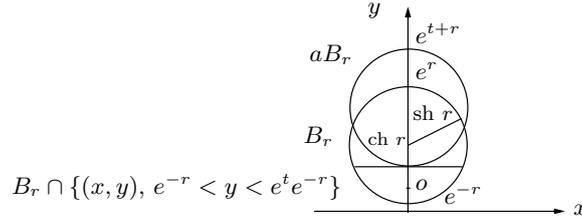


FIGURE 2. La boule $B(o, r)$.

Or, d'une part, lorsque r tend vers $+\infty$, le volume de la boule B_r est estimé par

$$(4.2.1) \quad \lambda(B_r) = c_d \int_0^r (\operatorname{sh} s)^d ds \sim c_d e^{dr}.$$

D'autre part, pour toute homothétie $a \in S_d$ de rapport $e^t > 1$, l'ensemble $B_r \setminus aB_r$ contient dans sa partie inférieure l'ensemble

$$B_r \cap \{(x, y); e^{-r} < y < e^t e^{-r}\}$$

dont le volume croît encore avec la variable r comme $c(t)e^{dr}$ pour une constante $c(t) > 0$. Par conséquent, le rapport $\frac{\lambda(B_r \setminus aB_r)}{\lambda(B_r)}$ ne tend pas vers 0 (cf. fig. 2).

On peut remarquer que, plus généralement, une suite (\mathcal{B}_n) de sous-ensembles radiaux de S_d ne peut former une suite de Følner sur S_d . En effet, le groupe des isométries directes de l'espace hyperbolique \mathbb{H}^{d+1} s'écrit $G = NAK$, où $K = SO(d+1)$ est le

⁽¹⁾Étant symétriques, elles ne forment pas non plus une famille de Følner à droite.

stabilisateur de o . Si (\mathcal{B}_n) était une suite de Følner à gauche de $S_d = NA$, formée de parties radiales, alors la suite $(\mathcal{B}_n K)$ serait une suite de Følner dans le groupe G , ce qui est impossible puisque le groupe simple G n'est pas moyennable.

4.2.2. Construction de suites de Følner. — Nous nous plaçons d'abord en dimension 2 (*i.e.* $d = 1$) et nous recherchons des suites de Følner formées de rectangles. Si A_1, A_2, B_1 et B_2 sont des réels donnés avec $0 < A_1 < A_2$ et $B_1 < B_2$, on note R_{A_1, A_2, B_1, B_2} le rectangle

$$R_{A_1, A_2, B_1, B_2} = \{(x, y) \in \mathbb{H}^2; B_1 \leq x \leq B_2, A_1 \leq y \leq A_2\}.$$

Comme pour les boules hyperboliques, une suite de rectangles ne peut former une suite de Følner à gauche. En effet, remarquons tout d'abord que l'image d'un rectangle par une affinité est encore un rectangle, puis choisissons une homothétie a de rapport strictement plus grand que 1. Pour tout rectangle $R = R_{A_1, A_2, B_1, B_2}$ avec $aA_1 \leq A_2$, l'ensemble $R \setminus aR$ contient alors (au bord près) le rectangle R_{A_1, aA_1, B_1, B_2} dont l'aire hyperbolique vaut $(B_2 - B_1) \left(\frac{1}{A_1} - \frac{1}{aA_1} \right)$. Cette aire ne peut être rendue négligeable relativement à l'aire totale de R , qui vaut $(B_2 - B_1) \left(\frac{1}{A_1} - \frac{1}{A_2} \right)$. D'autre part, si $aA_1 > A_2$, on a $R = R \setminus aR$.

Cependant, nous allons voir qu'il est possible de trouver des suites de rectangles formant des suites de Følner à droite. Notons tout d'abord que si $R = R_{A_1, A_2, B_1, B_2}$,

$$\lambda_d(R) = \int_{A_1}^{A_2} \frac{dy}{y} \int_{B_1}^{B_2} dx = (B_2 - B_1) \ln \frac{A_2}{A_1}.$$

Pour une affinité $g = (b, a)$, et pour $A_1^{-1}, A_2, |B_1|, |B_2|$ suffisamment grands, l'ensemble $Rg \cap R$, où Rg est l'image de R par l'action à droite de g sur R , est un quadrilatère Q , représenté en gris ci-dessous dans le cas $a > 1$ et $b > 0$,

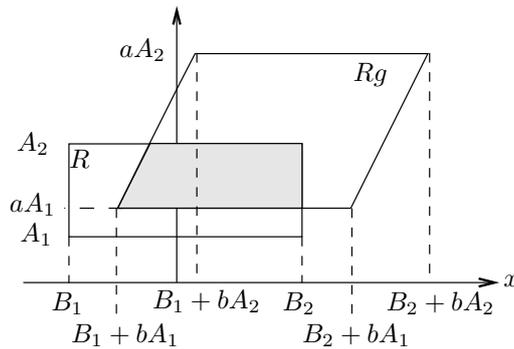


FIGURE 3. Un rectangle et son translaté.

On a dans ce cas,

$$\lambda_d(Q) = \int_{aA_1}^{A_2} \frac{dy}{y} \int_{B_1 + \frac{by}{a}}^{B_2} dx = \lambda_d(R) - (B_2 - B_1) \ln a - \frac{b}{a}(A_2 - aA_1).$$

Ceci implique

$$\frac{\lambda_d(Q)}{\lambda_d(R)} = 1 - \frac{\ln a}{\ln(A_2/A_1)} - \frac{b(A_2 - aA_1)}{a(B_2 - B_1) \ln(A_2/A_1)}.$$

Le calcul s'étend en dimension supérieure. Pour fixer les idées, dans ce chapitre nous choisissons comme norme sur $N \simeq \mathbb{R}^d$ la norme donnée par $\|x\| = \max_{1 \leq i \leq d} |x_i|$.

Proposition 4.2.1. — Soient $A \geq 1, B \geq 0$. Notons $P_{B,A}$ le parallélépipède de S_d défini par

$$P_{B,A} = \left\{ g = (b, a) \in S_d; \|b\| \leq B, \frac{1}{A} \leq a \leq A \right\}.$$

Alors la suite de parallélépipèdes $(P_{B_n, A_n})_{n \geq 1}$ forme une suite de Følner à droite sur S_d dès que $\lim_{n \rightarrow \infty} A_n = +\infty$ et que $\lim_{n \rightarrow \infty} \frac{A_n^d}{B_n^d \ln A_n} = 0$.

Notons qu'on a $P_{B,A} = R_{1/A, A, -B, B}$ pour $d = 1$. L'exemple le plus simple d'une suite satisfaisant à ces conditions est la suite de parallélépipèdes $P_{n,n}, n \in \mathbb{N}^*$, définie par

$$P_{n,n} = \left\{ (b, a) \in S_d; \|b\| \leq n, \frac{1}{n} \leq a \leq n \right\}.$$

et plus généralement la suite

$$P_{n^\beta, n} = \left\{ (b, a) \in S_d; \|b\| \leq n^\beta, \frac{1}{n} \leq a \leq n \right\}$$

lorsque $\beta \geq 1$.

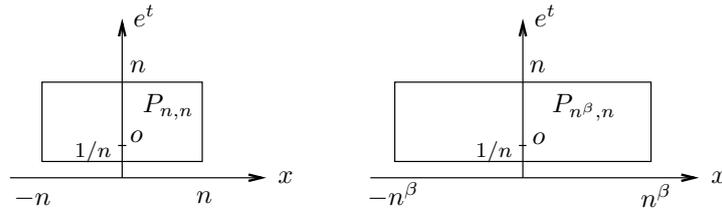


FIGURE 4

La suite croissante $(P_{n^\beta, n})$ n'est pas tempérée au sens de la définition 3.4.1 (adaptée au cas où la mesure de Haar est invariante à droite). En effet, considérons l'ensemble

$$\begin{aligned} E_n &= P_{(n+1)^\beta, n+1} (P_{n^\beta, n})^{-1} \\ &= \left\{ (b + ab', aa'); \|b\| \leq (n+1)^\beta, \|b'\| \leq n^\beta a', \frac{1}{n+1} \leq a \leq n+1, \frac{1}{n} \leq a' \leq n \right\}. \end{aligned}$$

On a l'inclusion suivante, pour tout $n \in \mathbb{N}^*$:

$$\{(ab', an); \|b'\| \leq n^{\beta+1}, 1 \leq a \leq n+1\} \subset E_n.$$

On en déduit, en utilisant l'invariance à droite de λ_d , que

$$\lambda_d(E_n) \geq |\mathcal{B}(0, n^{\beta+1})| \int_1^{n+1} a^d \frac{da}{a},$$

où $\mathcal{B}(0, r)$ désigne ici la boule de rayon r pour la norme $\|\cdot\|$ sur $N \simeq \mathbb{R}^d$ et $|\mathcal{B}(0, r)|$ désigne son volume euclidien. Comme son volume croît comme r^d , on obtient

$$\lambda_d(E_n) \geq cn^{d(\beta+1)}(n+1)^d$$

que l'on ne peut pas dominer par

$$\lambda_d(P_{(n+1)^\beta, n+1}) = c'(n+1)^{d\beta} \ln(n+1).$$

Ceci démontre l'assertion.

Remarque 4.2.2. — Terminons cette section en remarquant que d'après la proposition 4.2.1, pour tout $\beta \geq 1$, la suite $((P_{n^\beta, n})^{-1})_{n \geq 1}$ est une suite de Følner à gauche sur le groupe affine S_d . Mais elle n'est pas tempérée, comme on vient de le voir.

Dans la suite, on utilisera la famille de sous-ensembles $F_r^\beta = (P_{e^{\beta r}, e^r})^{-1}$, c'est-à-dire

$$F_r^\beta = \{n_x a_t; \|x\| \leq e^t e^{\beta r}, |t| \leq r\}.$$

Dans la décomposition $S_d = NA$ ce sont des "trapèzes", mais en utilisant les coordonnées (x', t) de S_d données par la décomposition $S_d = AN$, on obtient les parallélépipèdes (cf. fig. 5),

$$F_r^\beta = \{a_t n_{x'}; \|x'\| \leq e^{\beta r}, |t| \leq r\}.$$

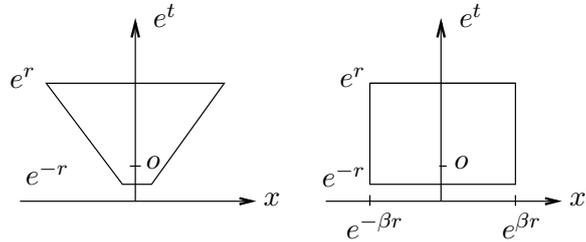


FIGURE 5. F_r^β en coordonnées (x, e^t) et en coordonnées (x', e^t)

4.3. Inégalités maximales pour les boules hyperboliques

Le but de cette section est de démontrer que l'opérateur maximal M^* pour les moyennes sur les translatés à gauche gB_r des boules de l'espace hyperbolique réel \mathbb{H}^{d+1} est borné sur L^p pour tout $1 < p \leq +\infty$ et de type faible $(1,1)$.

Comme nous n'avons jusqu'à présent considéré dans ce livre que des moyennes sur des translatés à droite, nous commençons par quelques remarques sur les différences entre moyennes à droite et moyennes à gauche.

4.3.1. Moyennes à droite et à gauche. — Étant donné un sous-ensemble mesurable E d'un groupe localement compact G tel que $0 < \lambda(E) < +\infty$, on peut considérer les deux opérateurs de moyenne

$$A_E \varphi(g) = \frac{1}{\lambda(Eg)} \int_{Eg} \varphi(g') d\lambda(g') = \frac{1}{\lambda(E)} \int_E \varphi(g'g) d\lambda(g')$$

et

$$M_E \varphi(g) = \frac{1}{\lambda(gE)} \int_{gE} \varphi(g') d\lambda(g') = \frac{1}{\lambda(E)} \int_E \varphi(gg') d\lambda(g').$$

Le premier opérateur est associé aux translatés à droite de E par l'action de G et le second opérateur est associé aux translatés à gauche de E par l'action de G .

Ce sont des opérateurs de convolution. Rappelons que si φ et ψ sont deux fonctions sur G et si Δ est la fonction modulaire de G , alors

$$\varphi * \psi(g) = \int_G \varphi(h) \psi(h^{-1}g) d\lambda(h) = \int_G \varphi(gh^{-1}) \psi(h) \Delta(h^{-1}) d\lambda(h).$$

Plus généralement pour toute mesure positive μ sur G , on a

$$\mu * \psi(g) = \int_G \psi(h^{-1}g) d\mu(h) \quad \text{et} \quad \varphi * \mu(g) = \int_G \varphi(gh^{-1}) \Delta(h^{-1}) d\mu(h).$$

Notons μ_E la mesure de probabilité $\frac{\mathbf{1}_E}{\lambda(E)} d\lambda$. Alors A_E est l'opérateur de convolution à gauche $\varphi \mapsto \check{\mu}_E * \varphi$ par la mesure de probabilité $\check{\mu}_E$ et M_E est l'opérateur de convolution à droite $\varphi \mapsto \varphi * (\Delta \check{\mu}_E)$ par la mesure $\Delta \check{\mu}_E$.

Pour un groupe unimodulaire G , l'application qui à une fonction φ associe la fonction $\check{\varphi}$ définie par $\check{\varphi}(g) = \varphi(g^{-1})$ est une isométrie de tous les espaces $L^p(G, \lambda)$. Si \mathcal{E} est une famille de parties mesurables de G , la relation

$$(4.3.1) \quad A_E \varphi(g^{-1}) = M_{E^{-1}} \check{\varphi}(g)$$

montre alors que les opérateurs maximaux $A_{\mathcal{E}}^*$ et $M_{\mathcal{E}^{-1}}^*$ associés respectivement à $\{A_E; E \in \mathcal{E}\}$ et à $\{M_E; E^{-1} \in \mathcal{E}\}$ sont simultanément bornés sur $L^p(G, \lambda)$.

Il n'en est pas de même sur un groupe non unimodulaire tel que S_d . En effet, l'égalité $d\check{\lambda}(g) = d\lambda_d(g) = \Delta(g^{-1}) d\lambda(g)$ montre qu'il faut introduire l'application

$$\mathcal{I}_p : \varphi \mapsto \Delta^{-\frac{1}{p}} \check{\varphi}$$

pour obtenir une isométrie de $L^p(G, \lambda)$.

Si μ est une mesure de probabilité sur G , nous noterons dans la suite M_μ l'opérateur défini par

$$M_\mu \varphi(g) = \int_G \varphi(gg') d\mu(g').$$

On a ainsi $M_E = M_{\mu_E}$. Dans un groupe non unimodulaire, la relation (4.3.1) doit être remplacée par

$$\mathcal{I}_p \circ A_E = M_{\mathcal{I}_p^*(\mu_E)} \circ \mathcal{I}_p$$

et l'étude des opérateurs de moyenne A_E se ramène à celle des opérateurs de convolution à droite relativement aux mesures $\mathcal{I}_p^*(\mu_E)$, $E \in \mathcal{E}$, où

$$\mathcal{I}_p^*(\mu_E)(dg) = \Delta^{\frac{1}{p}}(g) \check{\mu}_E(dg) = \frac{\mathbf{1}_{E^{-1}}(g)}{\lambda(E)} \Delta^{\frac{1}{p}-1}(g) d\lambda(g).$$

Comme on le voit, ces opérateurs diffèrent sensiblement des opérateurs de moyenne M_E .

Pour éviter toute ambiguïté, soulignons que dans la suite les opérateurs M^ et A^* seront toujours étudiés sur S_d muni de la mesure de Haar invariante à gauche λ .*

4.3.2. Inégalités maximales pour les boules hyperboliques. — Dans cette section, $B(z, r)$ désigne la boule hyperbolique de centre $z \in \mathbb{H}^{d+1}$ et de rayon $r > 0$ et $|B(z, r)|$ son volume hyperbolique. On note M_r l'opérateur de moyenne sur les boules de rayon r défini, pour toute fonction $\varphi \in L^1_{loc}(\mathbb{H}^{d+1}, dz)$, par

$$M_r \varphi(z) = \frac{1}{|B(z, r)|} \int_{B(z, r)} \varphi(z') dz'.$$

Lorsque \mathbb{H}^{d+1} est identifié avec S_d et le point $z = go$ avec l'affinité $g \in S_d$, la boule $B(z, r)$ s'identifie avec gB_r où $B_r = \{g' \in S_d; d(o, g'o) < r\}$, et l'opérateur de moyenne M_r apparaît comme l'opérateur de convolution à droite sur S_d :

$$M_r \varphi(g) = \frac{1}{\lambda(gB_r)} \int_{gB_r} \varphi(g') d\lambda(g') = \frac{1}{\lambda(B_r)} \int_{B_r} \varphi(gg') d\lambda(g').$$

L'opérateur maximal M^* est défini sur $L^1_{loc}(\mathbb{H}^{d+1}, dz) = L^1_{loc}(S_d, \lambda)$ par

$$M^* \varphi = \sup_{r>0} |M_r \varphi|.$$

L'inégalité L^p pour M^* a été démontrée par Clerc et Stein [CS74]. Par la suite Strömberg [Str81] a établi l'inégalité de type faible (1, 1). Évidemment l'inégalité L^p avec $1 < p < +\infty$ résulte par interpolation de l'inégalité de type faible (1, 1) et de l'inégalité triviale L^∞ . Nous commencerons néanmoins par la démonstration due à Clerc-Stein [CS74] de l'inégalité L^p ($1 < p < +\infty$), dans la mesure où elle met en lumière le rôle du phénomène de Kunze-Stein sur les espaces hyperboliques. On

donnera ensuite la démonstration de l'inégalité de type faible (1,1) due à Strömberg [Str81], telle qu'elle est présentée dans [ADY96].

Théorème 4.3.1. — *L'opérateur maximal M^* est de type faible (1,1) et il est borné sur L^p ($1 < p \leq \infty$).*

Pour démontrer ce théorème, on commence par analyser le comportement du volume des boules hyperboliques. Le volume de B_r vérifie

$$\lambda(B_r) \asymp \begin{cases} r^{d+1} & \text{pour } r \text{ petit} \\ e^{dr} & \text{pour } r \text{ grand} \end{cases}$$

où $f_1 \asymp f_2$ signifie que le rapport f_1/f_2 est borné supérieurement et inférieurement par des constantes strictement positives. Nous allons étudier séparément la partie locale

$$M_0^* \varphi(z) = \sup_{0 < r \leq 1} \frac{1}{\lambda(B_r)} \left| \int_{B(z,r)} \varphi(z') dz' \right|$$

et la partie à grande échelle

$$M_\infty^* \varphi(z) = \sup_{r > 1} \frac{1}{\lambda(B_r)} \left| \int_{B(z,r)} \varphi(z') dz' \right|$$

de la fonction maximale $M^* \varphi$. Comme

$$M^* \varphi \leq M_0^* \varphi + M_\infty^* \varphi,$$

les inégalités L^p et de type faible (1,1) pour $M^* \varphi$ résulteront de celles obtenues pour $M_0^* \varphi$ et $M_\infty^* \varphi$.

La partie locale s'étudie comme dans le cas euclidien traité dans le chapitre 3, puisque le volume possède la propriété de doublement local,

$$\lambda(B_{2r}) \leq c_R \lambda(B_r), \quad 0 < r \leq R.$$

Par contre, puisque le volume a une croissance exponentielle pour les grands rayons, la propriété de doublement du volume n'est plus satisfaite à l'infini. L'apport original du travail est la démonstration du fait que la fonction maximale $M_\infty^* \varphi$ est bornée sur $L^p(\mathbb{H}^{d+1}, dz)$ pour $1 < p < +\infty$ et de type faible (1,1). Dans une première étape, on va comparer ponctuellement $M_\infty^* \varphi(z)$ avec un opérateur de convolution à droite $\varphi \mapsto \varphi * \tau(z)$ sur S_d , dont le noyau τ est radial et possède une décroissance exponentielle précise,

$$\tau(g) \leq c \exp(-d\rho(g)),$$

où ρ désigne la distance à l'origine : $\rho(g) = d(o, go)$. Dans une seconde étape, on démontrera que de tels opérateurs de convolution sont bornés dans L^p ($1 < p \leq +\infty$) et de type faible (1,1).

Nous commençons par la comparaison avec l'opérateur de convolution.

Lemme 4.3.2. — Pour tout $\varphi \in L^1(S_d, \lambda)$, $\varphi \geq 0$, et pour tout $g \in S_d$, on a

$$M_\infty^* \varphi(g) \leq \varphi * \tau(g) = \int_{S_d} \varphi(gg') \tau(g') d\lambda(g'),$$

où $\tau(g) = \lambda(B_{\max(\rho(g), 1)})^{-1}$ vérifie

$$\tau(g^{-1}) = \tau(g) = O(\exp(-d\rho(g))), \quad \text{quand } \rho(g) \rightarrow \infty.$$

Démonstration. — On a l'inégalité évidente

$$M_\infty^* \varphi(g) = \sup_{r>1} \int_{S_d} \frac{\mathbf{1}_{B_r}(g')}{\lambda(B_r)} \varphi(gg') d\lambda(g') \leq \int_{S_d} \sup_{r>1} \frac{\mathbf{1}_{B_r}(g')}{\lambda(B_r)} \varphi(gg') d\lambda(g').$$

Pour tout $g' \in S_d$, il est clair que si $\rho(g') > 1$, alors

$$\frac{\mathbf{1}_{B_r}(g')}{\lambda(B_r)} \leq \frac{1}{\lambda(B_{\rho(g')})},$$

tandis que si $\rho(g') < 1$, on a

$$\frac{\mathbf{1}_{B_r}(g')}{\lambda(B_r)} \leq \frac{1}{\lambda(B_1)}$$

puisque $r > 1$. On obtient bien

$$\sup_{r>1} \frac{\mathbf{1}_{B_r}(g')}{\lambda(B_r)} \leq \lambda(B_{\max(\rho(g'), 1)})^{-1}.$$

L'estimation (4.2.1) du volume des grandes boules termine la démonstration. \square

Introduisons le noyau symétrique $k(g) = e^{-d\rho(g)}$ et posons

$$T\varphi(g) = \varphi * k(g) = \int_{S_d} \varphi(gg') e^{-d\rho(g')} d\lambda(g').$$

Théorème 4.3.3. — L'opérateur T est de type faible (1,1) et borné sur L^p ($1 < p < \infty$).

Le théorème 4.3.1 se déduit immédiatement du lemme 4.3.2 et du théorème 4.3.3.

Démonstration du théorème 4.3.3.

a) *L'inégalité L^p pour $1 < p < +\infty$.* — Elle va résulter du phénomène de Kunze-Stein sphérique transféré de \mathbb{H}^{d+1} à S_d .

Ce phénomène remarquable a été découvert par Kunze et Stein [KS60] pour le groupe $SL(2, \mathbb{R})$: si $1 \leq q < 2$, le produit de convolution par tout élément ψ de $L^q(SL(2, \mathbb{R}))$ est continu de $L^2(SL(2, \mathbb{R}))$ dans $L^2(SL(2, \mathbb{R}))$ ⁽²⁾. Par la suite, Stein [Ste70c] et Herz [Her70] ont établi la même propriété pour tout groupe de Lie G semi-simple, connexe, de centre fini, à condition de prendre ψ bi-invariante sous l'action d'un sous-groupe compact maximal. Finalement, Cowling [Cow78] a obtenu

⁽²⁾ On comparera ce résultat avec le fait que $L^q(\mathbb{R}^d) * L^2(\mathbb{R}^d) \subset L^2(\mathbb{R}^d)$ si et seulement si $q = 1$.

ce résultat pour tout ces groupes, sans la restriction de bi-invariance (voir [Cow97], pour un article de synthèse).

Désignons par $L^{q,\sharp}(S_d)$ le sous-espace des fonctions radiales dans $L^q(S_d)$. D'après ce que nous avons rappelé, et avec un argument supplémentaire d'interpolation, nous voyons que pour $1 \leq q < p \leq 2$, il existe $c > 0$ tel que, pour tous $\varphi \in L^p(S_d)$ et $\psi \in L^{q,\sharp}(S_d)$, on a

$$(4.3.2) \quad \|\varphi * \psi\|_p \leq c \|\varphi\|_p \|\psi\|_q.$$

Nous appliquons ce résultat au noyau radial $\psi = k$ ci-dessus. On vérifie facilement que k appartient à tous les espaces $L^{q,\sharp}(S_d)$ ($1 < q \leq +\infty$). En effet, la mesure invariante à gauche s'écrit

$$d\lambda = (\operatorname{sh} \rho)^d d\rho d\sigma$$

en coordonnées polaires, où $d\sigma$ est la mesure usuelle sur la sphère unité (voir (D.1.1) dans l'appendice D). La quantité

$$\|k\|_q^q = c \int_0^\infty \exp(-q\rho d) (\operatorname{sh} \rho)^d d\rho$$

est finie si et seulement si $q > 1$.

Fixons $p \in]1, 2]$. Soit q tel que $1 < q < p$. Comme $k \in L^{q,\sharp}(S_d)$, l'opérateur T est borné sur $L^p(S_d)$ d'après l'inégalité (4.3.2). Par ailleurs, le noyau k est symétrique. Par dualité et symétrie de l'opérateur, on voit que T est borné sur $L^{p'}(S_d)$ pour tout $2 \leq p' < +\infty$. Ceci termine la démonstration de l'inégalité L^p ($1 < p < \infty$). \square

b) *L'inégalité faible* (1, 1). — Pour la démontrer, nous allons utiliser les deux paramétrisations

$$g = n_x a_t = a_t n_{x'} \text{ avec } t \in \mathbb{R}, x, x' \in \mathbb{R}^d$$

de $S_d = NA = AN$. Rappelons que dans ces coordonnées, la mesure de Haar invariante à gauche s'écrit $d\lambda(g) = e^{-dt} dx dt = dt dx'$.

On a la majoration

$$(4.3.3) \quad e^{-d\rho(g)} \leq c_d e^{-dt} P_1(x'),$$

où P_1 est le noyau de Poisson euclidien de \mathbb{R}^d :

$$P_1(x') = c'_d \frac{1}{(1 + |x'|^2)^d}, \quad x' \in \mathbb{R}^d.$$

Ici $|x'|$ désigne la norme euclidienne de $x' \in \mathbb{R}^d$. Cette majoration résulte de la formule suivante (voir (D.3.2)) : pour tout $g \in G$ avec $g = a_t n_{x'}$, on a

$$\operatorname{ch} \rho(g) = \operatorname{ch} t + \frac{e^t}{2} |x'|^2.$$

On est amené naturellement à introduire les deux opérateurs suivants :

$$\Pi\varphi(g) = \int_{\mathbb{R}^d} \varphi(gn_x) P_1(x) dx \quad \text{et} \quad \Lambda\varphi(g) = \int_{\mathbb{R}} \varphi(ga_t) e^{-dt} dt.$$

Ce sont des opérateurs de convolution à droite, le premier par le noyau de Poisson $P_1(x)$ sur la composante \mathbb{R}^d de S_d et le second par la mesure de Lebesgue à poids e^{-dt} sur la composante \mathbb{R} de S_d .

Lemme 4.3.4. — *L'opérateur Π est une contraction sur $L^p(S_d, \lambda)$ pour tout $1 \leq p \leq +\infty$.*

Démonstration. — Cela résulte du fait que $P_1(x)dx$ est une mesure de probabilité sur \mathbb{R}^d et que la mesure de Haar invariante à gauche λ est aussi invariante par les translations à droite par les éléments de N . \square

Pour l'opérateur Λ , nous avons le résultat suivant :

Lemme 4.3.5. — *L'opérateur Λ est de type faible (1,1). Plus précisément, pour toute fonction $\varphi \in L^1(S_d, \lambda)$ et pour tout $\alpha > 0$, on a*

$$\lambda(\{|\Lambda\varphi| \geq \alpha\}) \leq \frac{1}{d} \frac{\|\varphi\|_1}{\alpha}.$$

Démonstration. — Revenons à la paramétrisation $g = n_x a_s$ de $S_d = NA$. Soit φ une fonction positive dans $L^1(S_d, \lambda)$. Un calcul simple établit la formule

$$\Lambda\varphi(n_x a_s) = e^{ds} \Lambda\varphi(n_x),$$

qui sera cruciale pour obtenir l'inégalité faible. Pour $\alpha > 0$, on pose

$$E_\alpha = \{g \in S_d; \Lambda\varphi(g) \geq \alpha\}.$$

On a, d'après la formule précédente,

$$E_\alpha = \{g = n_x a_s \in S_d; x \in \mathbb{R}^d \text{ et } s \geq s_0(n_x)\}$$

avec

$$s_0(n_x) = \begin{cases} \frac{1}{d} (\ln \alpha - \ln \Lambda\varphi(n_x)) & \text{si } \Lambda\varphi(n_x) > 0 \\ +\infty & \text{si } \Lambda\varphi(n_x) = 0. \end{cases}$$

On en déduit que

$$\begin{aligned} \lambda(E_\alpha) &= \int_{S_d} \mathbf{1}_{E_\alpha}(g) d\lambda(g) = \int_{\mathbb{R}^d} dx \left(\int_{s_0(n_x)}^{+\infty} e^{-ds} ds \right) = \frac{1}{d} \int_{\mathbb{R}^d} e^{-ds_0(n_x)} dx \\ &= \frac{1}{d\alpha} \int_{\mathbb{R}^d} \Lambda\varphi(n_x) dx = \frac{1}{d\alpha} \int_{\mathbb{R}^d} \int_{\mathbb{R}} \varphi(n_x a_t) e^{-dt} dt dx = \frac{1}{d\alpha} \|\varphi\|_1. \end{aligned}$$

L'inégalité de type faible (1,1) pour tout $\varphi \in L^1(S_d, \lambda)$ en résulte immédiatement. \square

Pour terminer, montrons que l'opérateur T est de type faible (1,1). Soit φ une fonction positive dans $L^1(S_d, \lambda)$ et soit $g \in S_d$. Dans l'intégrale

$$T\varphi(g) = \int_{S_d} \varphi(gg') e^{-d\rho(g')} d\lambda(g'),$$

écrivons $g' = a_t n_{x'}$ et majorons l'exponentielle à l'aide de l'inégalité (4.3.3) :

$$T\varphi(g) \leq c_d \int_{\mathbb{R}} \left(\int_{\mathbb{R}^d} \varphi(ga_t n_{x'}) P_1(x') dx' \right) e^{-dt} dt \leq c_d (\Lambda \circ \Pi) \varphi(g).$$

D'après les lemmes 4.3.4 et 4.3.5, l'opérateur $\Lambda \circ \Pi$ est de type faible (1, 1). Il en est donc de même pour l'opérateur T . Ceci conclut la démonstration du théorème 4.3.3. \square

4.4. Inégalités maximales pour des familles plus générales

Dans cette section nous considérons des familles d'ensembles dans S_d qui contiennent comme cas extrêmes les parallélépipèdes et les trapèzes de la section 4.2 et comme cas intermédiaire les boules de la section 4.3 (en fait des ensembles comparables à ces boules). Comme l'analyse à petite échelle est semblable au cas euclidien, nous nous limiterons dans cette section à des ensembles de grande taille, qui révèlent des phénomènes nouveaux dans le cas résoluble. Plus précisément, étant donné $0 \leq \epsilon \leq 1$ et $\beta > 0$, introduisons pour tout $r > 1$ les ensembles

$$\begin{aligned} F_r^{\epsilon, \beta} &= \{n_x a_t \in S_d; \|x\| \leq e^{t+\beta r}, |t| \leq r\} \\ &= \{a_t n_{x'} \in S_d; \|x'\| \leq e^{(\epsilon-1)t+\beta r}, |t| \leq r\}. \end{aligned}$$

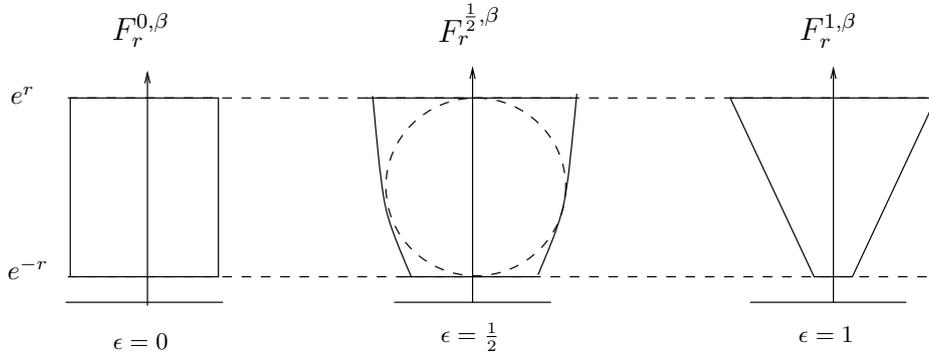


FIGURE 6. Les ensembles $F_r^{\epsilon, \beta}$ en coordonnées (x, e^t)

Remarque 4.4.1. — Pour $\epsilon = 0$, on retrouve la famille de parallélépipèdes $P_{B,A}$ avec $A = e^r$ et $B = e^{\beta r}$ considérée dans la proposition 4.2.1, tandis que pour $\epsilon = 1$, il s'agit de leurs inverses F_r^β introduits à la fin de la section 4.2. Notons plus généralement que $(F_r^{\epsilon, \beta})^{-1} = F_r^{1-\epsilon, \beta}$ et qu'en particulier les ensembles $F_r^{\epsilon, \beta}$ sont symétriques pour $\epsilon = 1/2$. De plus, pour $\epsilon = \beta = 1/2$, les ensembles $F_r^{\epsilon, \beta}$ sont

comparables aux boules hyperboliques ; plus précisément il existe une constante $c > 0$ telle que $B_r \subset F_r^{1/2, 1/2} \subset B_{r+c}$.

Évaluons la mesure des ensembles $F_r^{\epsilon, \beta}$:

$$(4.4.1) \quad \lambda(F_r^{\epsilon, \beta}) = \begin{cases} \frac{2}{(1-\epsilon)d} \operatorname{sh}((1-\epsilon)dr) e^{\beta dr} & \text{si } 0 \leq \epsilon < 1, \\ 2r e^{\beta dr} & \text{si } \epsilon = 1, \end{cases}$$

$$(4.4.2) \quad \asymp \begin{cases} e^{(1-\epsilon+\beta)dr} & \text{si } 0 \leq \epsilon < 1, \\ r e^{\beta dr} & \text{si } \epsilon = 1. \end{cases}$$

Étant donnés ϵ et β , notons M_∞^* l'opérateur maximal à grande échelle (*i.e.* pour lequel on se limite aux réels $r > 1$) associé aux moyennes à gauche $M_{F_r^{\epsilon, \beta}}$ et notons A_∞^* l'opérateur maximal à grande échelle associé aux moyennes à droite $A_{F_r^{\epsilon, \beta}}$. Les résultats présentés dans les quatre sous-sections suivantes sont dus à Gaudry, Giulini, Mantero et Sjögren [**GGM90**, **GS90**] et sont complétés par le théorème 4.4.5, qui est par contre nouveau. On notera la dissymétrie entre les opérateurs A_∞^* et M_∞^* , ainsi que la différence entre les cas $\epsilon < 1$ et $\epsilon = 1$, déjà apparent dans le comportement (4.4.1) du volume.

4.4.1. Inégalités maximales pour M_∞^* . —

Théorème 4.4.2 ([**GGM90**]). —

- (i) Si $0 \leq \epsilon < 1/2$, l'opérateur M_∞^* est borné sur L^p pour tout $p \geq 1$.
- (ii) Si $1/2 < \epsilon < 1$, l'opérateur M_∞^* n'est borné sur aucun L^p avec $1 \leq p < 1/2(1-\epsilon)$.
- (iii) Si $\epsilon = 1$, l'opérateur M_∞^* n'est borné sur aucun L^p avec $1 \leq p < +\infty$.

Remarque 4.4.3. — Compte-tenu du comportement à petite échelle, semblable au cas euclidien, l'opérateur maximal global M^* est seulement de type faible (1, 1) si $0 \leq \epsilon < 1/2$.

Démonstration. — (i) On reprend les arguments déjà utilisés dans la démonstration du théorème 4.3.1. Commençons par majorer, pour toute fonction mesurable positive φ ,

$$M_\infty^* \varphi(g) \leq \int_{S_d} \varphi(gg') \tau(g') d\lambda(g'),$$

où

$$\tau(g) = \sup_{r>1} \lambda(F_r^{\epsilon, \beta})^{-1} \mathbf{1}_{F_r^{\epsilon, \beta}}(g).$$

Ceci ramène l'étude de M_∞^* à celle d'un opérateur de convolution sur S_d . Le résultat suivant, valable pour tout groupe G localement compact, se démontre aisément à l'aide du théorème de Fubini.

Lemme 4.4.4. — Si $\tau : G \rightarrow \mathbb{R}$ est intégrable pour la mesure de Haar invariante à droite λ_d ou, de façon équivalente, si $\check{\tau} \in L^1(G, \lambda)$, alors l'opérateur de convolution à droite $\varphi \mapsto \int_G \varphi(\cdot g') \tau(g') d\lambda(g')$ est borné dans $L^1(G, \lambda)$.

Il suffit donc de vérifier que le noyau τ est intégrable relativement à λ_d pour obtenir que M_∞^* est borné dans L^1 , et par suite dans tout L^p .

Pour $g \in S_d$, posons $R(g) = \inf \{ r > 1 ; g \in F_r^{\epsilon, \beta} \}$. On a alors $\tau(g) = \lambda(F_{R(g)}^{\epsilon, \beta})^{-1}$. Un calcul simple donne, pour $g = n_x a_t \in S_d$,

$$R(g) = \max \left\{ 1, |t|, \frac{\ln \|x\| - \epsilon t}{\beta} \right\} = \max \{ 1, |t|, \beta^{-1} \ln \|z\| \},$$

après avoir effectué le changement de variable $x = e^{\epsilon t} z$.

En coordonnées (z, t) la mesure de Haar invariante à droite $dx dt$ s'écrit $e^{\epsilon dt} dz dt$. L'intégrale de τ par rapport à cette mesure, sur l'ensemble des $(z, t) \in \mathbb{R}^d \times \mathbb{R}$ où $R(z, t) = 1$, est finie. Notons Ω_1 l'ensemble des (z, t) où $R(z, t) = |t|$ et Ω_2 celui où $R(z, t) = \beta^{-1} \ln \|z\|$. En utilisant l'expression (4.4.1) du volume de $F_r^{\epsilon, \beta}$, on obtient d'une part

$$\begin{aligned} \int_{\Omega_1} \tau(g) d\lambda_d(g) &= c \int_1^\infty \frac{\text{ch}(\epsilon dt)}{\text{sh}((1-\epsilon)dt)} e^{-\beta dt} \left(\int_{\|z\| \leq e^{\beta t}} dz \right) dt \\ &\asymp \int_1^\infty e^{(2\epsilon-1)t} dt \end{aligned}$$

et d'autre part

$$\begin{aligned} \int_{\Omega_2} \tau(g) d\lambda_d(g) &= c \left(\int_0^1 \int_{\|z\| \geq e^{\beta t}} + \int_1^\infty \int_{\|z\| \geq e^{\beta t}} \right) \frac{\text{ch}(\epsilon dt)}{\|z\|^d \text{sh}((1-\epsilon)d\beta^{-1} \ln \|z\|)} dz dt \\ &\asymp \int_0^\infty e^{(2\epsilon-1)t} dt. \end{aligned}$$

On en déduit que $\tau \in L^1(S_d, \lambda_d)$ si et seulement si $0 \leq \epsilon < 1/2$.

Détaillons maintenant la démonstration de (iii). Celle de (ii) est similaire (voir [GGM90, Theorem 3]).

Par hypothèse $\epsilon = 1$. Pour tout $1 \leq p < +\infty$, on va construire une fonction $\varphi \in L^p(S_d, \lambda)$ telle que $M_\infty^* \varphi$ ne soit pas de puissance p -ième intégrable. On utilise la paramétrisation $(x, t) \mapsto n_x a_t$, pour laquelle la mesure de Haar invariante à gauche est $e^{-dt} dx dt$. La fonction φ , définie par

$$\varphi(x, t) = \mathbf{1}_{[-1, 0]}(t) \mathbf{1}_{\{\|x\| \geq 1\}}(x) \|x\|^{-\delta},$$

appartient à $L^p(S_d, \lambda)$ si et seulement si $\delta p > d$.

Considérons les translatés à gauche gF_r^β des ensembles $F_r^\beta = F_r^{1, \beta}$ par des éléments $g = n_x a_{-t} \in S_d$, avec $r > 1$, $t > 1$ et $\|x\| \geq 2e^{\beta t}$. Comme

$$gF_r^\beta = \{ n_y a_s ; \|y - x\| \leq e^{\beta r + s}, |s + t| \leq r \}$$

et $\lambda(gF_r^\beta) = \lambda(F_r^\beta)$ est proportionnel à $r e^{\beta dr}$, on a

$$\begin{aligned} M_\infty^* \varphi(g) &= C \sup_{r>1} \frac{1}{r e^{\beta dr}} \int_{-t-r}^{-t+r} \int_{\{\|y-x\| \leq e^{\beta r+s}\}} \varphi(y, s) e^{-ds} dy ds \\ &\geq C t^{-1} e^{-\beta dt} \int_0^{2t} \int_{\{\|y-x\| \leq e^{\beta t-s}\}} \varphi(y, -s) e^{ds} dy ds \end{aligned}$$

en prenant $r = t$ et en changeant s en $-s$. Comme $[0, 2t] \supset [0, 1]$ et que

$$\begin{cases} \|y\| \leq \|x\| + \|y-x\| \leq \|x\| + e^{\beta t} \leq \frac{3}{2} \|x\|, \\ \|y\| \geq \|x\| - \|y-x\| \geq \|x\| - e^{\beta t} \geq \frac{1}{2} \|x\| \geq e^{\beta t} \geq 1, \end{cases}$$

on aboutit à la minoration ponctuelle

$$M_\infty^* \varphi(g) \geq C t^{-1} e^{-\beta dt} \int_0^1 \int_{\{\|y-x\| \leq e^{\beta t-1}\}} \|y\|^{-\delta} dy ds \geq C t^{-1} \|x\|^{-\delta}.$$

On en déduit la minoration L^p

$$\begin{aligned} \|M_\infty^* \varphi\|_p^p &\geq C \int_1^{+\infty} \int_{\{\|x\| \geq 2e^{\beta t}\}} t^{-p} \|x\|^{-\delta p} e^{dt} dx dt \\ &\geq C \int_1^{+\infty} t^{-p} e^{dt} e^{(d-\delta p)\beta t} dt. \end{aligned}$$

La dernière intégrale diverge si $\delta p < (1 + 1/\beta)d$. Pour conclure, on choisit $\delta > 0$ en sorte que $d < \delta p < (1 + 1/\beta)d$. \square

Le résultat suivant, qui est nouveau, précise le comportement de M_∞^* dans le cas $\epsilon = 1/2$.

Théorème 4.4.5. — *Supposons $\epsilon = 1/2$ et soit $\beta > 0$. Alors l'opérateur maximal M_∞^* est borné sur L^p pour tout $1 < p \leq +\infty$.*

Démonstration. — Comme dans la démonstration des théorèmes 4.3.1 et 4.4.2, on commence par majorer, pour toute fonction mesurable positive φ ,

$$M_\infty^* \varphi(g) \leq \int_{S_d} \varphi(gg') \tau(g') d\lambda(g'),$$

où $\tau(h) = \lambda(F_{R(h)}^{1/2, \beta})^{-1}$ et $R(h) = \inf \{ r > 1; h \in F_r^{1/2, \beta} \}$. On obtient l'estimation suivante du noyau, en combinant le comportement (4.4.1) du volume avec l'expression $R(h) = \max \{ 1, |t|, \frac{\ln \|x\| + t/2}{\beta} \}$ en coordonnées $h = a_t n_x$:

$$(4.4.3) \quad \tau(a_t n_x) \leq \begin{cases} C e^{-(\beta + \frac{1}{2})d|t|} & \text{si } \|x\| \leq e^{\beta|t| - \frac{t}{2}} \\ C e^{-\frac{\beta+1/2}{2\beta}dt} \|x\|^{-\frac{\beta+1/2}{\beta}d} & \text{si } \|x\| \geq e^{\beta|t| - \frac{t}{2}} \end{cases}$$

i.e.

$$(4.4.4) \quad \tau(a_t n_x) \leq C \left(e^{\beta|t|} + e^{\frac{t}{2}} \|x\| \right)^{-(1 + \frac{1}{2\beta})d}.$$

Décomposons

$$M_\infty^* \varphi \leq \varphi * \check{\tau}_+ + \varphi * \check{\tau}_-,$$

où $\tau_\pm(a_t n_x) = \mathbf{1}_{\mathbb{R}_\pm}(t) \tau(a_t n_x)$. En utilisant les estimations précédentes, on vérifie facilement que le noyau τ_- est intégrable sur S_d pour la mesure de Haar invariante à droite $e^{dt} dx dt$. D'après le lemme 4.4.4, on en déduit que l'opérateur de convolution à droite par $\check{\tau}_-$ est borné dans L^1 , et par conséquent dans tous les espaces L^p . L'opérateur de convolution à droite par $\check{\tau}_+$ est plus délicat à étudier. À cet effet on réécrit l'estimation (4.4.4) pour τ_+ sous la forme suivante :

$$\tau_+(a_t n_x) \leq C e^{-dt} q_{u_t}(x),$$

où

$$q_u(x) = u^{-d} (1 + u^{-1} \|x\|)^{-(1+1/2\beta)d}$$

et $u_t = e^{(\beta-1/2)t}$. Notons que $q_u(x) = u^{-d} q_1(u^{-1}x)$ et que $u_t = 1$ si $\beta = 1/2$. On poursuit en estimant

$$(\varphi * \tau_+)(g) = \int_{S_d} \varphi(g a_t n_x) \tau_+(a_t n_x) dx dt \leq C (\Lambda' \circ \Pi') \varphi(g),$$

où

$$\Lambda' \varphi(g) = \int_0^{+\infty} e^{-td} \varphi(g a_t) dt \quad \text{et} \quad \Pi' \varphi(g) = \sup_{u>0} \int_{\mathbb{R}^d} \varphi(g n_x) q_u(x) dx.$$

On conclut comme dans la démonstration du théorème 4.3.3. L'opérateur maximal Π' , associé à la convolution à droite par la famille q_u , est de type faible $(1, 1)$ et donc borné sur L^p pour tout $1 < p \leq +\infty$. L'opérateur Λ' est aussi de type faible $(1, 1)$ (voir lemme 4.3.5) et donc borné sur L^p pour tout $1 < p \leq +\infty$. Par composition, on conclut que $\Lambda' \circ \Pi'$ est borné sur L^p pour tout $1 < p \leq +\infty$. Ceci termine la démonstration du théorème 4.4.5. \square

Remarque 4.4.6. — Le tableau ci-dessous fait la synthèse des résultats connus pour l'opérateur maximal M_∞^* associé aux translatés à gauche des ensembles $F_r^{\epsilon, \beta}$:

M_∞^*	$0 \leq \epsilon < \frac{1}{2}$	$\epsilon = \frac{1}{2}$	$\frac{1}{2} < \epsilon < 1$	$\epsilon = 1$
$L^1 \rightarrow L^1$	borné	type faible si $\beta = \frac{1}{2}$	non de type faible	non de type faible
$L^p \rightarrow L^p$ ($1 < p < \infty$)	borné	borné	non borné si $p < \frac{1}{2(1-\epsilon)}$	non borné

À l'exception du cas $\epsilon = 1/2$, tous ces résultats ont été établis dans [GGM90]. Si $\epsilon = \beta = 1/2$, les ensembles $F_r^{\epsilon, \beta}$ sont comparables aux boules hyperboliques B_r , comme signalé dans la remarque 4.4.1 ; dans ce cas, l'inégalité forte (p,p) avec $p > 1$ remonte à [CS74] et l'inégalité faible (1,1) a été obtenue ultérieurement dans [Str81]. Pour $\epsilon = 1/2$ et $\beta \neq 1/2$, l'inégalité forte pour $p > 1$ est démontrée pour la première fois dans ce livre (théorème 4.4.5).

Remarque 4.4.7. — Les problèmes suivants sont ouverts :

- Pour $\epsilon = 1/2$ et $\beta \neq 1/2$, M_∞^* est-il de type faible (1,1) ?
- Pour $1/2 < \epsilon < 1$, M_∞^* est-il de type faible (p_ϵ, p_ϵ) , où $p_\epsilon = 1/2(1-\epsilon)$? Plus généralement, quel est le comportement de M_∞^* sur les espaces L^p avec $p_\epsilon \leq p < +\infty$?

4.4.2. Inégalités maximales pour A_∞^* . —

Théorème 4.4.8 ([GGM90]). —

- (i) Si $\epsilon = 1$, l'opérateur A_∞^* est borné dans L^p pour tout $p > 1$.
- (ii) Si $0 \leq \epsilon < 1$, l'opérateur A_∞^* n'est borné dans aucun L^p avec $1 \leq p < +\infty$.

Démonstration. — (i) Par hypothèse $\epsilon = 1$. Rappelons que, dans les coordonnées $(x, t) \mapsto a_t n_x$, la mesure de Haar à gauche sur S_d coïncide avec la mesure de Lebesgue $dx dt$ sur $\mathbb{R}^d \times \mathbb{R}$ et que

$$F_r^{1,\beta} = F_r^\beta = \{ a_s n_y ; |s| \leq r, \|y\| \leq e^{\beta r} \}$$

est un parallélépipède centré en 0. Son translaté à droite

$$F_r^\beta g = \{ a_s n_y ; |s - t| \leq r, \|y - x\| \leq e^{\beta r - t} \}$$

par $g = a_t n_x$ est le parallélépipède de centre (x, t) , de largeur $2e^{\beta r - t}$ et de hauteur $2r$. Pour toute fonction mesurable positive φ sur S_d , on a

$$\begin{aligned} A_{F_r^\beta} \varphi(g) &= \frac{1}{2r |\mathcal{B}(x, e^{\beta r - t})|} \int_{t-r}^{t+r} \int_{\mathcal{B}(x, e^{\beta r - t})} \varphi(t', x') dt' dx' \\ &\leq \sup_{r>1} \frac{1}{2r} \int_{t-r}^{t+r} \left(\sup_{R>0} \frac{1}{|\mathcal{B}(x, R)|} \int_{\mathcal{B}(x, R)} |\varphi(t', x')| dx' \right) dt', \end{aligned}$$

où $|\mathcal{B}(x, R)|$ désigne le volume euclidien de la boule $\mathcal{B}(x, R)$ dans \mathbb{R}^d . La dernière expression est la composition de deux opérateurs maximaux de Hardy-Littlewood, l'un sur \mathbb{R} et l'autre sur \mathbb{R}^d . Comme ces deux opérateurs sont bornés sur $L^p(\mathbb{R})$ et sur $L^p(\mathbb{R}^d)$ respectivement, A_∞^* est borné sur $L^p(S_d, \lambda)$, pour tout $1 < p \leq +\infty$.

La démonstration de (ii) est similaire à celle de (iii) dans le théorème 4.4.2 (voir [GGM90, Theorem 5]). \square

Remarque 4.4.9. — Le cas $\epsilon = 1/2$ et $\beta = 1/2$ revient à étudier l'opérateur de moyenne A_∞^* sur les translatés à droite des boules hyperboliques. Contrairement à l'opérateur de moyenne sur leurs translatés à gauche étudié dans le théorème 4.3.1, A_∞^* n'est borné dans aucun espace L^p avec $1 \leq p < +\infty$.

Dans le cas $\epsilon = 1$ se pose la question de l'inégalité faible (1,1) pour l'opérateur A_∞^* . La réponse a été obtenue par Giuliani et Sjögren.

Théorème 4.4.10 ([GS90]). — L'opérateur maximal A_∞^* associé aux translatés à droite des ensembles $F_r^\beta = F_r^{1,\beta}$ est de type faible (1,1) si et seulement si $\beta > 1$.

Démonstration. — Nous conservons les notations utilisées dans la démonstration précédente.

Cas $\beta \leq 1$. Considérons la fonction indicatrice φ d'un ensemble compact de mesure positive dans S_d , par exemple

$$U = \{a_s n_y ; 0 \leq s \leq 1, \|y\| \leq 1\}.$$

On va montrer que

$$\sup_{\alpha > 0} \alpha \lambda(\{A_\infty^* \varphi \geq \alpha\}) = +\infty$$

et donc que l'opérateur maximal A_∞^* n'est pas de type faible (1, 1).

Soit $g = a_t n_x$ avec $t \geq 1$. Remarquons que, pour tout $r \geq 1$ tel que $e^{\beta r - t} \geq 1 + \|x\|$, on a $U \subset F_r^\beta g$. En effet cette condition implique $t \leq \beta r \leq r$ (puisque $\beta \leq 1$) ; alors, si $h = a_s n_y \in U$, on a $|s - t| = t - s \leq t \leq \beta r \leq r$ et $\|y - x\| \leq 1 + \|x\| \leq e^{\beta r - t}$, d'où l'inclusion.

Notons $R(g) = \frac{1}{\beta} (t + \ln(1 + \|x\|)) \geq 1$ le plus petit rayon r tel que $e^{\beta r - t} \geq 1 + \|x\|$. Dans la suite nous ne considérons que des éléments $g = a_t n_x$ tels que $t \geq 1$ et $\|x\| \geq e$, ce choix de x visant seulement à alléger les calculs. Nous obtenons alors

$$A_\infty^* \varphi(g) \geq \frac{\lambda(U)}{\lambda(F_{R(g)}^\beta g)} = \frac{c_1}{(t + \ln(1 + \|x\|))(1 + \|x\|)^d} \geq \frac{c}{t \|x\|^d \ln \|x\|}.$$

Nous introduisons la fonction strictement croissante $\Psi(r) = r^d \ln r$ sur $[1, +\infty[$ et, pour $\alpha > 0$, nous posons

$$D_\alpha = \left\{ a_t n_x ; \|x\| \geq e, 1 \leq t \leq \frac{c}{\alpha \Psi(\|x\|)} \right\} \subset \{A_\infty^* \varphi \geq \alpha\}.$$

Pour que D_α ne soit pas vide, il est nécessaire que $e \leq \Psi^{-1}(c/\alpha)$ et donc que $0 < \alpha \leq c/\Psi(e)$. En prenant α tel que $0 < \alpha \leq c/(2\Psi(e))$, nous obtenons la minoration

$$\begin{aligned} \lambda(D_\alpha) &= \int_{\{e \leq \|x\| \leq \Psi^{-1}(c/\alpha)\}} \left(\frac{c}{\alpha \Psi(\|x\|)} - 1 \right) dx \\ &\geq \int_{\{e \leq \|x\| \leq \Psi^{-1}(c/2\alpha)\}} \frac{c}{2\alpha \Psi(\|x\|)} dx. \end{aligned}$$

Nous en déduisons que

$$\begin{aligned} \sup_{\alpha > 0} \alpha \lambda(\{A_\infty^* \varphi \geq \alpha\}) &\geq \sup_{\alpha > 0} \alpha \lambda(D_\alpha) \\ &\geq \lim_{\alpha \rightarrow 0} \int_{\{e \leq \|x\| \leq \Psi^{-1}(c/2\alpha)\}} \frac{c}{2\alpha \Psi(\|x\|)} dx = c_2 \int_e^\infty \frac{dr}{r \ln r} = +\infty. \end{aligned}$$

Ceci termine la démonstration du cas $\beta \leq 1$.

Cas $\beta > 1$. La démonstration initiale de Giulini et Sjögren [GS90] a été simplifiée comme suit par Vallarino [Val06]. Etant donné $\gamma \geq 3 \frac{\beta+1}{\beta-1}$, on considère

les épaisissements suivants des ensembles $F_r^\beta = F_r^{1,\beta}$ et de leurs translatés à droite par $g = a_t n_x$:

$$\begin{aligned} (F_r^\beta)^\sim &= \{ a_s n_y ; |s| \leq \gamma r, \|y\| \leq \gamma e^{\beta r} \}, \\ (F_r^\beta g)^\sim &= (F_r^\beta)^\sim g = \{ a_s n_y ; |s-t| \leq \gamma r, \|y-x\| \leq \gamma e^{\beta r-t} \}. \end{aligned}$$

Notons que leurs mesures sont proportionnelles à celles des ensembles initiaux, avec un facteur de proportionnalité polynomial en γ :

$$\lambda((F_r^\beta g)^\sim) = (2\gamma)^{d+1} r e^{d(\beta r-t)} = \gamma^{d+1} \lambda(F_r^\beta g).$$

Dans le cas des boules d'un espace métrique, si $B(x_1, r_1) \cap B(x_2, r_2) \neq \emptyset$ et $r_1 \leq r_2$, alors $B(x_1, r_1) \subset B(x_2, 3r_2)$. La propriété suivante est un substitut à cette propriété :

$$(4.4.5) \quad \begin{aligned} &\text{Soient } F_1 = F_{r_1}^\beta a_{t_1} n_{x_1} \text{ et } F_2 = F_{r_2}^\beta a_{t_2} n_{x_2} \text{ tels que} \\ &F_1 \cap F_2 \neq \emptyset \text{ et } \lambda(F_1) \leq \lambda(F_2). \text{ Alors } F_1 \subset (F_2)^\sim. \end{aligned}$$

La condition $F_1 \cap F_2 \neq \emptyset$ équivaut à

$$(4.4.6) \quad \begin{cases} |t_1 - t_2| \leq r_1 + r_2 \\ \|x_1 - x_2\| \leq e^{\beta r_1 - t_1} + e^{\beta r_2 - t_2} \end{cases}$$

et la condition $\lambda(F_1) \leq \lambda(F_2)$ à

$$(4.4.7) \quad r_1 e^{d(\beta r_1 - t_1)} \leq r_2 e^{d(\beta r_2 - t_2)}.$$

Les inégalités (4.4.6) permettent d'estimer, pour tout $h = a_s n_y \in F_1$,

$$\begin{cases} |s - t_2| \leq |s - t_1| + |t_1 - t_2| \leq 2r_1 + r_2, \\ \|y - x_2\| \leq \|y - x_1\| + \|x_1 - x_2\| \leq 2e^{\beta r_1 - t_1} + e^{\beta r_2 - t_2}. \end{cases}$$

Utilisons (4.4.6) et (4.4.7) pour montrer que $h \in (F_2)^\sim$. On distingue à cet effet trois cas.

- Si $r_2 \geq \frac{\beta+1}{\beta-1} r_1$, on a d'une part $r_1 \leq r_2$, d'où

$$|s - t_2| \leq 3r_2 \leq \gamma r_2,$$

et d'autre part $\beta r_1 - t_1 \leq \beta r_1 + t_2 - t_1 - t_2 \leq (\beta+1)r_1 + r_2 - t_2 \leq \beta r_2 - t_2$, d'où

$$\|y - x_2\| \leq 3e^{\beta r_2 - t_2} \leq \gamma e^{\beta r_2 - t_2}.$$

- Si $r_1 \leq r_2 \leq \frac{\beta+1}{\beta-1} r_1$, on a de même

$$|s - t_2| \leq \gamma r_2.$$

D'après (4.4.7), on a d'autre part $e^{\beta r_1 - t_1} \leq \left(\frac{\beta+1}{\beta-1}\right)^{1/d} e^{\beta r_2 - t_2}$, d'où

$$\|y - x_2\| \leq 3 \left(\frac{\beta+1}{\beta-1}\right)^{1/d} e^{\beta r_2 - t_2} \leq \gamma e^{\beta r_2 - t_2}.$$

- Enfin, si $r_1 \geq r_2$, on déduit de (4.4.7) que $\beta r_1 - t_1 \leq \beta r_2 - t_2$, d'où

$$\|y - x_2\| \leq 3e^{\beta r_2 - t_2} \leq \gamma e^{\beta r_2 - t_2}.$$

De plus, on obtient $r_1 \leq \frac{\beta+1}{\beta-1} r_2$, en combinant $\beta r_1 - t_1 \leq \beta r_2 - t_2$ avec $|t_1 - t_2| \leq r_1 + r_2$. D'où

$$|s - t_2| \leq \frac{3\beta+1}{\beta-1} r_2 \leq \gamma r_2.$$

La fin de la démonstration du théorème 4.4.10 dans le cas $\beta > 1$ est semblable à celle du cas euclidien traité au chapitre 3. Plus précisément, on reprend la démonstration du théorème 3.1.2 après avoir déduit de la propriété (4.4.5) un lemme de type Vitali pour les translatés à droite $F_r^\beta g$. \square

Remarque 4.4.11. — Dans le cas $\beta > 1$, les translatés à droite $F_r^\beta g$ ont une propriété de recouvrement analogue à celle des boules dans l'espace euclidien, alors que la croissance du groupe S_d est exponentielle. Ce fait remarquable, observé initialement par Giuliani et Sjögren [GS90], a été exploité par Hebisch et Steger [HS03] pour élaborer une décomposition de Calderón–Zygmund et étudier des intégrales singulières sur le groupe S_d (ainsi que sur les arbres homogènes). Cette théorie a été approfondie et généralisée aux espaces de Damek–Ricci par Vallarino ([Val06], [Val07]). Notons que ces auteurs considèrent des opérateurs maximaux non centrés associés à des familles d'ensembles comparables à grande échelle aux translatés à droite $F_r^\beta g$, avec β variable entre deux constantes $1 < \beta_0 < \beta_1 < \infty$.

4.5. Un théorème ergodique sur le groupe S_d

Pour conclure ce chapitre, nous énonçons un théorème ergodique pour l'une des suites d'ensembles étudiées ci-dessus. Soit (X, \mathcal{B}, m) un espace mesuré sur lequel S_d agit à gauche en préservant la mesure σ -finie m . Si \mathcal{E} est une famille de parties mesurables de S_d , grâce au principe de transfert (voir section 2.3), toute inégalité maximale pour les opérateurs de moyenne A_E , $E \in \mathcal{E}$, associés à l'action à gauche de S_d sur lui-même, se transfère aux opérateurs de moyenne A_E associés à l'action de S_d sur (X, \mathcal{B}, m) .

Fixons $\beta \geq 1$ et notons A_r^β l'opérateur de moyenne défini sur $L^p(X, m)$ par

$$A_r^\beta f(x) = \frac{1}{\lambda(F_r^\beta)} \int_{F_r^\beta} f(gx) d\lambda(g).$$

D'après la remarque 4.2.2, F_r^β est une famille de Følner à gauche, non tempérée. Le théorème ergodique suivant pour une action de S_d se déduit alors des inégalités maximales obtenues dans les théorèmes 4.4.8 (i) et 4.4.10 et des résultats du chapitre 2 (voir le théorème 2.3.3 et les remarques 2.3.2 et 2.3.4).

Théorème 4.5.1. — Soient $p \in [1, +\infty[$, $\beta \geq 1$ et $f \in L^p(X, m)$.

- (i) Si $p > 1$, la famille à un paramètre $A_r^\beta f$ converge dans L^p et presque partout lorsque $r \rightarrow +\infty$.

- (ii) Si $p = 1$ et $\beta > 1$, la famille à un paramètre $A_r^\beta f$ converge presque partout lorsque $r \rightarrow +\infty$. Si m est une mesure de probabilité, la convergence a lieu aussi dans L^1 et la limite est l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I})$.

Notice bibliographique. — Il est naturel et intéressant de considérer les espaces symétriques G/K de rang quelconque et les groupes NA associés (voire des groupes résolubles plus généraux), où G désigne un groupe de Lie semi-simple connexe, non compact, de centre fini, et $G = NAK$ une décomposition d'Iwasawa. Le théorème 4.3.1 a été démontré dans ce cadre général par Clerc et Stein [CS74] pour $1 < p \leq \infty$ et par Strömberg [Str81] pour $p = 1$. Il reste encore beaucoup de travail à faire en rang supérieur, où des résultats ont été obtenus d'une part pour des opérateurs maximaux ([Ank92], [AJ99], [CGGM91], [GGHM88], [Ion00a], [Ion00b], [Ion05]) et d'autre part pour des intégrales singulières ([Ank90], [CGHM94], [CGM93], [CGM95], [CGM01], [CGM02], [Ion02], [Ion03], [Heb93], [Heb04], [MV08]). Voir aussi l'article de synthèse [Cow08].

CHAPITRE 5

MOYENNES SPHÉRIQUES

Nous décrivons dans ce chapitre les inégalités maximales et les théorèmes ergodiques relatifs aux moyennes sphériques sur \mathbb{R}^d ou \mathbb{Z}^d . Il s'agit donc d'effectuer des moyennes sur des ensembles de mesure de Lebesgue nulle dans \mathbb{R}^d . Dans cette situation, les lemmes de recouvrement de type Vitali sont inopérants pour obtenir des inégalités maximales. De nouvelles techniques d'analyse doivent être mises en place. Nous ferons en particulier appel à la théorie de Littlewood-Paley, qui se révèlera un excellent outil dans ce chapitre et les suivants.

Nous considérons dans la première section le cas classique des moyennes sur les sphères de \mathbb{R}^d . L'opérateur maximal correspondant est borné sur L^p pour $p > \frac{d}{d-1}$ et $d \geq 2$. Ce résultat est dû à Stein [**Ste70a**] lorsque $d \geq 3$ et à Bourgain [**Bou86**] si $d = 2$. Nous donnons ici une démonstration complète de l'inégalité maximale en dimension supérieure ou égale à 3, en nous inspirant de celle que donne Rubio de Francia dans [**Rub86**]. Le cas de la dimension 2 est significativement plus difficile et ne sera pas présenté dans ce livre.

Nous énonçons ensuite le théorème ergodique ponctuel pour ces moyennes sphériques. Celui-ci est dû à Jones [**Jon93**] en dimension supérieure ou égale à 3. La démonstration que nous en donnons est adaptée de l'argument de Rubio de Francia [**Rub86**]. Le cas de la dimension 2, qui est dû à Lacey [**Lac95**], est beaucoup plus difficile et nous ne le présenterons pas ici.

Le reste de ce chapitre est consacré au théorème maximal pour les moyennes sphériques de \mathbb{Z}^d obtenu par Magyar, Stein et Wainger [**MSW02**]. Ce théorème fait partie d'un ensemble de travaux consacrés aux analogues discrets de certains opérateurs classiques en analyse harmonique sur \mathbb{R}^d ([**SW99**], [**SW00**]). Si l'étude sur $\ell^p(\mathbb{Z}^d)$ de l'opérateur maximal discret

$$M_d(f)(n) = \sup_{r>0} \frac{c}{r^d} \sum_{m \in \mathbb{Z}^d, |m| < r} |f(n-m)|$$

associé aux moyennes sur les boules de \mathbb{Z}^d se ramène simplement à son analogue sur \mathbb{R}^d ([SW99, page 1292]), il n'en va pas de même pour l'opérateur maximal associé aux moyennes sur les sphères. Le passage du continu au discret doit ici être opéré avec soin et il apparaît des difficultés de nature arithmétique liées à la décomposition des entiers en sommes de carrés, qui nécessitent l'utilisation de méthodes issues de la théorie des nombres. Les théorèmes ergodiques associés aux moyennes sphériques pour l'action de \mathbb{Z}^d ont été obtenus par Magyar [Mag02].

Comme partout dans cet ouvrage, C (ou c) désignera une constante qui peut varier de ligne à ligne. On utilisera également les notations C_d et $C_{d,p}$ quand il y aura lieu de préciser que la constante ne dépend que de la dimension d ou que de la paire (dimension, exposant) (d, p) .

D'autre part, on notera $|x|$ la norme euclidienne de $x \in \mathbb{R}^d$.

5.1. Inégalités maximales pour les sphères de \mathbb{R}^d

Notations 5.1.1. — On notera \mathbb{S}^{d-1} la sphère unité de \mathbb{R}^d centrée en 0 et σ la mesure de probabilité uniforme portée par \mathbb{S}^{d-1} . Pour tout $r > 0$, nous désignerons par σ_r la mesure de probabilité uniforme sur la sphère $r\mathbb{S}^{d-1}$ de rayon r , ou encore l'image de σ par l'homothétie de centre 0 et de rayon r .

Pour φ dans l'espace de Schwartz $\mathcal{S}(\mathbb{R}^d)$ et pour $x \in \mathbb{R}^d$, nous posons

$$\sigma_r \cdot \varphi(x) = \varphi * \sigma_r(x) = \int_{r\mathbb{S}^{d-1}} \varphi(x+y) \, d\sigma_r(y) = \int_{\mathbb{S}^{d-1}} \varphi(x+ry) \, d\sigma(y).$$

Soit φ_σ^* l'opérateur maximal correspondant appliqué à φ , c'est-à-dire

$$\varphi_\sigma^*(x) = \sup_{r>0} |\sigma_r \cdot \varphi(x)|.$$

Nous nous intéresserons uniquement aux inégalités *a priori* concernant la fonction maximale φ_σ^* pour φ dans la classe de Schwartz. Notons que, si $f \in L^p(\mathbb{R}^d)$, les moyennes $\sigma_r \cdot f$ ne sont pas définies en général. Par contre, pour les valeurs de p pour lesquelles l'inégalité maximale *a priori* a lieu, on peut montrer que, en dehors d'un ensemble négligeable, la fonction $\sigma_r \cdot f$ et la fonction maximale f_σ^* sont bien définies pour tout r , et que l'inégalité maximale est encore satisfaite pour l'opérateur ainsi défini. Nous renvoyons le lecteur à l'appendice A, dans lequel ce passage aux fonctions de L^p est fait en détail.

5.1.1. Énoncé du théorème maximal sphérique. — Alors qu'en dimension 1 il n'y a pas d'inégalité maximale non triviale pour les sphères, qui sont réduites à deux points, on a en dimension supérieure le théorème suivant, appelé théorème maximal sphérique (ou circulaire si $d = 2$).

Théorème 5.1.2. — Soient $d \geq 2$ et $p > \frac{d}{d-1}$. Il existe une constante $C = C_{d,p} > 0$ telle que, pour toute fonction $\varphi \in \mathcal{S}(\mathbb{R}^d)$,

$$(5.1.1) \quad \|\varphi_\sigma^*\|_p \leq C\|\varphi\|_p.$$

Remarque 5.1.3. — Les exemples suivants montrent que les restrictions sur d et p sont optimales.

- Lorsque $d = 1$, alors $\sigma_r \cdot \varphi(x) = (\varphi(x+r) + \varphi(x-r))/2$, et seule est valide l'inégalité en norme uniforme. En effet, si $\varphi \in \mathcal{S}(\mathbb{R})$ est positive et $\varphi(0) = 1$, alors $\varphi_\sigma^* \geq \frac{1}{2}$ partout.
- Pour $d \geq 2$, on considère une fonction $\varphi \in \mathcal{S}(\mathbb{R}^d)$ positive et telle que $\varphi(x) \geq 1$ si $|x| \leq 1$. Soit alors $x \in \mathbb{R}^d$ avec $|x| \geq 1$. Écrivons $x = r\eta$ avec $\eta \in \mathbb{S}^{d-1}$ et $r \geq 1$. On a

$$\begin{aligned} \sigma_r \cdot \varphi(x) &= \int_{\mathbb{S}^{d-1}} \varphi(x - r\xi) \, d\sigma(\xi) \\ &\geq \int_{\{\xi \in \mathbb{S}^{d-1}; |\xi - \eta| \leq 1/r\}} d\sigma(\xi) \geq Cr^{1-d} \end{aligned}$$

où C est une constante ne dépendant que de la dimension. Il en résulte que la fonction φ_σ^* est minorée à l'infini par $\frac{C}{|x|^{\frac{d}{d-1}}}$ et n'est pas dans l'espace $L^p(\mathbb{R}^d)$ lorsque $p \leq \frac{d}{d-1}$.

Remarque 5.1.4. — Si on remplace la sphère par la frontière de l'hypercube, on voit facilement, comme en dimension 1, que l'inégalité maximale est fautive. Par contre ces théorèmes maximaux se généralisent à d'autres hypersurfaces à courbure positive (voir les références en fin de chapitre).

La démonstration du théorème 5.1.2 utilise deux méthodes classiques d'analyse réelle : dans un premier temps, nous aurons recours à une fonction de Littlewood-Paley auxiliaire et démontrerons le théorème dans le cas $p \geq 2$ et $d \geq 4$. Nous allierons ensuite cette démonstration à une décomposition de Littlewood-Paley pour établir le théorème dans le cas $d \geq 3$.

5.1.2. La démonstration dans un cas simple : $p \geq 2$ et $d \geq 4$. — Nous reproduisons ici la démonstration de Stein.

Sans avoir l'ambition de donner un historique de ces méthodes, rappelons que le terme de *fonction de Littlewood-Paley* a été d'abord utilisé pour la fonction $g(f)$ associée à une fonction f de $L^2(\mathbb{R}^d)$, définie par

$$g(f)(x) = \left(\int_0^\infty t \left| \frac{\partial}{\partial t} u(x, t) \right|^2 dt \right)^{1/2},$$

où u est le prolongement harmonique de f dans le demi-espace supérieur \mathbb{R}_+^{d+1} , donné par la convolution en la variable x avec le noyau de Poisson, $u(x, t) = P_t * f(x)$. Le

noyau de Poisson, qu'on peut voir comme la fonction harmonique dans le demi-espace supérieur dont la valeur au bord est la masse de Dirac en 0, est donné par

$$(5.1.2) \quad P_t(x) = \frac{c_d t}{(|x|^2 + t^2)^{\frac{d+1}{2}}},$$

la constante c_d étant choisie de sorte que l'intégrale en x soit égale à 1. Le lecteur intéressé peut consulter le livre de Stein [Ste70a, pages 61 et 82], ainsi que le livre de Zygmund [Zyg02].

Pour généraliser cette définition, on remarque que $t \frac{\partial}{\partial t} u(x, t)$ peut encore s'écrire $\psi_t * f(x)$, où ψ_t se déduit d'une même fonction ψ par homogénéité : $\psi_t(x) = t^{-d} \psi(x/t)$. Par extension, on appelle *fonction de Littlewood-Paley* associée à f , et on note encore quelquefois $g(f)$ quand il n'y a pas d'ambiguïté, la fonction

$$g_\psi(f)(x) = \left(\int_0^\infty |\psi_t * f(x)|^2 \frac{dt}{t} \right)^{1/2},$$

où ψ est une fonction de $L^1(\mathbb{R}^d)$ telle qu'il existe une constante C_ψ pour laquelle, pour tout $\xi \in \mathbb{R}^d$ non nul,

$$(5.1.3) \quad \left(\int_0^\infty |\widehat{\psi}(t\xi)|^2 \frac{dt}{t} \right)^{1/2} \leq C_\psi.$$

Cette condition intervient de façon très naturelle du fait de l'identité suivante, dont la démonstration est une conséquence immédiate de l'identité de Plancherel :

$$(5.1.4) \quad \|g_\psi(f)\|_2^2 = \int |\widehat{f}(\xi)|^2 \left(\int_0^\infty |\widehat{\psi}(t\xi)|^2 \frac{dt}{t} \right) d\xi.$$

Ainsi, sous la condition (5.1.3), on a, quel que soit $f \in L^2(\mathbb{R}^d)$,

$$(5.1.5) \quad \|g_\psi(f)\|_2 \leq C_\psi \|f\|_2.$$

Remarquons que dans le cas de la fonction de Littlewood-Paley classique, où $\widehat{\psi}(\xi) = -2\pi|\xi|e^{-2\pi|\xi|}$, le membre de gauche de (5.1.3) ne dépend pas de $\xi \neq 0$.

Ainsi, lorsque la condition (5.1.3) est réalisée, la norme de $g_\psi(f)$ dans $L^2(\mathbb{R}^d)$ est contrôlée par celle de f . Sous des hypothèses de régularité sur ψ , on peut montrer qu'il y a également équivalence des normes dans L^p , ces espaces devant être remplacés par les espaces de Hardy si $p \leq 1$ (voir [Ste93]). Enfin, remarquons que l'inégalité (5.1.5) est encore valable lorsque ψ est une distribution tempérée dont la transformée de Fourier satisfait à la condition (5.1.3), à condition qu'on suppose *a priori* la fonction f dans la classe de Schwartz. C'est dans ces conditions que nous allons d'abord l'utiliser.

Commençons par donner une démonstration du théorème 5.1.2 dans le cas simple où $d \geq 4$ et $p = 2$. Par interpolation (théorème 1.2.2) avec le cas trivial $p = +\infty$, on obtiendra le théorème pour $2 \leq p \leq +\infty$.

L'idée de Stein est de relier les moyennes sur les sphères aux moyennes sur les boules. Le passage de l'un à l'autre se fait à l'aide de la fonctionnelle de Littlewood-Paley généralisée que nous venons d'introduire : pour $\varphi \in \mathcal{S}(\mathbb{R}^d)$, on note

$$(5.1.6) \quad g(\varphi)(x) = \left(\int_0^\infty s \left| \frac{d}{ds} (\sigma_s \cdot \varphi)(x) \right|^2 ds \right)^{1/2}.$$

Le lien obtenu s'énonce alors de la manière suivante.

Lemme 5.1.5. — *Pour toute fonction $\varphi \in \mathcal{S}(\mathbb{R}^d)$, pour tout $r > 0$ et pour $x \in \mathbb{R}^d$,*

$$|\sigma_r \cdot \varphi(x)| \leq |\beta_r \cdot \varphi(x)| + \frac{1}{\sqrt{2d}} g(\varphi)(x),$$

où β_r désigne la moyenne sur la boule de rayon r .

Démonstration. — Pour $r > 0$ et $x \in \mathbb{R}^d$, on a

$$\begin{aligned} \sigma_r \cdot \varphi(x) &= r^{-d} \int_0^r \frac{d}{ds} (s^d \sigma_s \cdot \varphi(x)) ds \\ &= dr^{-d} \int_0^r s^{d-1} \sigma_s \cdot \varphi(x) ds + r^{-d} \int_0^r s^d \frac{d}{ds} (\sigma_s \cdot \varphi)(x) ds \\ &= \beta_r \cdot \varphi(x) + r^{-d} \int_0^r s^d \frac{d}{ds} (\sigma_s \cdot \varphi)(x) ds. \end{aligned}$$

On a donc

$$\begin{aligned} |\sigma_r \cdot \varphi(x)| &\leq |\beta_r \cdot \varphi(x)| + r^{-d} \left(\int_0^r s^{2d-1} ds \right)^{1/2} \times \left(\int_0^\infty s \left| \frac{d}{ds} (\sigma_s \cdot \varphi)(x) \right|^2 ds \right)^{1/2} \\ &\leq |\beta_r \cdot \varphi(x)| + \frac{1}{\sqrt{2d}} g(\varphi)(x). \end{aligned}$$

□

L'inégalité maximale L^2 pour les moyennes sur les sphères découle alors de l'inégalité maximale L^2 pour les moyennes sur les boules et du lemme suivant, qui donne l'analogie de l'inégalité (5.1.5) pour la fonction de Littlewood-Paley associée à σ .

Lemme 5.1.6. — *Pour $d \geq 4$, il existe une constante $C = C_d > 0$ telle que, pour toute fonction $\varphi \in \mathcal{S}(\mathbb{R}^d)$,*

$$\|g(\varphi)\|_2 \leq C \|\varphi\|_2.$$

La démonstration de ce lemme utilise les relations entre la transformée de Fourier de la mesure σ et les fonctions de Bessel, données par le lemme suivant :

Lemme 5.1.7. — *Notons J_ν la fonction de Bessel d'ordre ν . La transformée de Fourier de σ est donnée par :*

$$(5.1.7) \quad \forall \rho \geq 0, \forall \xi \in \mathbb{S}^{d-1}, \quad \widehat{\sigma}(\rho\xi) = 2\pi\rho^{1-d/2} J_{(d/2)-1}(2\pi\rho).$$

En particulier, on a les estimations suivantes, lorsque ρ tend vers l'infini :

$$(5.1.8) \quad \widehat{\sigma}(\rho\xi) = O(\rho^{-\frac{d-1}{2}})$$

et

$$(5.1.9) \quad \rho \frac{\partial}{\partial \rho} \widehat{\sigma}(\rho\xi) = O(\rho^{-\frac{d+1}{2}}).$$

Démonstration du lemme 5.1.7. — La formule (5.1.7) est classique (voir par exemple [SW71], [Ste93, page 347]). Les propriétés des fonctions de Bessel que nous utilisons ici se trouvent dans de nombreux ouvrages. Le lecteur pourra par exemple consulter [Leb72, Section 5.3] ou Stein-Weiss [SW71]. L'estimation (5.1.8) résulte alors de $J_\nu(\rho) = O\left(\frac{1}{\sqrt{\rho}}\right)$. Enfin, en utilisant la formule classique

$$\frac{d}{d\rho} \rho^{-\nu} J_\nu(\rho) = -\rho^{-\nu} J_{\nu+1}(\rho),$$

on obtient (5.1.9). □

Démonstration du lemme 5.1.6. — On a

$$\sigma_r \cdot \varphi(x) = \varphi * \sigma_r(x).$$

Il en résulte que

$$\widehat{\sigma_r \cdot \varphi}(\xi) = \widehat{\varphi}(\xi) \times \widehat{\sigma}(r\xi).$$

Comme $\varphi \in \mathcal{S}(\mathbb{R}^d)$, la transformée de Fourier de $\frac{d}{dr}(\sigma_r \cdot \varphi)$ au point ξ est égale à

$$\widehat{\varphi}(\xi) \times \frac{d}{dr} \widehat{\sigma}(r\xi).$$

Par le théorème de Plancherel, on a donc, comme dans la démonstration de (5.1.5),

$$\int_{\mathbb{R}^d} \left| \frac{d}{dr}(\sigma_r \cdot \varphi)(x) \right|^2 dx = \int_{\mathbb{R}^d} |\widehat{\varphi}(\xi)|^2 \left| \frac{d}{dr} \widehat{\sigma}(r\xi) \right|^2 d\xi.$$

En multipliant par r et en intégrant par rapport à r , on obtient

$$\int_{\mathbb{R}^d} |g_\sigma(\varphi)(x)|^2 dx = \int_{\mathbb{R}^d} |\widehat{\varphi}(\xi)|^2 \int_0^\infty r \left| \frac{d}{dr} \widehat{\sigma}(r\xi) \right|^2 dr d\xi$$

à l'aide du théorème de Fubini. Il suffit, pour conclure, de montrer que

$$\int_0^\infty r \left| \frac{d}{dr} \widehat{\sigma}(r\xi) \right|^2 dr$$

est fini, ce qui résulte immédiatement de (5.1.9) puisque $d \geq 4$. □

5.1.3. La démonstration pour $d \geq 3$. — Nous allons maintenant modifier la démonstration précédente pour obtenir le théorème 5.1.2 dans le cas $d \geq 3$ et $p > \frac{d}{d-1}$. Notons qu'il suffit d'établir ce théorème pour $\frac{d}{d-1} < p \leq 2$, les autres cas s'obtenant alors par interpolation avec le cas trivial $p = +\infty$. Pour cela, nous suivons l'idée de Rubio de Francia qui utilise, en plus des techniques développées dans la section 5.1.2, une décomposition de Littlewood-Paley.

5.1.3.1. Décomposition de Littlewood-Paley. — Nous voulons étudier l'opérateur

$$\varphi_\sigma^* = \sup_{r>0} |\sigma_r \cdot \varphi|.$$

Nous allons écrire σ_r comme une somme de termes, obtenus en décomposant la fonction $\hat{\sigma}$ selon des couronnes dyadiques. Pour cela, fixons une fonction $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ qui est la transformée de Fourier d'une fonction radiale, de classe \mathcal{C}^∞ , à support compact. Supposons de plus que $\psi_0(0) = 1$ et que, pour $1 \leq i < \frac{d}{2}$,

$$\left(\frac{\partial^i}{\partial r^i} \psi_0 \right) (0) = 0$$

où $\frac{\partial}{\partial r}$ désigne la dérivation dans la direction radiale. ⁽¹⁾

On peut obtenir une telle fonction en partant d'une fonction ψ qui est la transformée de Fourier d'une fonction \mathcal{C}^∞ à support compact, radiale, et telle que $\psi(0) \neq 0$, puis en posant, pour $\xi \in \mathbb{S}^{d-1}$,

$$\psi_0(r\xi) = \left(\sum_{j=0}^{[d/2]} a_j r^j \right) \psi(r\xi)$$

où les a_j sont choisis comme solutions d'un système linéaire triangulaire, de sorte que $\psi_0(0) = 1$ et que ψ_0 ait en 0 le nombre voulu de dérivées nulles ($[d/2]$ étant la partie entière de $d/2$).

Posons maintenant $\psi_1(\xi) = \psi_0(\xi/2) - \psi_0(\xi)$ et, pour $j \geq 1$, $\psi_j(\xi) = \psi_1(2^{-(j-1)}\xi)$. Ainsi, ψ_j est encore radiale et c'est la transformée de Fourier d'une fonction \mathcal{C}^∞ à support compact, notée $\check{\psi}_j$. De plus, il existe $\eta, c > 0$ tels que,

$$(5.1.10) \quad \text{pour } |\xi| < \eta, \quad |\psi_1(\xi)| \leq c|\xi|^{d/2}.$$

Enfin, on a, pour tout $\xi \in \mathbb{R}^d$,

$$(5.1.11) \quad \sum_{j=0}^{+\infty} \psi_j(\xi) = 1.$$

⁽¹⁾Pour démontrer le théorème 5.1.2 dans le cas $d \geq 3$, on pourrait prendre $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ telle que $0 \leq \psi_0 \leq 1$, $\psi_0(\xi) = 1$ si $|\xi| \leq 1$ et $\psi_0(\xi) = 0$ si $|\xi| \geq 2$, ce qui simplifierait un peu la démonstration. Mais ce choix interdirait que ψ_0 soit la transformée de Fourier d'une fonction à support compact. L'intérêt du choix que nous effectuons apparaîtra dans la section 5.1.4.

Posons $m_j = \widehat{\sigma}\psi_j$ et notons σ_j la fonction dans $\mathcal{S}(\mathbb{R}^d)$ telle que $\widehat{\sigma}_j = m_j$. Il résulte de (5.1.11) que

$$\sigma * \varphi = \sum_{j=0}^{\infty} \sigma_j * \varphi.$$

Soit $\sigma_{j,r}$ la fonction déduite de σ_j par homogénéité : $\sigma_{j,r}(x) = r^{-d}\sigma_j(x/r)$. On a, de la même manière que précédemment,

$$\sigma_r \cdot \varphi = \sum_{j=0}^{\infty} \sigma_{j,r} \cdot \varphi.$$

Pour $\varphi \in \mathcal{S}(\mathbb{R}^d)$, rappelons que $\sigma_{j,r} \cdot \varphi = \sigma_{j,r} * \varphi$. On pose $\varphi_{\sigma_j}^*(x) = \sup_{r>0} |\sigma_{j,r} \cdot \varphi(x)|$. On a alors

$$(5.1.12) \quad \varphi_{\sigma}^* \leq \sum_{j=0}^{+\infty} \varphi_{\sigma_j}^*.$$

Pour démontrer le théorème 5.1.2, il suffit donc d'établir une inégalité maximale de la forme

$$(5.1.13) \quad \left\| \varphi_{\sigma_j}^* \right\|_p \leq C_{j,p} \|\varphi\|_p$$

avec $\sum_j C_{j,p} < +\infty$.

Ceci sera obtenu en trois étapes, détaillées dans les sous-sections à venir, de 5.1.3.2 à 5.1.3.4. D'une part, le fait que $\psi_0 \in \mathcal{S}(\mathbb{R}^d)$ permet de comparer $\varphi_{\sigma_j}^*$ à un multiple de la fonction maximale de Hardy-Littlewood φ_{β}^* . Ceci nous permet d'établir, pour $j \geq 0$, une inégalité de type faible (1, 1) :

$$(5.1.14) \quad \left| \{ \varphi_{\sigma_j}^* > \alpha \} \right| \leq \frac{C_{j,1}}{\alpha} \|\varphi\|_1.$$

avec des constantes $C_{j,1}$ qui croissent suffisamment lentement avec j .

Ensuite, en utilisant les propriétés de décroissance de $\widehat{\sigma}$ et la localisation des ψ_j , on obtient, pour $j \geq 1$,

$$(5.1.15) \quad \left\| \varphi_{\sigma_j}^* \right\|_2 \leq C_{j,2} \|\varphi\|_2.$$

avec des constantes $C_{j,2}$ qui décroissent vite quand $j \rightarrow +\infty$.

Après interpolation, entre ces deux inégalités pour $j \geq 1$ et entre la première inégalité et l'inégalité de type fort (∞, ∞) pour $j = 0$, on obtient (5.1.13) avec des constantes $C_{j,p}$ satisfaisant à $\sum_j C_{j,p} < +\infty$ lorsque $p > \frac{d}{d-1}$ et $d \geq 3$.

5.1.3.2. Estimation de $\varphi_{\sigma_j}^$.* — Nous allons montrer que $\varphi_{\sigma_j}^*$ est majorée par la fonction maximale de Hardy-Littlewood. Plus précisément :

Lemme 5.1.8. — *Il existe une constante $C = C_d$ telle que, pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d)$, et tout $j \geq 0$,*

$$\varphi_{\sigma_j}^* \leq C 2^j |\varphi|_{\beta}^*.$$

Démonstration. — Par invariance par translation puis par dilatation, il suffit de montrer l'inégalité recherchée en 0, puis de montrer que, si φ est une fonction positive,

$$(5.1.16) \quad \int |\sigma_j(x)|\varphi(x) dx \leq C2^j \varphi_\beta^*(0).$$

Commençons par montrer le lemme suivant :

Lemme 5.1.9. — *Il existe une constante $C = C_d > 0$ telle que, pour tout $j \geq 0$ et tout $x \in \mathbb{R}^d$,*

$$|\sigma_j(x)| \leq C \frac{2^j}{(1+|x|)^{d+1}}.$$

Démonstration. — Rappelons que $\sigma_j = \sigma * \check{\psi}_j$ où, pour $j \geq 1$,

$$\check{\psi}_j(x) = 2^{(j-1)d} \check{\psi}_1(2^{j-1}x).$$

Comme $\check{\psi}_0$ et $\check{\psi}_1$ sont des fonctions C^∞ à support compact, on peut majorer $|\check{\psi}_0(x)|$ et $|\check{\psi}_1(x)|$ par $\frac{C}{(1+|x|)^{d+1}}$. Le lemme découle alors de l'estimation suivante :

$$(5.1.17) \quad \int_{\mathbb{S}^{d-1}} \frac{2^{jd}}{(1+2^j|x-\xi|)^{d+1}} d\sigma(\xi) \leq C \frac{2^j}{(1+|x|)^{d+1}}.$$

Pour $|x| > 2$, on a $|x-\xi| \geq \frac{|x|}{2}$, ce dont on déduit aisément une majoration par $C2^{-j}|x|^{-d-1}$, plus forte que (5.1.17).

Supposons donc $|x| \leq 2$ et montrons que le membre de gauche de (5.1.17) est majoré par $C2^j$. Pour cela, on découpe l'intégrale suivant des couronnes dyadiques et on la majore par

$$2^{jd} \int_{|\xi-x| \leq 2^{-j}} d\sigma(\xi) + \sum_{k=0}^{+\infty} 2^{jd} 2^{-(d+1)k} \int_{|\xi-x| \leq 2^{k-j+1}} d\sigma(\xi).$$

On conclut la démonstration de (5.1.17), et donc du lemme 5.1.9, en remarquant que $\sigma(\{\xi; |\xi-x| < r\}) \leq C_d r^{d-1}$. □

Revenons à la démonstration de (5.1.16). D'après le lemme 5.1.9, il suffit d'estimer l'intégrale

$$\int_{\mathbb{R}^d} \frac{\varphi(x)}{(1+|x|^2)^{\frac{d+1}{2}}} dx.$$

À nouveau, on découpe cette intégrale suivant des couronnes dyadiques pour la majorer par

$$\int_{|x| \leq 1} \varphi(x) dx + \sum_{k=0}^{+\infty} 2^{-k(d+1)} \int_{2^k \leq |x| \leq 2^{k+1}} \varphi(x) dx.$$

Mais $\int_{|x| \leq 1} \varphi(x) dx$ est majorée par $|B(0, 1)|\varphi_\beta^*(0)$ tandis que chacune des intégrales suivantes est majorée par

$$|B(0, 2^{k+1})|\varphi_\beta^*(0) = 2^{(k+1)d}|B(0, 1)|\varphi_\beta^*(0).$$

La majoration cherchée en découle aisément. \square

En combinant le lemme 5.1.8 avec l'estimation de la fonction maximale de Hardy-Littlewood donnée dans le théorème 3.1.2, nous obtenons le lemme suivant :

Lemme 5.1.10. — *Il existe une constante $C = C_d$ telle que, pour tout $\varphi \in \mathcal{S}(\mathbb{R}^d)$, tout $j \geq 0$ et tout $\alpha > 0$, on a*

$$(5.1.14) \quad \left| \{\varphi_{\sigma_j}^* > \alpha\} \right| \leq \frac{C2^j}{\alpha} \|\varphi\|_1.$$

5.1.3.3. *Estimation de $\|\varphi_{\sigma_j}^*\|_{L^2(\mathbb{R}^d)}$.* — Le but de cette sous-section est de démontrer le lemme suivant :

Lemme 5.1.11. — *Il existe une constante $C = C_d$ telle que, pour tout entier $j \geq 1$ et toute fonction $\varphi \in \mathcal{S}(\mathbb{R}^d)$, on a*

$$(5.1.15) \quad \left\| \varphi_{\sigma_j}^* \right\|_2 \leq C2^{-(d-2)j/2} \|\varphi\|_2.$$

Comme dans le cas simple étudié dans la section 5.1.2, la démonstration repose sur une majoration de $\varphi_{\sigma_j}^*$ en termes de fonctions de Littlewood-Paley adaptées que nous commençons par introduire.

Notations. — Pour $\varphi \in \mathcal{S}(\mathbb{R}^d)$, nous considérons $\sigma_{j,r} \cdot \varphi = \sigma_{j,r} * \varphi$, de sorte que $\widehat{\sigma_{j,r} \cdot \varphi}(\xi) = \widehat{\varphi}(\xi) \times m_j(r\xi)$. Notons

$$g_j(\varphi)(x) = \left(\int_0^{+\infty} |\sigma_{j,r} \cdot \varphi(x)|^2 \frac{dr}{r} \right)^{1/2}$$

la fonction de Littlewood-Paley associée à la fonction σ_j .

Posons $\tilde{\sigma}_{j,r}(x) = r \frac{d}{dr} \sigma_{j,r}(x)$ et $\tilde{\sigma}_{j,r} \cdot \varphi = \tilde{\sigma}_{j,r} * \varphi$. Introduisons enfin la fonction g associée :

$$\tilde{g}_j(\varphi)(x) = \left(\int_0^{+\infty} |\tilde{\sigma}_{j,r} \cdot \varphi(x)|^2 \frac{dr}{r} \right)^{1/2} = \left(\int_0^{+\infty} r \left| \frac{d}{dr} \sigma_{j,r} \cdot \varphi(x) \right|^2 dr \right)^{1/2}.$$

Nous avons vu en (5.1.4) que

$$\|g_j(\varphi)\|_2^2 \leq \|\varphi\|_2^2 \sup_{\xi \neq 0} \int_0^{+\infty} |m_j(r\xi)|^2 \frac{dr}{r}.$$

Comme la fonction m_j est radiale, l'intégrale considérée est indépendante de ξ . Il suffit donc de la calculer en un seul point de la sphère unité, noté encore ξ . En reprenant la définition de m_j , on voit que

$$\int_0^{+\infty} |m_j(r\xi)|^2 \frac{dr}{r} = \int_0^{+\infty} |\widehat{\sigma}(2^j r\xi)|^2 |\psi_1(r\xi)|^2 \frac{dr}{r}.$$

L'estimation (5.1.8) permet alors d'obtenir la majoration

$$\int_0^{+\infty} |m_j(r\xi)|^2 \frac{dr}{r} \leq C 2^{-j(d-1)} \int_0^{+\infty} \frac{|\psi_1(r\xi)|^2}{s^d} ds.$$

Cette dernière intégrale est finie grâce à (5.1.10) et à la décroissance en l'infini de ψ_1 . Il en résulte que

$$(5.1.18) \quad \|g_j(\varphi)\|_2 \leq C_d 2^{-(d-1)j/2} \|\varphi\|_2.$$

Le même raisonnement, en remplaçant $\xi \mapsto m_j(r\xi)$ par $\xi \mapsto r \frac{d}{dr} m_j(r\xi)$ et l'estimation (5.1.8) par (5.1.9), nous donne

$$(5.1.19) \quad \|\tilde{g}_j(\varphi)\|_2 \leq c_d 2^{j(1-\frac{d-1}{2})} \|\varphi\|_2.$$

A l'aide de ces estimations nous pouvons maintenant démontrer le lemme 5.1.11.

Démonstration du lemme 5.1.11. — Comme $\lim_{r \rightarrow +\infty} \sigma_{j,r} \cdot \varphi = 0$, on a

$$\begin{aligned} |\sigma_{j,r} \cdot \varphi(x)|^2 &= -2\Re \int_r^{+\infty} \frac{\sigma_{j,s} \cdot \varphi(x)}{\sigma_{j,s} \cdot \varphi(x)} \frac{d}{ds} \sigma_{j,s} \cdot \varphi(x) ds \\ &= -2\Re \int_r^{+\infty} \frac{\sigma_{j,s} \cdot \varphi(x)}{\sigma_{j,s} \cdot \varphi(x)} \tilde{\sigma}_{j,s} \cdot \varphi(x) \frac{ds}{s} \\ &\leq 2 \int_0^{+\infty} |\sigma_{j,s} \cdot \varphi(x)| |\tilde{\sigma}_{j,s} \cdot \varphi(x)| \frac{ds}{s}. \end{aligned}$$

Avec l'inégalité de Cauchy-Schwarz, on en déduit que

$$\sup_{r>0} |\sigma_{j,r} \cdot \varphi(x)|^2 \leq 2g_j(\varphi)(x) \tilde{g}_j(\varphi)(x).$$

En intégrant sur \mathbb{R}^d et en utilisant encore l'inégalité de Cauchy-Schwarz, on obtient

$$\left\| \varphi_{\sigma_j}^* \right\|_2^2 \leq 2 \|g_j(\varphi)\|_2 \|\tilde{g}_j(\varphi)\|_2 \leq C_d 2^{-j(d-2)} \|\varphi\|_2^2,$$

avec (5.1.18) et (5.1.19). \square

5.1.3.4. Conclusion. — Nous avons établi une inégalité de type faible (1, 1) pour $\varphi_{\sigma_j}^*$ dans le lemme 5.1.10. Par ailleurs, pour $j \geq 1$ une inégalité de type fort (2, 2) a été obtenue dans le lemme 5.1.11.

Par interpolation entre ces deux inégalités on obtient, pour chaque $1 < p \leq 2$, une constante $C = C_{p,d}$ ne dépendant que de p et de la dimension d , telle que, pour tout

$\varphi \in \mathcal{S}(\mathbb{R}^d)$, et tout $j \geq 1$,

$$(5.1.20) \quad \left\| \varphi_{\sigma_j}^* \right\|_p \leq C 2^{\left(\frac{d}{p} - d + 1\right)j} \|\varphi\|_p.$$

L'interpolation entre l'inégalité de type faible $(1, 1)$ et l'inégalité de type fort (∞, ∞) montre que (5.1.20) est encore vraie pour $j = 0$.

Mais alors, si $\frac{d}{p} - d + 1 < 0$, *i.e.* si $p > \frac{d}{d-1}$, la série

$$(5.1.21) \quad \sum_{j \geq 0} 2^{\left(\frac{d}{p} - d + 1\right)j}$$

converge. Ainsi, en combinant (5.1.12) et (5.1.20), on obtient

$$\|\varphi_{\sigma}^*\|_p \leq C \|\varphi\|_p.$$

ce qui conclut la démonstration du théorème 5.1.2 dans le cas $d \geq 3$.

Remarque 5.1.12. — Pour pouvoir appliquer le théorème d'interpolation, il nous faut $1 < p \leq 2$. Comme il nous faut également $p > \frac{d}{d-1}$ pour avoir convergence de la somme (5.1.21), on voit que cette démonstration n'est valide que si $d \geq 3$.

La principale difficulté pour démontrer ce théorème en dimension $d = 2$ provient du fait qu'il n'est pas valide pour $p = 2$, ce qui ne permet pas l'utilisation directe de techniques L^2 .

5.1.4. Le théorème ergodique pour les sphères. — Dans cette section, nous supposons que \mathbb{R}^d agit sur un espace (X, \mathcal{B}, m) en préservant la mesure de probabilité m . Le théorème ergodique ponctuel, dû à Jones si $d \geq 3$ et à Lacey si $d = 2$ s'énonce comme suit :

Théorème 5.1.13. — Soit $d \geq 2$ et $p > \frac{d}{d-1}$. Pour tout $f \in L^p(X, m)$, $\sigma_r \cdot f$ tend presque partout, lorsque r tend vers l'infini, vers $\mathbb{E}(f|\mathcal{I})$, l'espérance conditionnelle de f relativement à la tribu des invariants \mathcal{I} .

Pour donner un sens à l'énoncé précédent, il faut déjà savoir que si $d \geq 2$, $p > \frac{d}{d-1}$ et $f \in L^p(X, m)$, alors la famille $(\sigma_r \cdot f)_{r>0}$ est m -presque sûrement bien définie. C'est garanti par l'inégalité maximale du théorème 5.1.2, d'après le théorème 2.3.1 et l'appendice A.

Remarque 5.1.14. — En ce qui concerne le théorème ergodique en moyenne, un argument de densité déjà utilisé dans la remarque 1.3.5 montre qu'il s'étend à tous les espaces L^p , $p \geq 1$, une fois établi dans L^2 . Un argument classique de théorie spectrale, et les seuls faits que $\widehat{\sigma_r}(0) = 1$ et $\lim_{r \rightarrow +\infty} \widehat{\sigma_r}(\xi) = 0$ pour tout $\xi \neq 0$, suffisent pour démontrer le théorème ergodique en moyenne dans L^2 , c'est-à-dire que pour tout $f \in L^2(X, m)$, on a

$$(5.1.22) \quad \lim_{r \rightarrow +\infty} \|\sigma_r \cdot f - \mathbb{E}(f|\mathcal{I})\|_2 = 0.$$

En effet soit m_f la mesure spectrale sur \mathbb{R}^d associée à f et à la représentation unitaire de \mathbb{R}^d dans $L^2(X, m)$ définie par l'action de \mathbb{R}^d sur (X, \mathcal{B}, m) . On a alors

$$\|\sigma_r \cdot f\|_2^2 = \int_{\mathbb{R}^d} |\widehat{\sigma_r}(\xi)|^2 dm_f(\xi).$$

Si f est invariante par l'action de \mathbb{R}^d , l'égalité (5.1.22) est bien sûr vraie. Si f est orthogonale à l'espace des fonctions invariantes, l'origine de \mathbb{R}^d est de masse nulle relativement à m_f (voir le lemme B.2.3 de l'appendice B). Le théorème de convergence dominée de Lebesgue permet de conclure que $\lim_{r \rightarrow +\infty} \|\sigma_r \cdot f\|_2 = 0$ dans ce cas.

Nous utiliserons à nouveau cette méthode spectrale dans les chapitres suivants, pour les actions du groupe libre \mathbb{F}_d (théorème 6.2.1) et du groupe $SO^o(d, 1)$ des isométries de l'espace hyperbolique réel (théorème 7.1.1).

Nous ne démontrerons le théorème 5.1.13 que dans le cas $d \geq 3$. Les moyennes n'étant pas effectuées sur des familles de Følner, ce théorème n'est pas une conséquence de l'inégalité maximale. Notre démonstration prolonge celle que nous avons donnée pour l'inégalité maximale. Elle peut s'adapter immédiatement aux mesures radiales dont la transformée de Fourier décroît suffisamment vite à l'infini. Elle est plus simple que celle de Jones, même si elle repose sur le même principe, à savoir l'utilisation du théorème ergodique de Wiener. Cela se fait comme dans l'article de Jones à l'aide du lemme suivant, qui utilise les liens entre moyennes sphériques et moyennes sur des boules.

Lemme 5.1.15. — Soit $k \in C^\infty(\mathbb{R}^d)$ une fonction radiale à support compact et, pour $r > 0$, soit K_r l'opérateur défini sur $L^p(X, m)$ par

$$K_r f(x) = \int_{\mathbb{R}^d} f((rg)x) k(g) dg = \int_{\mathbb{R}^d} f(gx) k_r(g) dg = k_r \cdot f(x),$$

où $k_r(g) = \frac{1}{r^d} k(g/r)$. Alors, pour tout $p \geq 1$ et pour tout $f \in L^p(X, m)$, on a

$$\lim_{r \rightarrow +\infty} K_r f = \left(\int_{\mathbb{R}^d} k(y) dy \right) \mathbb{E}(f|\mathcal{I})$$

presque partout.

Démonstration. — Rappelons que σ est la mesure de Lebesgue normalisée sur la sphère \mathbb{S}^{d-1} . En écrivant $k(g) = k_0(|g|)$ et en passant en coordonnées polaires, on a donc

$$K_r f(x) = d|B(0, 1)| \int_0^{+\infty} k_0(\rho) \rho^{d-1} \int_{\mathbb{S}^{d-1}} f((r\rho\zeta)x) d\sigma(\zeta) d\rho.$$

Comme k_0 est à support compact, une intégration par parties donne alors

$$\begin{aligned} K_r f(x) &= - \int_0^{+\infty} k_0'(\rho) \left(d|B(0, 1)| \int_0^\rho t^{d-1} \int_{\mathbb{S}^{d-1}} f((rt\zeta)x) d\sigma(\zeta) dt \right) d\rho \\ &= - \int_0^{+\infty} k_0'(\rho) r^{-d} \int_{B(0, r\rho)} f(gx) dg d\rho \end{aligned}$$

en repassant en coordonnées cartésiennes. Comme $|B(0, r\rho)| = (r\rho)^d |B(0, 1)|$, on en déduit que

$$K_r f(x) = -|B(0, 1)| \int_0^{+\infty} k'_0(\rho) \rho^d \beta_{r\rho} \cdot f(x) d\rho.$$

D'après le théorème de Wiener, $\beta_{r\rho} \cdot f$ tend vers $\mathbb{E}(f|\mathcal{I})$ presque partout quand r tend vers $+\infty$. Avec le théorème de convergence dominée, il en résulte que

$$\lim_{r \rightarrow +\infty} K_r f = -|B(0, 1)| \left(\int_0^{+\infty} k'_0(\rho) \rho^d d\rho \right) \mathbb{E}(f|\mathcal{I})$$

presque partout. Une nouvelle intégration par parties et un passage en coordonnées cartésiennes donnent alors

$$\lim_{r \rightarrow +\infty} K_r f = \left(\int_{\mathbb{R}^d} k(y) dy \right) \mathbb{E}(f|\mathcal{I})$$

presque partout, ce qui conclut la démonstration du lemme. \square

Démonstration du théorème 5.1.13. — Nous conservons les notations de la section 5.1.3. Le lemme 5.1.15, appliqué à $k = \sigma_j$, donne

$$\lim_{r \rightarrow +\infty} \sigma_{j,r} \cdot f = \mathbb{E}(f|\mathcal{I}) \text{ si } j = 0 \quad \text{et} \quad \lim_{r \rightarrow +\infty} \sigma_{j,r} \cdot f = 0 \text{ sinon,}$$

puisque $\int_{\mathbb{R}^d} \sigma_j(u) du = \widehat{\sigma}_j(0)$ vaut 1 si $j = 0$ et 0 sinon.

Rappelons enfin que σ_j est à support compact⁽²⁾. Par suite, pour tout $R > 0$, les opérateurs $\varphi_{\sigma_j, R}^* = \sup_{0 < r < R} |\sigma_{j,r} \cdot \varphi|$ sont semi-locaux et on peut appliquer le principe de transfert à l'inégalité (5.1.20) (voir la remarque 2.3.4). Il existe donc $C > 0$ tel que, pour tous $j \geq 0$, $R > 0$ et tout $f \in L^p(X)$, ($1 < p \leq 2$),

$$\left\| \sup_{0 < r < R} |\sigma_{j,r} \cdot f| \right\|_p \leq C 2^{\left(\frac{d}{p} - d + 1\right)j} \|f\|_p.$$

En interpolant cette inégalité avec l'inégalité (triviale) de type fort (∞, ∞) , on en déduit que, pour tout $p > \frac{d}{d-1}$, il existe $C, Q > 0$ tels que, pour tous $j \geq 0$, $R > 0$ et tout $f \in L^p(X)$,

$$\left\| \sup_{0 < r < R} |\sigma_{j,r} \cdot f| \right\|_p \leq C 2^{-Qj} \|f\|_p.$$

Le membre de droite ne dépendant pas de R , on a également

$$\left\| \sup_{0 < r < +\infty} |\sigma_{j,r} \cdot f| \right\|_p \leq C 2^{-Qj} \|f\|_p.$$

⁽²⁾Nous avons pris ψ_0 (et donc ψ_j) transformée de Fourier d'une fonction à support compact. C'est ici que ce choix devient crucial.

Il en résulte que

$$\left\| \sup_{0 < r < +\infty} \left| \sigma_r \cdot f - \sum_{j=0}^J \sigma_{j,r} \cdot f \right| \right\|_p \leq \left\| \sum_{j=J+1}^{+\infty} \sup_{0 < r < +\infty} |\sigma_{j,r} \cdot f| \right\|_p \rightarrow 0$$

quand $J \rightarrow +\infty$.

Pour voir que $\lim_{r \rightarrow +\infty} \sigma_r \cdot f = \mathbb{E}(f|\mathcal{I})$ presque partout, introduisons $\Delta f(x) = \lim_{R \rightarrow +\infty} \sup_{r \geq R} |\sigma_r \cdot f(x) - \mathbb{E}(f|\mathcal{I})(x)|$. Pour tout entier J , nous avons, pour presque tout x ,

$$\begin{aligned} \Delta f(x) &\leq \lim_{R \rightarrow +\infty} \sup_{r \geq R} \left| \sigma_r \cdot f(x) - \sum_{j=0}^J \sigma_{j,r} \cdot f(x) \right| \\ &\quad + \lim_{R \rightarrow +\infty} \sup_{r \geq R} \left| \sum_{j=0}^J \sigma_{j,r} \cdot f(x) - \mathbb{E}(f|\mathcal{I})(x) \right| \\ &\leq \sup_{0 < r < +\infty} \left| \sigma_r \cdot f(x) - \sum_{j=0}^J \sigma_{j,r} \cdot f(x) \right| \end{aligned}$$

puisque

$$\lim_{r \rightarrow +\infty} \sum_{j=0}^J \sigma_{j,r} \cdot f = \mathbb{E}(f|\mathcal{I}).$$

Or, quel que soit $\varepsilon > 0$,

$$m(\{\Delta f \geq \varepsilon\}) \leq \frac{\|\Delta f\|_p^p}{\varepsilon^p}$$

et le membre de droite est majoré par une quantité arbitrairement petite. On en déduit que $\Delta f = 0$ presque partout, c'est-à-dire que $\sigma_r \cdot \varphi$ converge presque partout vers $\mathbb{E}(f|\mathcal{I})$. Ceci conclut la démonstration du théorème de Jones. \square

5.2. Inégalités maximales pour les sphères de \mathbb{Z}^d

Pour tout réel $r > 0$, la sphère de \mathbb{Z}^d de rayon r est, par définition, l'ensemble $\{n = (n_1, \dots, n_d) \in \mathbb{Z}^d; |n|^2 = \sum n_i^2 = r^2\}$. Nous noterons $N_d(r)$ son cardinal et Λ l'ensemble $\{r > 0; N_d(r) \neq 0\}$. Si $r \in \Lambda$, alors $r^2 \in \mathbb{N}$ et $N_d(r)$ est le nombre de décompositions de r^2 en somme de d carrés d'entiers. (On s'écarte donc de la notation classique en théorie des nombres). Pour $d \geq 4$, Λ est exactement l'ensemble des réels strictement positifs de carré entier (théorème de Lagrange) et, pour $d \geq 5$ on a des inégalités ([Gro85], [Wal57])

$$C'_d r^{d-2} \leq N_d(r) \leq C_d r^{d-2}, \quad r \in \Lambda.$$

Les opérateurs de moyenne sur les sphères de \mathbb{Z}^d sont les opérateurs de convolution définis, pour toute fonction f sur \mathbb{Z}^d par

$$A_r f(n) = \frac{1}{N_d(r)} \sum_{|\ell|=r} f(n - \ell), \quad r \in \Lambda.$$

La fonction maximale associée à cette famille est

$$A^* f(n) = \sup_{r \in \Lambda} |A_r f(n)|.$$

Le résultat est alors le suivant :

Théorème 5.2.1. — *L'opérateur maximal A^* est borné de $\ell^p(\mathbb{Z}^d)$ dans $\ell^p(\mathbb{Z}^d)$ si et seulement si $d \geq 5$ et $p > \frac{d}{d-2}$.*

Remarque 5.2.2. — Comme on voit, les limites de validité de ce théorème sont différentes de celles du théorème correspondant sur les sphères de \mathbb{R}^d . La limitation portant sur la dimension est clairement en rapport avec la régularité de la fonction N_d qui demande $d \geq 5$. Quant à la condition $p > \frac{d}{d-2}$ (au lieu de $p > \frac{d}{d-1}$ dans le cas des sphères de \mathbb{R}^d), elle est liée à l'ordre de grandeur en r^{d-2} du cardinal de la sphère de \mathbb{Z}^d de rayon r . Pour voir que ces limitations sont optimales, il suffit de considérer δ_0 , la fonction de Dirac en 0. Alors, pour $n \in \mathbb{Z}^d$ non nul, $A_{|n|}\delta_0(n) = \frac{1}{N_d(|n|)}$. Or, d'après un résultat classique [HW79, p. 314], on a $N_4(2^k) = 24$, pour tout $k \geq 1$. Pour $d \leq 4$, $N_d \leq N_4$ et $A^*\delta_0$ n'est donc dans aucun $\ell^p(\mathbb{Z}^d)$ pour $p < \infty$. Pour $d \geq 5$, comme $A^*\delta_0(n) \geq \frac{1}{N_d(|n|)} \geq \frac{C^{-1}}{|n|^{\frac{d}{d-2}}}$ pour tout $n \in \mathbb{Z}^d$, $A^*\delta_0$ ne peut appartenir à $\ell^p(\mathbb{Z}^d)$ que si $p > \frac{d}{d-2}$.

On suppose donc désormais que $d \geq 5$. Modifiant légèrement la notation, on remplacera les moyennes sur les sphères de \mathbb{Z}^d par les opérateurs équivalents

$$A_r f(n) = \frac{1}{r^{d-2}} \sum_{|\ell|=r} f(n - \ell), \quad r \in \Lambda,$$

et on notera encore A^* l'opérateur maximal qui leur est associé. La démonstration du théorème 5.2.1 consiste à se ramener *in fine* au théorème maximal pour les sphères de \mathbb{R}^d . On utilise pour cela un théorème de comparaison entre certains opérateurs de convolution sur \mathbb{R}^d et une version discrétisée de ceux-ci. Nous commencerons par donner en préliminaire, comme dans [MSW02], ces outils généraux qui relèvent de l'analyse harmonique classique. Ensuite, la démarche consistera à décomposer les opérateurs A_r suivant la *méthode du cercle* de Hardy et Littlewood, puis à approcher les opérateurs apparaissant dans cette décomposition par des opérateurs de convolution par les mesures uniformes sur les sphères de \mathbb{R}^d .

Notations 5.2.3. — Dans la suite de cette section on notera Q le cube $]-\frac{1}{2}, \frac{1}{2}]^d$, domaine fondamental de \mathbb{Z}^d dans \mathbb{R}^d . On utilisera le même symbole \wedge pour les transformées de Fourier des mesures et fonctions sur \mathbb{R}^d et celles des fonctions sur \mathbb{Z}^d .

Ainsi, on écrira $\widehat{\sigma}_r(\xi) = \int_{\mathbb{R}^d} e^{-2\pi i x \cdot \xi} d\sigma_r(x)$ pour la transformée de Fourier de la probabilité uniforme sur la sphère de \mathbb{R}^d de rayon r et, pour une fonction $f \in \ell^1(\mathbb{Z}^d)$, on aura $\widehat{f}(\xi) = \sum_{n \in \mathbb{Z}^d} f(n) e^{-2i\pi n \cdot \xi}$; cette dernière fonction pourra être vue suivant les cas comme une fonction sur Q ou comme une fonction \mathbb{Z}^d -périodique sur \mathbb{R}^d .

5.2.1. Opérateurs de convolution sur \mathbb{R}^d et \mathbb{Z}^d . — Notre but, dans cette sous-section, est d'associer à certains opérateurs de convolution sur \mathbb{R}^d des opérateurs de convolution sur \mathbb{Z}^d . Nous allons définir deux opérations, sur les fonctions et sur les opérateurs de convolution, permettant les passages entre \mathbb{Z}^d et \mathbb{R}^d . Nous nous placerons dans le cadre des fonctions à valeurs vectorielles, que nous utiliserons dans la démonstration du théorème maximal.

5.2.1.1. Extension de fonctions de \mathbb{Z}^d à \mathbb{R}^d . — On définit d'abord un opérateur standard de prolongement des fonctions de $\ell_B^1(\mathbb{Z}^d)$ (à valeurs dans un espace de Banach B , la norme dans B étant notée $|\cdot|_B$) en fonctions définies sur \mathbb{R}^d , ceci par convolution avec un bon noyau. Plus précisément, soit Δ la fonction définie, pour $x = (x_1, \dots, x_d) \in \mathbb{R}^d$, par

$$\Delta(x) = \prod_{i=1}^d \left(\frac{\sin \pi x_i}{\pi x_i} \right)^2$$

lorsque $x \neq 0$ et $\Delta(0) = 1$. Sa transformée de Fourier

$$\widehat{\Delta}(\xi) = \prod_{i=1}^d (1 - |\xi_i|)^+$$

a son support dans le cube $2Q$. La décroissance à l'infini de la fonction Δ permet de trouver une constante A pour laquelle on a

$$\int_{\mathbb{R}^d} \Delta(x) dx \leq A \quad \text{et} \quad \sup_{x \in \mathbb{R}^d} \sum_{n \in \mathbb{Z}^d} \Delta(x - n) \leq A.$$

D'autre part, fixons une fois pour toutes une fonction D sur \mathbb{R}^d dont la transformée de Fourier \widehat{D} est de classe \mathcal{C}^∞ , à support compact et égale à 1 sur le cube $2Q$, de sorte que $\widehat{D}\widehat{\Delta} = \widehat{\Delta}$ et donc $D * \Delta = \Delta$. Quitte à augmenter la valeur de A , on peut supposer que

$$\int_{\mathbb{R}^d} |D(x)| dx \leq A \quad \text{et} \quad \sup_{x \in \mathbb{R}^d} \sum_{n \in \mathbb{Z}^d} |D(x - n)| \leq A.$$

À toute fonction $f \in \ell_B^1(\mathbb{Z}^d)$ on associe la fonction définie sur \mathbb{R}^d par

$$f_{\text{ext}}(x) = \sum_{n \in \mathbb{Z}^d} f(n) \Delta(x - n),$$

dont on voit aisément que c'est un prolongement de f : pour tout $n \in \mathbb{Z}^d$,

$$f_{\text{ext}}(n) = f(n).$$

De plus les normes de f et f_{ext} sont comparables :

Lemme 5.2.4. — Soit $f \in \ell_B^p(\mathbb{Z}^d)$, avec $1 \leq p \leq \infty$. Alors $f_{\text{ext}} \in L_B^p(\mathbb{R}^d)$ et

$$A^{-1} \|f\|_{\ell_B^p} \leq \|f_{\text{ext}}\|_{L_B^p} \leq A \|f\|_{\ell_B^p}$$

Démonstration. — L'inégalité $\|f_{\text{ext}}\|_{L_B^p} \leq A \|f\|_{\ell_B^p}$ est immédiate pour $p = \infty$. Pour $p < \infty$, elle se déduit, par intégration, de l'inégalité de Hölder

$$|f_{\text{ext}}(x)|_B^p \leq \left(\sum_{n \in \mathbb{Z}^d} |f(n)|_B^p \Delta(x-n) \right) \left(\sum_{n \in \mathbb{Z}^d} \Delta(x-n) \right)^{p-1}.$$

Pour montrer l'inégalité inverse, on utilise la relation

$$(5.2.1) \quad f(n) = \int_{\mathbb{R}^d} f_{\text{ext}}(x) D(n-x) dx$$

et l'inégalité de Hölder, comme précédemment, ainsi que les majorations sur D .

Pour démontrer l'identité 5.2.1, il suffit de voir par exemple que

$$\widehat{f_{\text{ext}}} = \widehat{f} \widehat{\Delta} = \widehat{f_{\text{ext}}} \widehat{D}.$$

Remarquons que la démonstration permet aussi de voir que f_{ext} est bien définie pour $f \in \ell_B^p(\mathbb{Z}^d)$. □

5.2.1.2. Discrétisation d'opérateurs de convolution. — Inversement, nous associons maintenant à certains opérateurs de convolution sur \mathbb{R}^d leur *discrétisé*, c'est-à-dire un opérateur de convolution sur \mathbb{Z}^d qui est en quelque sorte leur restriction. Soient B_1 et B_2 deux espaces normés de dimension finie et \mathcal{B} l'espace des applications linéaires de B_1 dans B_2 . Soit T un opérateur de convolution de $L_{B_1}^p(\mathbb{R}^d)$ dans $L_{B_2}^p(\mathbb{R}^d)$ défini par un noyau K (à valeurs dans \mathcal{B}) dont la transformée de Fourier $m = \widehat{K}$ est une fonction bornée et à support dans le cube Q . La fonction m est encore appelée *multiplicateur* associé à K . Comme \widehat{K} est à support compact, K est de classe C^∞ . De plus les valeurs prises par K aux points de \mathbb{Z}^d sont égales aux coefficients de Fourier de la fonction \mathbb{Z}^d -périodique qui coïncide avec \widehat{K} sur Q . On appellera *discrétisé* de l'opérateur T vérifiant les conditions précédentes l'opérateur de convolution de $\ell_{B_1}^p(\mathbb{Z}^d)$ dans $\ell_{B_2}^p(\mathbb{Z}^d)$, noté T_{dis} , dont le noyau est la restriction de K à \mathbb{Z}^d .

Ainsi, si $f \in \ell_{B_1}^p(\mathbb{Z}^d)$, alors $\widehat{T_{\text{dis}}(f)}$ est la fonction \mathbb{Z}^d -périodique qui coïncide avec $\xi \mapsto m(\xi) \cdot \widehat{f}(\xi)$ sur Q (c'est-à-dire avec l'action de l'application linéaire $m(\xi)$ sur le vecteur $\widehat{f}(\xi)$). Autrement dit, pour $\xi \in \mathbb{R}^d$, on a

$$\widehat{T_{\text{dis}}(f)}(\xi) = m_{\text{per}}(\xi) \cdot \widehat{f}(\xi),$$

où m_{per} est la \mathbb{Z}^d -périodisée de m , définie par

$$m_{\text{per}}(\xi) = \sum_{n \in \mathbb{Z}^d} m(\xi - n).$$

On a alors :

Proposition 5.2.5. — Soient $1 \leq p \leq \infty$, B_1 et B_2 des espaces normés de dimension finie. Si T est un opérateur borné de $L_{B_1}^p(\mathbb{R}^d)$ dans $L_{B_2}^p(\mathbb{R}^d)$ comme ci-dessus, alors T_{dis} est borné de $\ell_{B_1}^p(\mathbb{Z}^d)$ dans $\ell_{B_2}^p(\mathbb{Z}^d)$ et on a l'inégalité de normes

$$\|T_{\text{dis}}\|_{\ell_{B_1}^p \rightarrow \ell_{B_2}^p} \leq C_d \|T\|_{L_{B_1}^p \rightarrow L_{B_2}^p}$$

où la constante C_d ne dépend ni de p , ni de B_1 , ni de B_2 .

Démonstration. — Par définition

$$((T_{\text{dis}}(f))_{\text{ext}})^\wedge(\xi) = \widehat{\Delta}(\xi) (T_{\text{dis}}(f))^\wedge(\xi) = \widehat{\Delta}(\xi) m_{\text{per}}(\xi) \cdot \widehat{f}(\xi).$$

Posons

$$\tilde{m}(\xi) = \sum_{\ell \in \{-1, 0, 1\}^d} m(\xi + \ell),$$

qui a son support dans le cube

$$3Q = \bigcup_{\ell \in \{-1, 0, 1\}^d} (Q + \ell).$$

En utilisant le fait que le support de $\widehat{\Delta}$ est inclus dans $3Q$, on peut encore écrire que

$$((T_{\text{dis}}(f))_{\text{ext}})^\wedge(\xi) = \widehat{\Delta}(\xi) \tilde{m}(\xi) \cdot \widehat{f}(\xi) = \tilde{m}(\xi) \cdot \widehat{f_{\text{ext}}}(\xi).$$

Soit \tilde{T} l'opérateur de convolution sur \mathbb{R}^d dont le multiplicateur de Fourier est \tilde{m} . Du fait de l'identité précédente, on a

$$(T_{\text{dis}}(f))_{\text{ext}} = \tilde{T}(f_{\text{ext}}).$$

En utilisant cette égalité en même temps que le lemme 5.2.4, on obtient

$$\begin{aligned} \|T_{\text{dis}}(f)\|_{\ell_{B_2}^p} &\leq A \|(T_{\text{dis}}(f))_{\text{ext}}\|_{L_{B_2}^p} = A \|\tilde{T}(f_{\text{ext}})\|_{L_{B_2}^p} \\ &\leq A \|\tilde{T}\|_{L_{B_1}^p \rightarrow L_{B_2}^p} \|f_{\text{ext}}\|_{L_{B_1}^p} \\ &\leq A^2 \|\tilde{T}\|_{L_{B_1}^p \rightarrow L_{B_2}^p} \|f\|_{\ell_{B_1}^p}. \end{aligned}$$

D'autre part il est clair, en découpant \tilde{m} en 3^d termes, que

$$\|\tilde{T}\|_{L_{B_1}^p \rightarrow L_{B_2}^p} \leq 3^d \|T\|_{L_{B_1}^p \rightarrow L_{B_2}^p}.$$

Ceci permet d'achever la démonstration avec la constante $C_d = 3^d A^2$. \square

Nous aurons besoin d'une version en quelque sorte "contractée" de la proposition précédente. Fixons un entier $q \geq 1$, et considérons un opérateur T défini comme précédemment mais en supposant cette fois que le support du multiplicateur associé m est contenu dans le cube Q/q . On définit

$$m_{\text{per}}^q(\xi) = \sum_{\ell \in \mathbb{Z}^d} m(\xi - \ell/q)$$

comme le $q^{-1}\mathbb{Z}^d$ -périodisé de m et on note T_{dis}^q l'opérateur de convolution sur \mathbb{Z}^d associé au multiplicateur m_{per}^q . On a alors la majoration suivante :

Proposition 5.2.6. — Pour un opérateur de convolution T vérifiant la condition précédente, on a

$$\|T_{\text{dis}}^q\|_{\ell_{B_1}^p \rightarrow \ell_{B_2}^p} \leq C_d \|T\|_{L_{B_1}^p \rightarrow L_{B_2}^p}$$

où la constante C_d ne dépend pas de p , B_1 , B_2 et pas non plus de q .

Démonstration. — Par hypothèse, T est l'opérateur de convolution de noyau K et multiplicateur associé m , les deux étant reliés par la relation

$$K(x) = \int_{Q/q} e^{2\pi i x \cdot \xi} m(\xi) d\xi.$$

L'opérateur T_{dis}^q est l'opérateur de convolution sur \mathbb{Z}^d de noyau

$$n \mapsto K_q(n) = \int_Q e^{2\pi i n \cdot \xi} m_{\text{per}}^q(\xi) d\xi.$$

Remarquons en premier lieu que ce dernier noyau est porté par le sous-groupe $q\mathbb{Z}^d$. En effet $K_q(n) = 0$ si chaque coordonnée de n n'est pas divisible par q : c'est une propriété élémentaire des coefficients de Fourier des fonctions \mathbb{Z}^d -périodiques qui sont en fait $q\mathbb{Z}^d$ -périodiques. De plus $K_q(n) = q^d K(n)$ si $n \in q\mathbb{Z}^d$.

Introduisons les opérateurs supplémentaires suivants :

- T_{dis} : opérateur de convolution sur \mathbb{Z}^d , de noyau $n \mapsto K(n)$.
- U : opérateur de convolution sur \mathbb{R}^d , associé au multiplicateur $m(\xi/q)$.
- U_{dis} : opérateur de convolution sur \mathbb{Z}^d , “discretisé” de U , de noyau $n \mapsto q^d K(qn)$.

On montre de façon élémentaire que les opérateurs T et U , qui se déduisent l'un de l'autre par changement d'échelle, ont même norme :

$$\|T\|_{L_{B_1}^p(\mathbb{R}^d) \rightarrow L_{B_2}^p(\mathbb{R}^d)} = \|U\|_{L_{B_1}^p(\mathbb{R}^d) \rightarrow L_{B_2}^p(\mathbb{R}^d)}$$

En effet, si on note θ_{B_i} l'isométrie de $L_{B_i}^p(\mathbb{R}^d)$ sur lui-même définie par $\theta_{B_i} f(x) = q^{-d/p} f(x/q)$, on vérifie que l'on a $\theta_{B_2} U = T \theta_{B_1}$.

La proposition 5.2.5 appliquée à U permet d'autre part d'écrire

$$\|U_{\text{dis}}\|_{\ell_{B_1}^p(\mathbb{Z}^d) \rightarrow \ell_{B_2}^p(\mathbb{Z}^d)} \leq C_d \|U\|_{L_{B_1}^p(\mathbb{R}^d) \rightarrow L_{B_2}^p(\mathbb{R}^d)}.$$

Pour finir la démonstration, il suffit donc de montrer que

$$(5.2.2) \quad \|T_{\text{dis}}^q\|_{\ell_{B_1}^p(\mathbb{Z}^d) \rightarrow \ell_{B_2}^p(\mathbb{Z}^d)} = \|U_{\text{dis}}\|_{\ell_{B_1}^p(\mathbb{Z}^d) \rightarrow \ell_{B_2}^p(\mathbb{Z}^d)}.$$

Afin de simplifier l'écriture, on notera $\|f\|_p$ la norme $\|f\|_{\ell_B^p}$ d'une fonction f dans l'espace $\ell_B^p(\mathbb{Z}^d)$. Pour montrer l'égalité (5.2.2), considérons, pour tout $a \in \mathbb{Z}^d \cap qQ$ (donc pour a décrivant un ensemble de représentants des classes de \mathbb{Z}^d modulo $q\mathbb{Z}^d$), le sous-espace $\ell_{a, B_i}^p(\mathbb{Z}^d)$ de $\ell_{B_i}^p(\mathbb{Z}^d)$ constitué des fonctions à support dans $a + q\mathbb{Z}^d$. Alors, toute fonction f de $\ell_{B_i}^p(\mathbb{Z}^d)$ se décompose sous la forme $f = \sum_a f_a$, avec $f_a \in \ell_{a, B_i}^p(\mathbb{Z}^d)$ et $\|f\|_p^p = \sum_a \|f_a\|_p^p$. On vérifie facilement que T_{dis}^q laisse invariants les sous-espaces $\ell_{a, B_i}^p(\mathbb{Z}^d)$ et que $(T_{\text{dis}}^q f)_a = T_{\text{dis}}^q f_a$. Notons θ_{a, B_i} l'isométrie de $\ell_{B_i}^p(\mathbb{Z}^d)$ sur $\ell_{a, B_i}^p(\mathbb{Z}^d)$

définie par $(\theta_{a,B_i} f)(a + qm) = f(m)$, $m \in \mathbb{Z}^d$. On vérifie que $\theta_{a,B_2} U_{\text{dis}} = T_{\text{dis}}^q \theta_{a,B_1}$. Alors, dans un sens, pour $f \in \ell_{B_1}^p(\mathbb{Z}^d)$,

$$\begin{aligned} \|T_{\text{dis}}^q f\|_p^p &= \sum_a \|T_{\text{dis}}^q f_a\|_p^p = \sum_a \|\theta_{a,B_2} U_{\text{dis}} \theta_{a,B_1}^{-1} f_a\|_p^p \\ &\leq \sum_a \|U_{\text{dis}}\|_{\ell_{B_1}^p \rightarrow \ell_{B_2}^p}^p \|\theta_{a,B_1}^{-1} f_a\|_p^p \\ &\leq \|U_{\text{dis}}\|_{\ell_{B_1}^p \rightarrow \ell_{B_2}^p}^p \|f\|_p^p. \end{aligned}$$

Dans l'autre sens

$$\begin{aligned} \|U_{\text{dis}} f\|_p^p &= \|\theta_{0,B_2} U_{\text{dis}} f\|_p^p = \|T_{\text{dis}}^q \theta_{0,B_1} f\|_p^p \\ &\leq \|T_{\text{dis}}^q\|_{\ell_{B_1}^p \rightarrow \ell_{B_2}^p}^p \|\theta_{0,B_1} f\|_p^p \\ &\leq \|T_{\text{dis}}^q\|_{\ell_{B_1}^p \rightarrow \ell_{B_2}^p}^p \|f\|_p^p. \end{aligned}$$

Ceci achève la démonstration de la proposition 5.2.6. \square

Nous terminons ce paragraphe par un énoncé technique nous aurons besoin dans la démonstration du lemme 5.2.17.

Lemme 5.2.7. — On considère un opérateur de convolution T sur \mathbb{Z}^d dont le multiplicateur de Fourier est défini par la fonction \mathbb{Z}^d -périodique

$$m(\xi) = \sum_{\ell \in \mathbb{Z}^d \cap qQ} \gamma_\ell \Psi(\xi - \ell/q)$$

avec les hypothèses suivantes :

- (a) Ψ est une fonction C^∞ , \mathbb{Z}^d -périodique, à support dans $(1/q)Q$ modulo \mathbb{Z}^d , dont le développement en série de Fourier $\Psi(\xi) = \sum_{m \in \mathbb{Z}^d} \psi(m) e^{2\pi i m \cdot \xi}$ vérifie $\sum_{m \in \mathbb{Z}^d} |\psi(m)| \leq A$.
- (b) $\gamma = (\gamma_\ell)_{\ell \in \mathbb{Z}^d}$ est une fonction sur \mathbb{Z}^d qui est $q\mathbb{Z}^d$ -périodique (et donc s'identifie à une fonction sur le groupe $\mathbb{Z}^d/q\mathbb{Z}^d$).

On note

$$s \mapsto \widehat{\gamma}_s = \sum_{\ell \in \mathbb{Z}^d/q\mathbb{Z}^d} \gamma_\ell e^{-2\pi i s \cdot \ell/q}$$

la transformée de Fourier de γ considérée comme fonction sur le groupe $\mathbb{Z}^d/q\mathbb{Z}^d$. Sous ces hypothèses, pour $1 \leq p \leq 2$, on a

$$\|T\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^p(\mathbb{Z}^d)} \leq AC_d \left(\sup_\ell |\gamma_\ell| \right)^{2-2/p} \left(\sup_s |\widehat{\gamma}_s| \right)^{2/p-1}.$$

Démonstration. — Pour $p = 2$, les termes de la somme $\sum_{\ell \in \mathbb{Z}^d} \gamma_\ell \Psi(\xi - \ell/q)$ étant à supports deux à deux disjoints, on a

$$\sup_{\xi \in \mathbb{R}^d} |m(\xi)| \leq \sup_{\xi \in \mathbb{R}^d} |\Psi(\xi)| \sup_{\ell \in \mathbb{Z}^d} |\gamma_\ell| \leq \left(\sum_{m \in \mathbb{Z}^d} |\psi(m)| \right) \sup_{\ell \in \mathbb{Z}^d} |\gamma_\ell| \leq A \sup_{\ell \in \mathbb{Z}^d} |\gamma_\ell|.$$

Donc

$$\|T\|_{\ell^2(\mathbb{Z}^d) \rightarrow \ell^2(\mathbb{Z}^d)} \leq A \sup_{\ell \in \mathbb{Z}^d} |\gamma_\ell|.$$

Pour $p = 1$, on calcule la norme dans $\ell^1(\mathbb{Z}^d)$ du noyau K de l'opérateur T . On a

$$\begin{aligned} K(n) &= \int_Q m(\xi) e^{2\pi i n \cdot \xi} d\xi = \int_{\mathbb{R}^d/\mathbb{Z}^d} \left(\sum_{\ell \in \mathbb{Z}^d/q\mathbb{Z}^d} \gamma_\ell \Psi(\xi - \ell/q) \right) e^{2\pi i n \cdot \xi} d\xi \\ &= \psi(-n) \left(\sum_{\ell \in \mathbb{Z}^d/q\mathbb{Z}^d} \gamma_\ell e^{2\pi i n \cdot \ell/q} \right) = \psi(-n) \widehat{\gamma}_{-n}. \end{aligned}$$

Donc

$$\|T\|_{\ell^1(\mathbb{Z}^d) \rightarrow \ell^1(\mathbb{Z}^d)} = \sum_{n \in \mathbb{Z}^d} |K(n)| \leq A \sup_n |\widehat{\gamma}_n|.$$

Le résultat pour $1 \leq p \leq 2$ s'en déduit par le théorème d'interpolation de Marcinkiewicz. \square

5.2.2. Décomposition de l'opérateur A_r . — Dans la suite de cette section et sans que ce soit toujours rappelé, le paramètre réel positif r sera astreint à appartenir à l'ensemble Λ , c'est-à-dire sera tel que r^2 est entier.

Le multiplicateur de Fourier associé à l'opérateur de convolution A_r , défini par la série de Fourier

$$a_r(\xi) = \frac{1}{r^{d-2}} \sum_{n \in \mathbb{Z}^d, |n|=r} e^{-2\pi i n \cdot \xi}, \quad \xi \in \mathbb{R}^d$$

peut s'écrire

$$a_r(\xi) = \frac{1}{r^{d-2}} \sum_{n \in \mathbb{Z}^d} e^{2\pi i n \cdot \xi} \int_0^1 e^{2\pi i (|n|^2 - r^2)t} dt.$$

Pour $\varepsilon > 0$ (qui sera choisi plus tard) on peut introduire un facteur de régularisation et écrire

$$\begin{aligned} a_r(\xi) &= \frac{1}{r^{d-2}} \sum_{n \in \mathbb{Z}^d} e^{2\pi i n \cdot \xi} e^{-2\pi \varepsilon (|n|^2 - r^2)} \int_0^1 e^{2\pi i (|n|^2 - r^2)t} dt \\ &= \frac{1}{r^{d-2}} \int_0^1 \left(\sum_{n \in \mathbb{Z}^d} e^{-2\pi (\varepsilon - it)|n|^2} e^{2\pi i n \cdot \xi} \right) e^{2\pi (\varepsilon - it)r^2} dt. \end{aligned}$$

L'opérateur A_r est ainsi représenté par l'intégrale sur $[0, 1] \sim \mathbb{R}/\mathbb{Z}$,

$$(5.2.3) \quad A_r = \frac{1}{r^{d-2}} \int_0^1 e^{2\pi (\varepsilon - it)r^2} T_t dt,$$

où (T_t) est la famille des opérateurs de convolution par les fonctions $n \mapsto e^{-2\pi (\varepsilon - it)|n|^2}$.

Dans une première étape, étant donné un entier R , on étudiera l'opérateur maximal partiel $f \mapsto \sup_{R \leq r \leq 2R} |A_r f|$. On obtiendra ainsi le premier résultat de Magyar (Théorème 5.2.10).

En tout dernier lieu (section 5.2.8), en donnant à R les valeurs 2^k , $k \geq 0$ et en recollant les estimations relatives aux intervalles $]2^k, 2^{k+1}]$, on obtiendra le résultat global du théorème 5.2.1.

Un entier R étant donc fixé, l'application de la *méthode du cercle* s'appuiera sur la décomposition de l'intégrale sur $[0, 1]$ de l'expression (5.2.3) en une somme d'intégrales prises chacune sur un des intervalles associés aux nombres de Farey d'ordre inférieur à R (cf. [HW79, chap. III] ou [Gro85, chap. 12]).

Précisons les notations. Soit F_R l'ensemble des nombres de Farey d'ordre inférieur à R , c'est-à-dire l'ensemble des fractions irréductibles s/q , avec $1 \leq q \leq R$ et $0 \leq s \leq q$. Si $\frac{h}{k} < \frac{s}{q} < \frac{h'}{k'}$ sont trois termes consécutifs de cet ensemble, on sait que $sk - qh = h'q - sk' = 1$. On introduit les nombres de Farey intermédiaires $\frac{h+s}{k+q}$ et $\frac{s+h'}{q+k'}$ qui sont d'ordre supérieur à R , ce qui fait que $q+k$ et $q+k'$ sont compris entre R et $2R$. On obtient

$$\frac{h}{k} < \frac{h+s}{k+q} < \frac{s}{q} < \frac{s+h'}{q+k'} < \frac{h'}{k'}$$

et on définit l'intervalle $I(s/q)$, pour $1 \leq s < q$, par

$$s/q + I(s/q) = \left[\frac{h+s}{k+q}, \frac{s+h'}{q+k'} \right].$$

Pour $s = q = 1$, on prend $I(1/1) = [-\frac{1}{R+1}, \frac{1}{R+1}]$ qui est égal à $[0, \frac{1}{R+1}] \cup [1 - \frac{1}{R+1}, 1]$, modulo 1. Il vient alors :

$$\left[-\frac{1}{2qR}, \frac{1}{2qR} \right] \subset I(s/q) \subset \left[-\frac{1}{qR}, \frac{1}{qR} \right].$$

Compte tenu des notations ainsi introduites, on peut donc écrire

$$\begin{aligned} a_r(\xi) &= \frac{1}{r^{d-2}} \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \int_{s/q + I(s/q)} \left(\sum_{n \in \mathbb{Z}^d} e^{-2\pi(\varepsilon-it)|n|^2} e^{2\pi in \cdot \xi} \right) e^{2\pi(\varepsilon-it)r^2} dt \\ &= \frac{1}{r^{d-2}} \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \int_{I(s/q)} \left(\sum_{n \in \mathbb{Z}^d} e^{-2\pi(\varepsilon-i\tau-is/q)|n|^2} e^{2\pi in \cdot \xi} \right) e^{2\pi(\varepsilon-i\tau-is/q)r^2} d\tau. \end{aligned}$$

On voit ainsi apparaître dans le développement de a_r l'expression $F(\varepsilon - i\tau - is/q, \xi)$ où $F(z, \xi)$ est la fonction analytique dans le demi-plan $\Re z > 0$:

$$F(z, \xi) = \sum_{n \in \mathbb{Z}^d} e^{-2\pi|n|^2 z} e^{2\pi in \cdot \xi}.$$

Il s'agit d'une *fonction thêta* qui, classiquement, peut être transformée par la formule sommatoire de Poisson comme l'indique le lemme suivant.

Lemme 5.2.8. — Pour toute fraction irréductible s/q de $]0, 1]$, on a la relation

$$F(z - is/q, \xi) = \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) \frac{1}{(2z)^{d/2}} e^{-\pi|\xi - \ell/q|^2 / (2z)}$$

où $G(s/q, \ell)$ est la somme de Gauss normalisée

$$G(s/q, \ell) = \frac{1}{q^d} \sum_{m \in \mathbb{Z}^d / q\mathbb{Z}^d} e^{2\pi i |m|^2 s/q} e^{2i\pi m \cdot \ell / q}.$$

Démonstration. — On a

$$F(z - is/q, \xi) = \sum_{m \in \mathbb{Z}^d \cap qQ} \sum_{k \in \mathbb{Z}^d} e^{-2\pi z |m+qk|^2} e^{2i\pi |m|^2 s/q} e^{2i\pi (m+qk) \cdot \xi},$$

où $\mathbb{Z}^d \cap qQ$ est l'ensemble, qui s'identifie à $\mathbb{Z}^d / q\mathbb{Z}^d$, des $(n_1, \dots, n_d) \in \mathbb{Z}^d$ avec $-q/2 \leq n_i < q/2$ pour $i = 1, \dots, d$.

On applique alors la formule sommatoire de Poisson sur \mathbb{R}^d à la fonction

$$f(x) = e^{-2\pi z |m+qx|^2} e^{2i\pi (m+qx) \cdot \xi}$$

dont la transformée de Fourier est

$$\widehat{f}(y) = \frac{1}{q^d (2z)^{d/2}} e^{2i\pi m \cdot y / q} e^{-\pi |y/q - \xi|^2 / (2z)}.$$

On trouve

$$\sum_{k \in \mathbb{Z}^d} e^{-2\pi z |m+qk|^2} e^{2i\pi (m+qk) \cdot \xi} = \frac{1}{q^d} \frac{1}{(2z)^{d/2}} \sum_{\ell \in \mathbb{Z}^d} e^{2i\pi m \cdot \ell / q} e^{-\pi |\ell/q - \xi|^2 / (2z)}$$

qui donne, en remplaçant, le résultat du lemme. \square

Au bout du compte, pour un entier R fixé et pour la dissection de Farey associée, on obtient la représentation suivante du multiplicateur a_r :

$$a_r = \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s, q) = 1}} a_r^{s/q}$$

où, pour tout $\xi \in \mathbb{R}^d$,

$$a_r^{s/q}(\xi) = \frac{1}{r^{d-2}} \int_{I(s/q)} \left(\sum_{n \in \mathbb{Z}^d} e^{-2\pi(\varepsilon - i\tau - is/q)|n|^2} e^{2i\pi n \cdot \xi} \right) e^{2\pi(\varepsilon - i\tau)r^2} e^{-2\pi i r^2 s/q} d\tau,$$

qui s'écrit aussi

$$a_r^{s/q}(\xi) = e^{-2\pi i r^2 s/q} \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) J_r(s/q, \xi - \ell/q)$$

avec

$$J_r(s/q, \xi) = \frac{1}{r^{d-2}} \int_{I(s/q)} e^{2\pi r^2(\varepsilon - i\tau)} (2(\varepsilon - i\tau))^{-d/2} e^{-\pi|\xi|^2 / (2(\varepsilon - i\tau))} d\tau.$$

Rappelons enfin la majoration classique des sommes de Gauss, qui joue ici un rôle essentiel (cf. [Gro85] ou [Wal57, p. 17]).

Théorème 5.2.9. — Il existe une constante C_d telle que

$$(5.2.4) \quad |G(s/q, \ell)| \leq C_d q^{-d/2}$$

pour toute fraction irréductible s/q et tout $\ell \in \mathbb{Z}^d$.

5.2.3. Inégalité maximale partielle. — Les premières inégalités maximales qui ont été obtenues sur ces moyennes sphériques l'ont été pour des opérateurs maximaux partiels, relatifs à des tranches $[R, 2R]$ de valeurs du rayon r . Elles sont données par le théorème suivant [Mag97] :

Théorème 5.2.10. — Pour $d \geq 5$ et $p > \frac{d}{d-2}$, il existe une constante $C_{d,p}$ telle que, pour tout $R > 0$ et toute $f \in \ell^p(\mathbb{Z}^d)$,

$$\left\| \sup_{R \leq r \leq 2R} |A_r(f)| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}.$$

Pour $\tau \in I(s/q)$ et $\varepsilon = R^{-2}$, notons S_τ l'opérateur de convolution par la fonction $n \mapsto e^{-2\pi(\varepsilon - i\tau - is/q)|n|^2}$. On peut écrire la formule (5.2.3) sous la forme

$$A_r = \frac{1}{r^{d-2}} \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \int_{I(s/q)} e^{2\pi(\varepsilon - i\tau - is/q)r^2} S_\tau \, d\tau.$$

Lemme 5.2.11. — Avec les notations précédentes, on a

$$\|S_\tau\|_{\ell^2 \rightarrow \ell^2} \leq C_d q^{-d/2} \min(R^d, |\tau|^{-d/2}).$$

Démonstration. — Le multiplicateur de Fourier associé à S_τ s'écrit, d'après le lemme 5.2.8,

$$\begin{aligned} s_\tau(\xi) &= \sum_{n \in \mathbb{Z}^d} e^{-2\pi(\varepsilon - i\tau - is/q)|n|^2} e^{-2\pi i n \cdot \xi} \\ &= \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) \frac{1}{(2(\varepsilon - i\tau))^{d/2}} e^{-\pi|\xi - \ell/q|^2 / (2(\varepsilon - i\tau))}. \end{aligned}$$

D'où, en appliquant l'inégalité (5.2.4),

$$|s_\tau(\xi)| \leq C_d q^{-d/2} \frac{1}{(\varepsilon^2 + \tau^2)^{d/4}} \sum_{\ell \in \mathbb{Z}^d} e^{-\pi|\xi q - \ell|^2 \varepsilon / (2q^2(\varepsilon^2 + \tau^2))}.$$

Sous les conditions $1 \leq q \leq R$, $\varepsilon = \frac{1}{R^2}$ et $\tau \in I(s/q)$, (donc $|\tau| \leq \frac{1}{qR}$), on obtient

$$\sum_{\ell \in \mathbb{Z}^d} e^{-\frac{\pi|\xi q - \ell|^2 \varepsilon}{2q^2(\varepsilon^2 + \tau^2)}} \leq \sum_{\ell \in \mathbb{Z}^d} e^{-\frac{\pi|\xi q - \ell|^2 R^{-2}}{2q^2(R^{-4} + q^{-2}R^{-2})}} \leq \sum_{\ell \in \mathbb{Z}^d} e^{-\frac{\pi|\xi q - \ell|^2}{4}} \leq C_d.$$

Donc

$$\sup_{\xi \in \mathbb{R}^d} |s_\tau(\xi)| \leq C_d q^{-d/2} \frac{1}{(R^{-4} + \tau^2)^{d/4}} \leq C_d q^{-d/2} \min(R^d, |\tau|^{-d/2}),$$

ce qui fournit la majoration de la norme de S_τ comme opérateur sur $\ell^2(\mathbb{Z}^d)$. \square

Démonstration du théorème 5.2.10. — Gardons les notations du lemme. Commençons par majorer la norme de S_τ comme opérateur sur $\ell^p(\mathbb{Z}^d)$. Pour $p = 1$, on a

$$\|S_\tau\|_{\ell^1 \rightarrow \ell^1} \leq \sum_{n \in \mathbb{Z}^d} |e^{-2\pi(\varepsilon - i\tau - is/q)|n|^2}| = \sum_{n \in \mathbb{Z}^d} e^{-2\pi R^{-2}|n|^2}.$$

En appliquant la formule de Poisson, on obtient immédiatement

$$\sum_{n \in \mathbb{Z}^d} e^{-2\pi R^{-2}|n|^2} = R^d 2^{-d/2} \sum_{m \in \mathbb{Z}^d} e^{-\frac{\pi}{2} R^2 |m|^2} \leq R^d 2^{-d/2} \sum_{m \in \mathbb{Z}^d} e^{-\frac{\pi}{2} |m|^2}$$

et donc

$$\|S_\tau\|_{\ell^1 \rightarrow \ell^1} \leq C_d R^d.$$

Pour $1 < p \leq 2$, soit α tel que $\frac{1}{p} = \frac{\alpha}{2} + (1 - \alpha)$, c'est-à-dire $\alpha = \frac{2(p-1)}{p}$. Le théorème d'interpolation de Marcinkiewicz appliqué à l'opérateur S_τ donne

$$\|S_\tau\|_{\ell^p \rightarrow \ell^p} \leq C_{d,p} q^{-d\alpha/2} R^d \min(1, (|\tau|R^2)^{-d\alpha/2}).$$

On en tire, pour $p > \frac{d}{d-1}$, la majoration

$$\begin{aligned} \int_{I(s/q)} \|S_\tau\|_{\ell^p \rightarrow \ell^p} d\tau &\leq C_{d,p} q^{-d\alpha/2} R^d \left(\int_0^{R^{-2}} d\tau + \int_{R^{-2}}^{+\infty} (|\tau|R^2)^{-d\alpha/2} d\tau \right) \\ &\leq C_{d,p} q^{-d\alpha/2} R^{d-2}. \end{aligned}$$

Rappelons que, pour $R \leq r \leq 2R$ et $\varepsilon = R^{-2}$, l'opérateur A_r peut s'écrire

$$A_r = \frac{1}{r^{d-2}} \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \int_{I(s/q)} e^{2\pi(\varepsilon - i\tau - is/q)r^2} S_\tau d\tau.$$

Pour toute fonction f de $\ell^p(\mathbb{Z}^d)$ on a donc

$$\begin{aligned} \left\| \sup_{R \leq r \leq 2R} |A_r f| \right\|_{\ell^p} &\leq C_{d,p} R^{-d+2} \left(\sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \int_{I(s/q)} \|S_\tau\|_{\ell^p \rightarrow \ell^p} d\tau \right) \|f\|_{\ell^p} \\ &\leq C_{d,p} \sum_{1 \leq q \leq R} q^{1-d\alpha/2} \|f\|_{\ell^p}. \\ &\leq C_{d,p} \sum_{q \geq 1} q^{1-d\alpha/2} \|f\|_{\ell^p}. \end{aligned}$$

Donc

$$\left\| \sup_{R \leq r \leq 2R} |A_r f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}$$

pour $1 - d\alpha/2 < -1$, c'est-à-dire $p > \frac{d}{d-2}$. Ceci achève la démonstration du théorème 5.2.10, le cas $p > 2$ se déduisant par interpolation entre l'inégalité obtenue dans le lemme 5.2.11 pour $p = 2$ et l'inégalité triviale en norme infinie. \square

5.2.4. Approximations des opérateurs de convolution $A_r^{s/q}$. — Revenons à la décomposition des opérateurs A_r , pour un entier R donné, en la somme

$$A_r = \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} A_r^{s/q}.$$

On va effectuer successivement deux approximations des opérateurs $A_r^{s/q}$ en contrôlant à chaque fois la norme $\ell^2 - \ell^2$ des termes d'erreur des opérateurs maximaux partiels associés. Pour la première de ces approximations, on fixe une fonction Φ , de classe \mathcal{C}^∞ sur \mathbb{R}^d , à valeurs dans $[0, 1]$, à support dans $Q/2$ (où Q est toujours le cube $]-\frac{1}{2}, \frac{1}{2}]^d$) et valant 1 sur $Q/4$. Pour tout entier $q \geq 1$ et tout $\xi \in \mathbb{R}^d$, on pose $\Phi_q(\xi) = \Phi(q\xi)$. À partir du multiplicateur de Fourier

$$a_r^{s/q}(\xi) = e^{-2\pi i r^2 s/q} \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) J_r(s/q, \xi - \ell/q),$$

on définit le multiplicateur localisé

$$b_r^{s/q}(\xi) = e^{-2\pi i r^2 s/q} \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) \Phi_q(\xi - \ell/q) J_r(s/q, \xi - \ell/q)$$

La seconde approximation consiste à étendre à la droite \mathbb{R} toute entière les intégrales J_r , c'est-à-dire à approcher $J_r(s/q, \xi)$ par

$$(5.2.5) \quad I_r(\xi) = \frac{1}{r^{d-2}} \int_{-\infty}^{+\infty} e^{2\pi r^2(\varepsilon - i\tau)} (2(\varepsilon - i\tau))^{-d/2} e^{-\pi|\xi|^2/(2(\varepsilon - i\tau))} d\tau.$$

Comme on le verra plus loin (Proposition 5.2.15), $I_r(\xi)$ ne dépend pas de ε et c'est, à un facteur constant près, la transformée de Fourier de la mesure de probabilité uniforme sur la sphère de \mathbb{R}^d de centre 0 et de rayon r . On pourra ainsi appliquer le théorème maximal connu pour les moyennes sur les sphères de \mathbb{R}^d . On définit donc le multiplicateur

$$(5.2.6) \quad c_r^{s/q}(\xi) = e^{-2\pi i r^2 s/q} \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) \Phi_q(\xi - \ell/q) I_r(\xi - \ell/q).$$

On note respectivement $B_r^{s/q}$, $C_r^{s/q}$ les opérateurs de convolution agissant sur les fonctions sur \mathbb{Z}^d , définis par les multiplicateurs de Fourier $b_r^{s/q}$ et $c_r^{s/q}$. On pose naturellement aussi $B_r = \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} B_r^{s/q}$ et $C_r = \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} C_r^{s/q}$. On montre

alors la majoration suivante.

Proposition 5.2.12. — Pour $d \geq 5$ et $f \in \ell^2(\mathbb{Z}^d)$ on a, pour tout $R > 0$,

$$\left\| \sup_{R \leq r \leq 2R} |(A_r - C_r)f| \right\|_{\ell^2} \leq C_d R^{2-d/2} \|f\|_{\ell^2}.$$

Cette inégalité va résulter des deux lemmes suivants.

Lemme 5.2.13. — On a

$$\sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \left\| \sup_{R \leq r \leq 2R} \left| (A_r^{s/q} - B_r^{s/q}) f \right| \right\|_{\ell^2} \leq C_d R^{2-d/2} \|f\|_{\ell^2}.$$

Démonstration. — On pose, pour $\tau \in I(s/q)$,

$$\mu_\tau(\xi) = \sum_{\ell \in \mathbb{Z}^d} (1 - \Phi_q(\xi - \ell/q)) e^{-\pi|\xi - \ell/q|^2 / (2(\varepsilon - i\tau))} G(s/q, \ell).$$

Le terme d'indice ℓ de cette somme est porté par l'ensemble $\{\xi; |\xi - \ell/q| \geq 1/8q\}$ puisque Φ_q vaut 1 sur la boule de centre 0 et de rayon $1/8q$. Sur cet ensemble, pour la constante $c = \frac{\pi}{256}$, on a

$$\frac{\pi}{2} |q\xi - \ell|^2 \geq \frac{\pi}{4} |q\xi - \ell|^2 + c$$

et donc

$$\begin{aligned} \left| e^{-\pi|\xi - \ell/q|^2 / (2(\varepsilon - i\tau))} \right| &= e^{-\pi|q\xi - \ell|^2 \varepsilon / (2q^2(\varepsilon^2 + \tau^2))} \\ &\leq e^{-c\varepsilon / (q^2(\varepsilon^2 + \tau^2))} e^{-\pi|q\xi - \ell|^2 \varepsilon / (4q^2(\varepsilon^2 + \tau^2))}. \end{aligned}$$

Sous les conditions $1 \leq q \leq R$, $\varepsilon = \frac{1}{R^2}$ et $\tau \in I(s/q)$, (donc $|\tau| \leq \frac{1}{qR}$), on obtient, comme dans la démonstration du lemme 5.2.11,

$$\sum_{\ell \in \mathbb{Z}^d} e^{-\pi|q\xi - \ell|^2 \varepsilon / (4q^2(\varepsilon^2 + \tau^2))} \leq C_d.$$

Comme $G(s/q, \ell) \leq C_d q^{-d/2}$, on a finalement

$$\sup_{\xi \in \mathbb{R}^d} |\mu_\tau(\xi)| \leq C_d q^{-d/2} e^{-c\varepsilon / (q^2(\varepsilon^2 + \tau^2))}.$$

Soit $f \in \ell^2(\mathbb{Z}^d)$ et, pour tout $\tau \in I(s/q)$, soit F_τ la fonction de $\ell^2(\mathbb{Z}^d)$ de transformée de Fourier $\mu_\tau \hat{f}$. On a

$$\|F_\tau\|_{\ell^2} \leq C_d q^{-d/2} e^{-c\varepsilon / (q^2(\varepsilon^2 + \tau^2))} \|f\|_{\ell^2}.$$

Alors,

$$(A_r^{s/q} - B_r^{s/q}) f = e^{-2\pi i r^2 s/q} \frac{e^{2\pi \varepsilon r^2}}{r^{d-2}} \int_{I(s/q)} e^{-2\pi i r^2 \tau} (2(\varepsilon - i\tau))^{-d/2} F_\tau d\tau.$$

Pour $\varepsilon = \frac{1}{R^2}$ et $R \leq r \leq 2R$,

$$\begin{aligned} \left| (A_r^{s/q} - B_r^{s/q}) f \right| &\leq \frac{e^{8\pi 2^{-d/2}}}{r^{d-2}} \int_{I(s/q)} (R^{-4} + \tau^2)^{-d/4} |F_\tau| d\tau, \\ \left\| \sup_{R \leq r \leq 2R} \left| (A_r^{s/q} - B_r^{s/q}) f \right| \right\|_{\ell^2} &\leq C_d R^{-d+2} \int_{I(s/q)} (R^{-4} + \tau^2)^{-d/4} \|F_\tau\|_{\ell^2} d\tau \\ &\leq C_d q^{-d/2} R^{-d+2} \int_{I(s/q)} (R^{-4} + \tau^2)^{-d/4} e^{-cR^{-2} / (q^2(R^{-4} + \tau^2))} d\tau \|f\|_{\ell^2}. \end{aligned}$$

En utilisant une majoration de la forme $e^{-u} \leq C u^{-d/4}$ il reste donc

$$\left\| \sup_{R \leq r \leq 2R} \left| (A_r^{s/q} - B_r^{s/q})f \right| \right\|_{\ell^2} \leq C_d R^{2-d/2} |I(s/q)| \|f\|_{\ell^2}$$

d'où, puisque les intervalles $I(s/q)$ convenablement translatés forment une partition de $[0, 1]$, l'inégalité cherchée

$$\sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \left\| \sup_{R \leq r \leq 2R} \left| (A_r^{s/q} - B_r^{s/q})f \right| \right\|_{\ell^2} \leq C_d R^{2-d/2} \|f\|_{\ell^2}.$$

□

Lemme 5.2.14. — On a

$$\sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \left\| \sup_{R \leq r \leq 2R} \left| (B_r^{s/q} - C_r^{s/q})f \right| \right\|_{\ell^2} \leq C_d R^{2-d/2} \|f\|_{\ell^2}.$$

Démonstration. — Le raisonnement est analogue au précédent. On pose, pour $\tau \in \mathbb{R} \setminus I(s/q)$,

$$\nu_\tau(\xi) = \sum_{\ell \in \mathbb{Z}^d} \Phi_q(\xi - \ell/q) e^{-\pi|\xi - \ell/q|^2/(2(\varepsilon - i\tau))} G(s/q, \ell).$$

Cette somme est en fait réduite à un seul terme et on a immédiatement

$$\sup_{\xi \in \mathbb{R}^d} |\nu_\tau(\xi)| \leq C_d q^{-d/2}.$$

Pour $f \in \ell^2(\mathbb{Z}^d)$ la fonction G_τ de $\ell^2(\mathbb{Z}^d)$ de transformée de Fourier $\nu_\tau \hat{f}$ vérifie donc

$$\|G_\tau\|_{\ell^2} \leq C_d q^{-d/2} \|f\|_{\ell^2}.$$

Alors

$$(B_r^{s/q} - C_r^{s/q})f = e^{-2\pi i r^2 s/q} \frac{e^{2\pi \varepsilon r^2}}{r^{d-2}} \int_{\mathbb{R} \setminus I(s/q)} e^{-2\pi i r^2 \tau} (2(\varepsilon - i\tau))^{-d/2} G_\tau d\tau.$$

Et, pour $\varepsilon = \frac{1}{R^2}$ et $R \leq r \leq 2R$,

$$\begin{aligned} \left| (B_r^{s/q} - C_r^{s/q})f \right| &\leq \frac{e^{8\pi 2^{-d/2}}}{r^{d-2}} \int_{\mathbb{R} \setminus I(s/q)} (R^{-4} + \tau^2)^{-d/4} |G_\tau| d\tau \\ &\leq \frac{e^{8\pi 2^{-d/2}}}{r^{d-2}} \int_{|\tau| \geq 1/(2qR)} (R^{-4} + \tau^2)^{-d/4} |G_\tau| d\tau. \end{aligned}$$

Par conséquent

$$\begin{aligned} \left\| \sup_{R \leq r \leq 2R} \left| (B_r^{s/q} - C_r^{s/q})f \right| \right\|_{\ell^2} &\leq C_d R^{-d+2} \int_{|\tau| \geq 1/(2qR)} (R^{-4} + \tau^2)^{-d/4} \|G_\tau\|_{\ell^2} d\tau \\ &\leq C_d q^{-d/2} R^{-d+2} \int_{|\tau| \geq 1/(2qR)} \frac{d\tau}{(R^{-4} + \tau^2)^{d/4}} \|f\|_{\ell^2}, \end{aligned}$$

qui est majoré par $C_d R^{1-d/2} q^{-1} \|f\|_{\ell^2}$. En sommant ces majorations, on obtient donc

$$\begin{aligned} \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \left\| \sup_{R \leq r \leq 2R} |(B_r^{s/q} - C_r^{s/q})f| \right\|_{\ell^2} &\leq C_d R^{1-d/2} \left(\sum_{1 \leq q \leq R} \sum_{1 \leq r \leq q} q^{-1} \right) \|f\|_{\ell^2} \\ &\leq C_d R^{2-d/2} \|f\|_{\ell^2}. \end{aligned}$$

□

Démonstration de la proposition 5.2.12. — Soit $f \in \ell^2(\mathbb{Z})$. On a

$$\sup_{R \leq r \leq 2R} |(A_r - C_r)(f)| \leq \sup_{R \leq r \leq 2R} |(A_r - B_r)(f)| + \sup_{R \leq r \leq 2R} |(B_r - C_r)(f)|$$

donc

$$\left\| \sup_{R \leq r \leq 2R} |(A_r - C_r)(f)| \right\|_{\ell^2} \leq \left\| \sup_{R \leq r \leq 2R} |(A_r - B_r)(f)| \right\|_{\ell^2} + \left\| \sup_{R \leq r \leq 2R} |(B_r - C_r)(f)| \right\|_{\ell^2}$$

et il ne reste qu'à majorer chaque terme du second membre par $C_d R^{2-d/2} \|f\|_{\ell^2}$ en vertu des résultats des deux lemmes précédents. □

5.2.5. Une expression de la transformée de Fourier $\widehat{\sigma}$ de la mesure superficielle d'une sphère. — Remarquons d'abord que l'intégrale $I_r(\xi)$ définie par la formule (5.2.5) ne dépend pas de ε (appliquer le théorème de Cauchy à deux chemins de la forme $\tau \mapsto \varepsilon - i\tau$). En prenant $\varepsilon = r^{-2}$ puis en faisant le changement de variable $\tau = r^{-2}t$ on obtient la nouvelle expression

$$I_r(\xi) = \int_{-\infty}^{+\infty} e^{2\pi(1-it)} (2(1-it))^{-d/2} e^{-\pi r^2 |\xi|^2 / (2(1-it))} dt.$$

Or on a le résultat suivant :

Proposition 5.2.15. — *Pour tout réel r , la transformée de Fourier $\widehat{\sigma}_r$ de la probabilité uniforme σ_r sur la sphère de \mathbb{R}^d centrée en 0 et de rayon r admet l'expression*

$$\widehat{\sigma}_r(\xi) = \frac{\Gamma(d/2)}{\pi^{d/2}} \int_{-\infty}^{+\infty} e^{2\pi(1-it)} (2(1-it))^{-d/2} e^{-\pi r^2 |\xi|^2 / (2(1-it))} dt = \frac{\Gamma(d/2)}{\pi^{d/2}} I_r(\xi).$$

Démonstration. — Pour appliquer les formules d'inversion de Fourier on introduit dans l'intégrale $I_r(\xi)$ un facteur de régularisation $e^{-\pi\delta t^2}$ et on pose

$$I_r^\delta(\xi) = \int_{-\infty}^{+\infty} e^{2\pi(1-it)} (2(1-it))^{-d/2} e^{-\pi r^2 |\xi|^2 / (2(1-it))} e^{-\pi\delta t^2} dt.$$

Il est clair que $\lim_{\delta \rightarrow 0} I_r^\delta = I_r$. Deux calculs successifs de transformée de Fourier de gaussiennes donnent la transformée de Fourier inverse de I_r^δ :

$$\int_{\mathbb{R}^d} \int_{-\infty}^{+\infty} e^{2\pi i \xi \cdot x} e^{2\pi(1-it)} (2(1-it))^{-d/2} e^{-\pi r^2 |\xi|^2 / (2(1-it))} e^{-\pi\delta t^2} dt d\xi$$

$$\begin{aligned}
&= \int_{-\infty}^{+\infty} r^{-d} e^{-2\pi(1-it)|x|^2/r^2} e^{2\pi(1-it)t} e^{-\pi\delta t^2} dt \\
&= r^{-d} \delta^{-1/2} e^{2\pi(1-|x|^2/r^2)} e^{-\frac{\pi}{\delta}(1-|x|^2/r^2)^2} = \psi_\delta(x).
\end{aligned}$$

La fonction ψ_δ ainsi définie est positive, intégrable, radiale. Lorsque δ tend vers 0, ψ_δ tend vers 0, uniformément sur le complémentaire de tout voisinage de la sphère de rayon r et son intégrale tend vers $I_r(0)$.

La fonction $t \mapsto (1-it)^{-d/2}$ est la transformée de Fourier inverse de la fonction $g : x \mapsto \frac{(2\pi)^{d/2}}{\Gamma(d/2)} x^{-1+d/2} e^{-2\pi x} \mathbf{1}_{]0,+\infty[}(x)$. D'où

$$I_r(0) = \int_{-\infty}^{+\infty} e^{2\pi(1-it)} (2(1-it))^{-d/2} dt = 2^{-d/2} e^{2\pi} g(1) = \frac{\pi^{d/2}}{\Gamma(d/2)}.$$

Par conséquent, $\frac{\Gamma(d/2)}{\pi^{d/2}} \psi_\delta$ tend faiblement vers σ_r et I_r^δ tend vers $\frac{\pi^{d/2}}{\Gamma(d/2)} \widehat{\sigma}_r$. \square

5.2.6. Le terme principal du multiplicateur a_r . — Revenons à l'opérateur de convolution $C_r = \sum_{1 \leq q \leq R} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} C_r^{s/q}$, partie principale de l'opérateur A_r . Il résulte de la proposition précédente que le multiplicateur de Fourier de $C_r^{s/q}$ défini par (5.2.6) peut s'écrire

$$c_r^{s/q}(\xi) = c_d e^{-2\pi i r^2 s/q} \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) \Phi_q(\xi - \ell/q) \widehat{\sigma}_r(\xi - \ell/q).$$

Proposition 5.2.16. — Si $d \geq 5$ et $\frac{d}{d-2} < p \leq 2$, on a pour $f \in \ell^p(\mathbb{Z}^d)$,

$$\left\| \sup_{0 < r < \infty} |C_r(f)| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}.$$

Démonstration. — Comme

$$\sup_{0 < r < \infty} |C_r(f)| \leq \sum_{q=1}^{\infty} \sum_{\substack{1 \leq s \leq q \\ (s,q)=1}} \sup_{0 < r < \infty} |C_r^{s/q}(f)|,$$

la proposition résulte du lemme suivant. \square

Lemme 5.2.17. — Si $d \geq 3$ et $\frac{d}{d-2} < p \leq 2$, on a pour $f \in \ell^p(\mathbb{Z}^d)$,

$$\left\| \sup_{0 < r < \infty} |C_r^{s/q}(f)| \right\|_{\ell^p} \leq C_{d,p} q^{-d(p-1)/p} \|f\|_{\ell^p}.$$

Démonstration. — Soit $\Phi'(\xi)$ une fonction de classe \mathcal{C}^∞ , à support dans Q et valant 1 sur $Q/2$, ce qui fait que $\Phi' \Phi = \Phi$. On pose $\Phi'_q(\xi) = \Phi'(q\xi)$. L'opérateur $C_r^{s/q}$ est, à un facteur constant près, le produit de deux opérateurs de convolution M_1 et $M_{2,r}$ de multiplicateurs respectifs

$$m_1(\xi) = \sum_{\ell \in \mathbb{Z}^d} G(s/q, \ell) \Phi'_q(\xi - \ell/q)$$

et

$$m_{2,r}(\xi) = \sum_{\ell \in \mathbb{Z}^d} \Phi_q(\xi - \ell/q) \widehat{\sigma}_r(\xi - \ell/q).$$

À l'opérateur M_1 on applique le lemme 5.2.7 avec pour γ_ℓ la somme de Gauss normalisée $G(s/q, \ell)$ et pour Ψ la fonction Φ'_q . La transformée de Fourier $\widehat{\Phi}'$ est à décroissance rapide, donc

$$\sum_{n \in \mathbb{Z}^d} |\widehat{\Phi}'_q(n)| \leq q^{-d} \sum_{n \in \mathbb{Z}^d} |\widehat{\Phi}'(n/q)| \leq Cq^{-d} \sum_{n \in \mathbb{Z}^d} (1 + |n/q|^2)^{-d} \leq A$$

avec une constante A indépendante de q . Par ailleurs, pour $\gamma_\ell = G(s/q, \ell)$, la majoration fondamentale (5.2.4) des sommes de Gauss entraîne que $\sup_{\ell \in \mathbb{Z}^d} |\gamma_\ell| \leq C_d q^{-d/2}$. Comme $\widehat{\gamma}_n = e^{2i\pi|n|^2 s/q}$, donc $\sup_n |\widehat{\gamma}_n| \leq 1$, on a finalement

$$\|M_1\|_{\ell^p(\mathbb{Z}^d) \rightarrow \ell^p(\mathbb{Z}^d)} \leq C_d q^{-(d/2)(2-2/p)} \leq C_d q^{-d(p-1)/p}.$$

Soient d'autre part un entier $N > 0$, E_N l'ensemble fini des réels r de carré entier vérifiant $0 < r \leq N$ et B_N l'espace normé $\ell^\infty(E_N)$. On peut considérer la famille d'opérateurs de convolution $(M_{2,r})_{r \in E_N}$ comme un opérateur de convolution agissant sur l'espace $\ell^p_{B_N}(\mathbb{Z}^d)$. Mais, d'une part, il résulte de l'inégalité maximale pour les sphères de \mathbb{R}^d que, pour $p > \frac{d}{d-1}$, l'opérateur $(\sigma_r)_{r \in E_N}$ est un opérateur de convolution de $L^p(\mathbb{R}^d)$ dans $L^p_{B_N}(\mathbb{R}^d)$ uniformément borné en N . D'autre part, $\Phi_q(\xi)$ est à support dans Q/q et c'est le multiplicateur de Fourier d'un opérateur de $L^p(\mathbb{R}^d)$ dans $L^p(\mathbb{R}^d)$ de norme $\|\widehat{\Phi}_q\|_{L^1}$ indépendante de q . On peut donc appliquer la proposition 5.2.6 au multiplicateur $m = (\phi_q \widehat{\sigma}_r)_{r \in E_N}$: il existe une constante $C_{d,p}$ telle que, pour tout $q \geq 1$, tout $N > 0$ et toute $f \in \ell^p(\mathbb{Z}^d)$,

$$\left\| \sup_{0 < r \leq N} |M_{2,r} f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}.$$

En faisant tendre N vers l'infini et en utilisant le résultat obtenu sur M_1 on obtient l'inégalité voulue. \square

5.2.7. Démonstration du théorème principal 5.2.1. — Commençons par réunir les résultats obtenus, toujours en dimension $d \geq 5$. En norme ℓ^2 , la proposition 5.2.12 nous dit que, pour tout $f \in \ell^2(\mathbb{Z}^d)$,

$$\left\| \sup_{R \leq r \leq 2R} |(A_r - C_r)f| \right\|_{\ell^2} \leq C_d R^{2-d/2} \|f\|_{\ell^2}.$$

Pour $\frac{d}{d-2} < p \leq 2$ et $f \in \ell^p(\mathbb{Z}^d)$ on a d'une part

$$\left\| \sup_{R \leq r \leq 2R} |A_r f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}$$

d'après le théorème 5.2.10 de Magyar et, d'autre part,

$$\left\| \sup_{R \leq r \leq 2R} |C_r f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}$$

comme conséquence de la proposition 5.2.16. En ajoutant, on a donc

$$\left\| \sup_{R \leq r \leq 2R} |(A_r - C_r)f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}.$$

Soit alors $\frac{d}{d-2} < p \leq 2$. On peut choisir p' tel que $\frac{d}{d-2} < p' < p \leq 2$. En utilisant la proposition 5.2.12 et le théorème d'interpolation de Marcinkiewicz entre p' et 2, on voit qu'il existe un réel $\varepsilon(p) > 0$ tel que

$$\left\| \sup_{R \leq r \leq 2R} |(A_r - C_r)f| \right\|_{\ell^p} \leq C_{d,p} R^{-\varepsilon(p)} \|f\|_{\ell^p}.$$

Comme

$$\sup_{1 \leq r < \infty} |(A_r - C_r)f| \leq \sum_{k=0}^{\infty} \sup_{2^k \leq r < 2^{k+1}} |(A_r - C_r)f|$$

et $\sum_{k \geq 0} 2^{-\varepsilon(p)k} < \infty$, on obtient

$$\left\| \sup_{1 \leq r < \infty} |(A_r - C_r)f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}, \text{ pour tout } p \text{ vérifiant } \frac{d}{d-2} < p \leq 2.$$

D'où, en écrivant

$$\sup_{1 \leq r} |A_r f| \leq \sup_{1 \leq r} |(A_r - C_r)f| + \sup_{1 \leq r} |C_r f|$$

et en utilisant à nouveau la proposition 5.2.16, il vient

$$\left\| \sup_{1 \leq r < \infty} |A_r f| \right\|_{\ell^p} \leq C_{d,p} \|f\|_{\ell^p}, \text{ pour tout } p \text{ vérifiant } \frac{d}{d-2} < p \leq 2.$$

L'inégalité maximale dans le cas $p > 2$ s'obtient par interpolation entre cette inégalité pour $p = 2$ et l'inégalité triviale en norme infinie. Ceci achève la démonstration du théorème 5.2.1. \square

5.2.8. Le théorème ergodique pour les sphères discrètes. — Un théorème ergodique pour les moyennes sphériques pour une action de \mathbb{Z}^d a été établi par Magyar dans [Mag02]. C'est le théorème ergodique naturellement associé au théorème maximal 5.2.1. Pour éviter une obstruction de nature arithmétique à la convergence des moyennes ergodiques, il est toutefois nécessaire d'ajouter une hypothèse spectrale.

Soit $(T_\ell)_{\ell \in \mathbb{Z}^d}$ une action du groupe \mathbb{Z}^d sur un espace probabilisé (X, \mathcal{B}, m) , avec préservation de la mesure. Si r est un réel de carré entier l'opérateur de moyenne sur la sphère de rayon r s'écrit, pour une fonction f définie sur X ,

$$A_r f(x) = \frac{1}{N_d(r)} \sum_{|\ell|=r} f(T_\ell x).$$

Soit $g \in L^2(X, m)$. On dit que g est une fonction propre associée au spectre rationnel s'il existe $\theta \in \mathbb{Q}^d$ tel que, pour tout $\ell \in \mathbb{Z}^d$, $g \circ T_\ell = e^{2i\pi \ell \cdot \theta} g$. On note \mathcal{R} la sous-tribu de \mathcal{B} engendrée par les fonctions propres associées au spectre rationnel. Pour tout $f \in L^2(X, m)$, on note ν_f la mesure spectrale de f pour l'action unitaire

de \mathbb{Z}^d sur l'espace de Hilbert $L^2(X, m)$, induite par l'action ponctuelle (T_ℓ) . Il y a équivalence entre les trois propriétés suivantes :

- Pour tout $\theta \in \mathbb{Q}^d$, $\nu_f(\theta) = 0$.
- Toute fonction propre g associée au spectre rationnel est orthogonale à f .
- $\mathbb{E}(f|\mathcal{R}) = 0$.

Le théorème ergodique s'écrit comme suit.

Théorème 5.2.18. — Soient $d \geq 5$ et $p > \frac{d}{d-2}$. Si $f \in L^p(X, m)$ et $\mathbb{E}(f|\mathcal{R}) = 0$, alors $\lim_{r \rightarrow \infty} A_r f = 0$ dans $L^p(m)$ et m -presque partout.

(La convergence en moyenne s'étend à tous les espaces L^p , pour $p \geq 1$.)

Nous ne donnerons pas ici la démonstration de ce théorème. Les arguments utilisés par A. Magyar prolongent directement ceux utilisés pour établir les inégalités maximales. Il s'agit d'obtenir de bonnes estimations, transférables, de

$$\left\| \max_{r_k \leq r \leq r_{k+1}} |A_r f| \right\|_2$$

pour une suite (r_k) croissant très rapidement vers l'infini.

Notice bibliographique. — Nous avons donné au fur et à mesure les références des résultats exposés dans ce chapitre. Nous mentionnons maintenant quelques prolongements de ces résultats.

Le théorème 5.1.2 a été étendu à de nombreuses autres hypersurfaces que les sphères, et même à des opérateurs de moyenne sur des familles d'hypersurfaces qui ne se déduisent pas forcément d'une hypersurface donnée par dilatation-translation. On doit alors imposer une borne inférieure sur la courbure de la variété en chaque point. Nous renvoyons pour cette étude au chapitre XI du livre [Ste93], à l'article de Stein et Wainger [SW78] et aux références qui y sont données. Une autre direction d'étude consiste à ne regarder la fonction maximale sphérique que sur une famille lacunaire de rayons. Des résultats récents ont été obtenus par Seeger, Tao et Wright [STW03].

L'étude d'inégalités maximales sphériques sur des espaces L^p à poids a été initiée par Rubio de Francia, et on pourra consulter les résultats récents de Cowling, García-Cuerva et Gunaman [CGCG02] où l'on trouvera d'autres références.

Notons également qu'il y a deux autres démonstrations du théorème maximal de Bourgain. La première, due à Mockenhaupt, Seeger et Sogge [MSS92], utilise des arguments de régularisation locale, la seconde, due à Schlag [Sch98], utilise des arguments géométriques et combinatoires.

Notons enfin que le théorème ergodique pour les moyennes sur les sphères de \mathbb{R}^d a été renforcé par Jones, Seeger et Wright qui ont établi dans [JSW08] des inégalités variationnelles et des inégalités de saut.

L'extension de ces travaux au cas du groupe de Heisenberg de dimension $2d + 1$ a été d'abord étudiée par Nevo et Thangavelu dans [NT97]. Ils ont obtenu le théorème

maximal sphérique de type fort (p, p) et le théorème ergodique ponctuel associé pour $d \geq 2$ et $p > (2d - 1)/(2d - 2)$. Plus tard, Narayanan et Thangavelu [NT04] et, dans un cadre plus général, Müller et Seeger [MS04], ont démontré que, pour $d \geq 2$, le théorème maximal de type fort (p, p) est satisfait si et seulement si $p > 2d/(2d - 1)$. Ceci entraîne le théorème ergodique ponctuel pour ces valeurs de p et d .

Le théorème maximal et le théorème ergodique pour les moyennes sur les sphères de \mathbb{Z}^d ont aussi été étendus dans diverses directions.

Dans [Mag02], Magyar étudie les moyennes ergodiques sur les “hypersurfaces discrètes” de \mathbb{Z}^d définies par une équation $Q(n_1, n_2, \dots, n_d) = \lambda$, où Q est un polynôme homogène à coefficients entiers et où le paramètre λ varie parmi les valeurs admissibles. Dans ce cadre, qui généralise largement le cas des sphères, Magyar obtient les théorèmes maximaux et les théorèmes ergodiques attendus. La démonstration fait appel à des résultats difficiles de théorie des nombres.

Dans [Ion04], Ionescu met un point final à la recherche des meilleurs espaces fonctionnels dans lesquels l’inégalité maximale pour les moyennes sphériques discrètes est satisfaite. Les espaces de Lorentz (qui sont des espaces d’interpolation entre les L^p) sont considérés.

CHAPITRE 6

THÉORÈMES ERGODIQUES POUR UNE ACTION D'UN GROUPE LIBRE

Ce chapitre se place dans le cadre d'une généralisation très naturelle des théorèmes de von Neumann et de Birkhoff et peut être lu à la suite du chapitre 1. Nous considérons le groupe libre \mathbb{F}_d engendré par un ensemble A de $d \geq 2$ générateurs libres. On pose $S = A \cup A^{-1}$. Rappelons que \mathbb{F}_d est l'ensemble des mots réduits utilisant l'alphabet S (une lettre n'est jamais suivie par son inverse) ; le produit correspond à la juxtaposition des mots suivie des simplifications appropriées. La longueur $|w|$ d'un élément $w \in \mathbb{F}_d$ est le nombre de lettres (distinctes ou non) appartenant à S nécessaires pour écrire w sous forme réduite. Le groupe \mathbb{F}_d est un espace métrique pour la distance définie par la longueur des mots : la distance de v à w est $|v^{-1}w|$.

Les théorèmes ergodiques pour des moyennes de nature géométrique sur le groupe \mathbb{F}_d sont remarquables, car ce groupe n'est pas moyennable. Ils sont établis à partir de l'observation fondamentale suivante, initialement due à Arnold et Krylov : l'algèbre de convolution engendrée par les mesures de probabilité uniformes σ_n sur les sphères de rayon n (autrement dit l'algèbre $\ell^1_{Rad}(\mathbb{F}_d)$ des fonctions radiales sommables) est commutative et admet une théorie spectrale explicite. Cette algèbre est utilisée dans [AK63] afin d'obtenir un théorème d'équidistribution pour certaines actions par isométries de \mathbb{F}_2 sur la sphère euclidienne de rayon 1 dans \mathbb{R}^3 . Dans [Gui69], Guivarc'h a ensuite développé l'étude du théorème ergodique sur l'espace L^2 relatif aux moyennes sphériques σ_n , obtenant ainsi une généralisation du théorème de von Neumann.

Les techniques conduisant aux théorèmes ergodiques ponctuels pour les actions sur (X, m) préservant la mesure finie m utilisent ici, et dans le chapitre 7, la méthode générale suivante :

- 1) on établit le théorème ergodique en moyenne dans $L^2(X, m)$ à l'aide de la théorie spectrale ;
- 2) on établit l'inégalité maximale de type fort (2, 2) ainsi que la convergence ponctuelle pour les fonctions d'une partie dense dans $L^2(X, m)$;

- 3) par interpolation, on passe ensuite au cas L^p en utilisant une inégalité maximale (p, p) , qu'on a souvent pu démontrer en même temps que l'inégalité (2, 2).

Après avoir donné dans la première section la démonstration très simple du théorème d'équidistribution d'Arnold et Krylov, nous analysons dans la deuxième section la théorie spectrale de $\ell_{Rad}^1(\mathbb{F}_d)$ et nous en déduisons des théorèmes ergodiques en moyenne dans L^2 . Les résultats concernent les moyennes relatives aux suites (σ_{2n}) , (β_{2n}) et (ν_n) , où β_n est la mesure de probabilité uniforme sur la boule de rayon n et $\nu_n = \frac{1}{n+1} \sum_{k=0}^n \sigma_k$ est la moyenne de Cesàro des σ_k . Les moyennes relatives aux suites (σ_{2n+1}) et (β_{2n+1}) se comportent comme leurs analogues d'indice pair, mais avec des limites différentes en général.

Dans la section 3, les théorèmes ergodiques ponctuels dans L^2 sont établis en suivant l'article [Nev94a] de Nevo. Bien que les théorèmes maximaux dont on a besoin se trouvent aussi dans la section suivante, on présente ici une démonstration spécifique au cas L^2 . En utilisant la sous-additivité de la suite de moyennes relatives aux mesures ν_n , Nevo obtient d'abord que l'opérateur maximal $f \mapsto f_\nu^*$ correspondant est borné dans L^2 . À l'aide de la méthode de Littlewood-Paley-Stein déjà rencontrée dans le chapitre précédent, il en déduit ensuite la même propriété pour l'opérateur maximal correspondant à σ_n . L'analyse spectrale de $\ell_{Rad}^1(\mathbb{F}_d)$ intervient à nouveau ici, ainsi que dans l'étude de la convergence ponctuelle pour les fonctions appartenant à une partie dense de $L^2(X, m)$.

La section 4 est consacrée à l'étude des résultats correspondants dans L^p , dus à Nevo et Stein [NS94]. On démontre directement, à l'aide de l'inégalité maximale de Hopf-Dunford-Schwartz que l'opérateur maximal $f \mapsto f_\nu^*$ relatif à (ν_n) est de type faible $(1, 1)$. On rappelle ensuite les techniques d'interpolation complexe permettant d'établir que l'opérateur maximal pour les moyennes sphériques est de type fort (p, p) pour $p > 1$. On en déduit le théorème de Nevo-Stein, c'est-à-dire la convergence presque partout de la suite $(\sigma_{2n} \cdot f)$ des moyennes sur les sphères de rayon pair pour $f \in L^p(X, m)$ et $p > 1$.

Les techniques décrites dans ce chapitre ont une portée très générale. Elles s'adaptent à de nombreuses situations comme, par exemple, à celles des moyennes radiales pour les actions de groupes semi-simples. On les verra à l'œuvre pour le groupe des isométries de l'espace hyperbolique réel de dimension d dans le chapitre suivant.

Nous terminons ce chapitre par une approche entièrement différente du théorème de Nevo-Stein, due à Bufetov. Pour l'étude de la convergence presque partout de la suite $(\sigma_{2n} \cdot f)$, la méthode de Bufetov est notablement plus simple et permet d'atteindre le cas où f appartient à la classe $L \log L$. En outre cette méthode est bien adaptée au cas des actions de semi-groupes libres. L'idée de la démonstration consiste à relier les opérateurs de moyenne sur les sphères de rayon pair aux puissances d'un même opérateur de Markov, pour lequel on applique le théorème de convergence de Rota rappelé dans l'appendice C.

6.1. Actions par isométries sur un espace métrique compact

Dans cette section nous considérons un exemple particulier d'action du groupe libre à deux générateurs sur la sphère \mathbb{S}^2 . Nous présentons la démonstration due à Guivarc'h du théorème d'équidistribution d'Arnold-Krylov. Ce résultat est à comparer avec ceux obtenus pour les actions de groupes moyennables dans le corollaire 2.2.8 et les remarques 2.2.9 et 2.2.10.

Soient X un espace métrique compact et $\text{Iso}(X)$ le groupe compact (pour la topologie de la convergence uniforme sur X) de ses isométries. Ce groupe agit par isométries sur l'espace $C(X)$: pour $g \in \text{Iso}(X)$ et $f \in C(X)$, on pose $T_g f = f \circ g$. Si l'action de $\text{Iso}(X)$ sur X est transitive, elle laisse invariante une unique mesure de probabilité, qui sera notée m (voir la remarque 2.2.10). Dans ce cas, l'opérateur T_g se prolonge en un opérateur unitaire de $L^2(X, m)$, et nous conservons la même notation pour le prolongement.

Soient a, b deux isométries de X et soit G le sous-groupe de $\text{Iso}(X)$ qu'elles engendrent. Observons que si l'action de G sur X est *topologiquement transitive*, c'est-à-dire s'il existe $x_0 \in X$ tel que l'orbite Gx_0 soit dense dans X , alors $\text{Iso}(X)$ agit transitivement sur X .

Pour tout entier $n \geq 1$, nous notons M_n le sous-ensemble de G formé des isométries de la forme $a^{p_1} b^{q_1} \dots a^{p_k} b^{q_k}$ avec $k \geq 1$, $p_i, q_i \in \mathbb{N}$ et $\sum_{i=1}^k (p_i + q_i) = n$ (mots en a et b de longueur n). Enfin, G_0 désigne le sous-groupe de G formé des isométries $a^{p_1} b^{q_1} \dots a^{p_k} b^{q_k}$ avec $k \geq 1$, $p_i, q_i \in \mathbb{Z}$ et $\sum_{i=1}^k (p_i + q_i) = 0$ (mots de degré total nul).

Théorème 6.1.1 ([Gui69]). — *Si l'action de G_0 sur X est topologiquement transitive alors, pour toute fonction continue f sur X , on a*

$$\lim_{n \rightarrow \infty} \left(\frac{T_a + T_b}{2} \right)^n f = \int_X f \, dm$$

uniformément sur X .

Démonstration. — Posons $T = \frac{T_a + T_b}{2}$. La fonction $T^n f$ est une moyenne de 2^n fonctions de la forme $f \circ g$ avec $g \in \text{Iso}(X)$. Par conséquent l'ensemble $\{T^n f; n \in \mathbb{N}\}$ est équicontinu. Étant borné, il est relativement compact, d'après le théorème d'Ascoli. Soit alors $(T^{n_k} f)_{k \geq 1}$ une suite extraite de $(T^n f)_{n \geq 1}$, qui converge vers un élément f_0 de $C(X)$. Le point crucial de la démonstration est la vérification que f_0 est G_0 -invariante. Admettons ce point pour le moment. Comme l'action de G_0 est topologiquement transitive, on en déduit que f_0 est constante, et la G_0 -invariance de m entraîne aussitôt que cette constante vaut $\int_X f \, dm$. Enfin, comme la limite ne dépend pas du choix de la suite extraite, on conclut que $(T^n f)$ converge vers $\int_X f \, dm$.

Vérifions donc la G_0 -invariance de f_0 . Remarquons d'abord que $(T^{n_k} f)_{k \geq 1}$ converge aussi vers f_0 dans $L^2(X, m)$. Fixons un entier $r > 0$. Quitte à remplacer $(T^{n_k} f)_{k \geq 1}$ par une sous-suite, on peut supposer que $n_{k+1} - n_k \geq r$ pour tout k .

Puisque T est une contraction de $L^2(X, m)$, on a

$$\|f_0\|_2 = \lim_{k \rightarrow \infty} \|T^{n_{k+1}} f\|_2 \leq \lim_{k \rightarrow \infty} \|T^r T^{n_k} f\|_2 = \|T^r f_0\|_2 \leq \|f_0\|_2.$$

Observons que $T^r f_0$ est une combinaison convexe des éléments $T_g f_0$, où g décrit M_r , tous de norme L^2 égale à $\|f_0\|_2$. En utilisant la convexité stricte de $L^2(X, m)$, nous obtenons $T_{g_1} f_0 = T_{g_2} f_0$ pour tous $g_1, g_2 \in M_r$.

Ainsi, f_0 est invariante par les éléments de $\bigcup_n M_n M_n^{-1}$. Pour conclure, nous montrons que G_0 est contenu dans l'adhérence de $\bigcup_n M_n M_n^{-1}$. Soit $g = a^{p_1} b^{q_1} \dots a^{p_r} b^{q_r}$ avec $r \geq 1$, $p_i, q_i \in \mathbb{Z}$ et $\sum_{i=1}^r (p_i + q_i) = 0$. Puisque $\text{Iso}(X)$ est un groupe compact, nous pouvons choisir des suites d'entiers (m_k) et (n_k) qui tendent vers $+\infty$, telles que les suites (a^{m_k}) et (b^{n_k}) convergent dans $\text{Iso}(X)$ vers l'élément neutre. On a alors

$$g = \lim_{k \rightarrow \infty} a^{p_1 + m_k} b^{q_1 + n_k} \dots a^{p_r + m_k} b^{q_r + n_k} (a^{m_k} b^{n_k})^{-r}.$$

Il suffit maintenant d'observer que pour k assez grand, $a^{p_1 + m_k} b^{q_1 + n_k} \dots a^{p_r + m_k} b^{q_r + n_k}$ et $(a^{m_k} b^{n_k})^r$ appartiennent à $M_{r(m_k + n_k)}$. \square

Indiquons maintenant comment le résultat d'équidistribution d'Arnold-Krylov se déduit du théorème précédent.

Théorème 6.1.2 ([Gui69]). — Soient a et b deux rotations dans $SO(3)$. On suppose que le sous-groupe de $SO(3)$ qu'elles engendrent a une orbite dense dans la sphère euclidienne unité \mathbb{S}^2 de \mathbb{R}^3 . Alors, pour tout $x \in \mathbb{S}^2$, les $M_n x$ s'équirépartissent suivant la mesure de probabilité uniforme m sur \mathbb{S}^2 . Plus précisément, pour toute fonction continue f sur \mathbb{S}^2 on a

$$\lim_{n \rightarrow \infty} \left(\frac{T_a + T_b}{2} \right)^n f = \int_{\mathbb{S}^2} f \, dm$$

uniformément sur \mathbb{S}^2

Démonstration. — Notons toujours G le sous-groupe de $SO(3)$ engendré par les isométries a et b . Il suffit de vérifier que si l'action de G sur \mathbb{S}^2 est topologiquement transitive, il en est de même pour l'action de G_0 .

Si H est un sous-groupe de G et si $x \in \mathbb{S}^2$, observons d'abord que $\overline{Hx} = \overline{Hx}$, car l'adhérence \overline{H} de H dans $SO(3)$ est compacte et \overline{Hx} est donc fermé. L'action de H sur \mathbb{S}^2 est donc topologiquement transitive si et seulement si celle de \overline{H} est transitive.

Puisque les éléments de la forme $g_1 g_2 g_1^{-1} g_2^{-1}$ avec $g_1, g_2 \in G$ sont contenus dans G_0 , le sous-groupe G' qu'ils engendrent, c'est-à-dire le sous-groupe dérivé de G , est contenu dans G_0 .

Le groupe \overline{G} agit transitivement sur \mathbb{S}^2 . Comme $SO(3)$ n'a pas d'autre sous-groupe fermé que lui-même agissant transitivement sur \mathbb{S}^2 , on voit que $\overline{G} = SO(3)$. Le groupe G' est donc dense dans le sous-groupe dérivé de $SO(3)$, qui n'est autre que $SO(3)$. Par conséquent, l'action de G' sur \mathbb{S}^2 est topologiquement transitive, et celle de G_0 aussi. \square

Remarque 6.1.3. — Conservons les notations et les hypothèses du théorème 6.1.1. Le groupe \mathbb{F}_2 agit par isométries sur X , puisqu'il s'envoie naturellement sur G . Pour tout n entier, S_n^+ désigne le sous-ensemble de \mathbb{F}_2 formé des mots de longueur n , où n'interviennent que des puissances positives des deux générateurs. Si $f \in C(X)$ et $w \in \mathbb{F}_2$, rappelons que $w \cdot f(x) = f(wx)$. Le théorème 6.1.1 énonce que

$$\lim_{n \rightarrow \infty} \frac{1}{2^n} \sum_{w \in S_n^+} w \cdot f = \int_X f \, dm$$

uniformément sur X . Ce résultat d'unique ergodicité s'étend bien sûr au cas de tout groupe libre \mathbb{F}_d à $d \geq 2$ générateurs.

Le théorème 6.1.2 s'interprète de la manière suivante. *Supposons que l'action de \mathbb{F}_2 sur \mathbb{S}^2 engendrée par a et b soit topologiquement transitive. Si B est une partie borélienne de \mathbb{S}^2 de frontière négligeable et si $x \in \mathbb{S}^2$, la proportion du nombre de points w de S_n^+ tels que $wx \in B$ tend vers la mesure de Lebesgue de B lorsque n tend vers l'infini.*

6.2. Théorèmes ergodiques en moyenne dans L^2

Nous étudions maintenant les théorèmes ergodiques en moyenne dans l'espace L^2 (ou plus généralement dans un espace hilbertien H) pour les moyennes sur les boules ou sur les sphères de \mathbb{F}_d .

6.2.1. Énoncé du théorème. Application à l'équidistribution. — Précisons d'abord les notations utilisées. Tout élément $\mu \in \ell^1(\mathbb{F}_d)$ peut être vu comme une mesure bornée : $\mu = \sum_{w \in \mathbb{F}_d} \mu(w) \delta_w$, où δ_w est la mesure de Dirac en w . Pour simplifier, nous utiliserons en général l'écriture $\mu = \sum_{w \in \mathbb{F}_d} \mu(w) w$. Si $x \in \mathbb{F}_d$, rappelons que $|x|$ désigne la longueur du mot x . Si E est une partie de \mathbb{F}_d , on note comme d'habitude $|E|$ son cardinal.

Posons $S_n = \{x \in \mathbb{F}_d; |x| = n\}$ et $B_n = \{x \in \mathbb{F}_d; |x| \leq n\}$. Pour $n > 0$, l'ensemble S_n contient $(q+1)q^{n-1}$ éléments, où $q = 2d - 1$. Pour $n \in \mathbb{N}$, nous considérerons les mesures de probabilité suivantes

$$\begin{aligned} \sigma_n &= \frac{1}{|S_n|} \sum_{w \in S_n} w \text{ (mesure uniforme sur la sphère de rayon } n) ; \\ \sigma'_n &= \frac{1}{2} (\sigma_n + \sigma_{n+1}) ; \\ \beta_n &= \frac{1}{|B_n|} \sum_{w \in B_n} w \text{ (mesure uniforme sur la boule de rayon } n) ; \\ \nu_n &= \frac{1}{n+1} \sum_{k=0}^n \sigma_k, \text{ (moyenne de Cesàro des } \sigma_i). \end{aligned}$$

Soit π une représentation unitaire du groupe \mathbb{F}_d dans un espace hilbertien H , que l'on prolonge en une représentation π de $\ell^1(\mathbb{F}_d)$ en posant $\pi(\mu) = \sum_{w \in \mathbb{F}_d} \mu(w) \pi(w)$ si $\mu = \sum_{w \in \mathbb{F}_d} \mu(w) w$. Nous notons E_1 le projecteur orthogonal sur le noyau de $\pi(\sigma_1) - I$ et E_{-1} celui sur le noyau de $\pi(\sigma_1) + I$. Grâce à la stricte convexité de la boule unité de H , nous voyons que E_1 (respectivement E_{-1}) est le projecteur orthogonal sur le sous-espace des vecteurs v de H tels que $\pi(w)v = v$ (resp. $\pi(w)v = -v$) pour tout

$w \in S$. Observons que $E_1 E_{-1} = 0$ et que E_1 projette sur le sous-espace des vecteurs invariants sous l'action de \mathbb{F}_d .

Les résultats de convergence en moyenne s'énoncent ainsi :

Théorème 6.2.1 ([Gui69]). — Soit π une représentation unitaire du groupe \mathbb{F}_d .

- (i) La suite $(\pi(\sigma_{2n}))$ converge fortement vers le projecteur $E_1 + E_{-1}$ et la suite $(\pi(\sigma_{2n+1}))$ converge fortement vers l'opérateur $E_1 - E_{-1}$.
- (ii) Les suites $(\pi(\nu_n))$ et $(\pi(\sigma'_n))$ convergent fortement vers E_1 .
- (iii) La suite $(\pi(\beta_{2n}))$ converge fortement vers l'opérateur $E_1 + \frac{q-1}{q+1}E_{-1}$ et la suite $(\pi(\beta_{2n+1}))$ converge fortement vers l'opérateur $E_1 - \frac{q-1}{q+1}E_{-1}$ (où $q = 2d - 1$).

Remarque 6.2.2. — Comme $\text{Ker}(\pi(\sigma_1) - I) \oplus \text{Ker}(\pi(\sigma_1) + I) = \text{Ker}(\pi(\sigma_1)^2 - I)$, on voit, en utilisant la stricte convexité de H , que $E_1 + E_{-1}$ est le projecteur sur le sous-espace des vecteurs invariants par le sous-groupe $\mathbb{F}_d^{(2)}$ de \mathbb{F}_d formé des mots de longueur paire.

Avant de passer à la démonstration du théorème 6.2.1, donnons tout de suite l'application suivante, obtenue par Guivarc'h.

Théorème 6.2.3 ([Gui69]). — Soit X un espace métrique compact sur lequel \mathbb{F}_d agit par isométries. On suppose que le sous-groupe $\mathbb{F}_d^{(2)}$ de \mathbb{F}_d formé des mots de longueur paire agit topologiquement transitivement sur X . Si m désigne l'unique mesure de probabilité sur X invariante par \mathbb{F}_d , pour tout $f \in C(X)$ on a

$$\lim_{n \rightarrow \infty} \pi(\sigma_n) f = \int_X f dm$$

uniformément sur X .

Démonstration. — L'existence et l'unicité de m sont justifiées dans la remarque 2.2.10, puisque l'action de \mathbb{F}_d est topologiquement transitive. On en déduit aussi que m est l'unique mesure de probabilité sur X invariante par l'action groupe $\mathbb{F}_d^{(2)}$. Nous notons π la représentation unitaire associée de \mathbb{F}_d dans $L^2(X, m)$.

Observons que l'espace $E_1(L^2(X, m))$ des éléments \mathbb{F}_d -invariants de $L^2(X, m)$ est l'espace des fonctions constantes sur X , puisque l'action de \mathbb{F}_d est ergodique.

Par ailleurs, si f appartient au sous-espace $E_{-1}(L^2(X, m))$ des fonctions h transformées en $-h$ par les générateurs de \mathbb{F}_d , on voit que f est $\mathbb{F}_d^{(2)}$ -invariante et donc constante, par ergodicité de l'action de $\mathbb{F}_d^{(2)}$. Mais cette constante est nulle puisque $\pi(w)f = -f$ pour tout générateur w . Ainsi $E_{-1} = 0$.

Il résulte du théorème 6.2.1 que $\lim_{n \rightarrow \infty} \pi(\sigma_n) f = \int_X f dm$ dans $L^2(X, m)$, pour tout $f \in L^2(X, m)$. L'utilisation du théorème d'Ascoli, comme dans la démonstration du théorème 6.1.1, montre alors que pour tout $f \in C(X)$ la convergence a lieu uniformément sur X . \square

Remarque 6.2.4. — Lorsque X est connexe, la transitivité topologique de l'action de \mathbb{F}_d sur X entraîne celle de $\mathbb{F}_d^{(2)}$. En effet, avec les notations de la démonstration précédente, $G/G^{(2)}$ est isomorphe à un quotient de $\mathbb{F}_d/\mathbb{F}_d^{(2)}$, donc fini. Sous l'hypothèse $X = \overline{G}x$, on voit que X est réunion d'un nombre fini d'orbites de $\overline{G^{(2)}}$, d'où la conclusion.

Le théorème 6.2.1 est démontré à l'aide de la théorie spectrale. Bien que cela ne nécessite pas une étude spectrale aussi poussée que celle qui est développée dans les sections 6.2.2 et 6.2.3 (voir la remarque 6.2.10), nous aurons besoin de cette étude par la suite, pour établir les théorèmes ergodiques ponctuels. Nous suivons l'article [Nev94a] de Nevo. Toutes les mesures considérées ci-dessus sont *radiales*, c'est-à-dire qu'elles ne dépendent que de la longueur des mots. On note $\mathbb{C}_{Rad}[\mathbb{F}_d]$ l'algèbre involutive formée des fonctions radiales à valeurs complexes à support fini et $\ell^1_{Rad}(\mathbb{F}_d)$ son adhérence dans $\ell^1(\mathbb{F}_d)$.

Tout élément f de $\mathbb{C}_{Rad}[\mathbb{F}_d]$ s'écrit de façon unique $f = \sum_{n \geq 0} c_n \sigma_n$, avec une somme finie, et en outre on a $\|f\|_1 = \sum_{n \geq 0} |c_n|$.

Le fait crucial est la commutativité de l'algèbre $\ell^1_{Rad}(\mathbb{F}_d)$ démontrée dans le lemme 6.2.5 ci-dessous. Nous aurons à estimer la norme d'éléments auto-adjoints de l'algèbre involutive $\pi(\ell^1_{Rad}(\mathbb{F}_d))$ et pour cela nous allons déterminer les caractères auto-adjoints de l'algèbre de Banach involutive et commutative $\ell^1_{Rad}(\mathbb{F}_d)$ (voir B.2.5 dans l'appendice B).

6.2.2. Détermination des caractères auto-adjoints de $\ell^1_{Rad}(\mathbb{F}_d)$. — Nous rappelons ci-dessous les points essentiels. Nous renvoyons par exemple à [FTP83, Chap. 3] pour une étude plus détaillée et à l'appendice B pour les généralités concernant la théorie spectrale.

Lemme 6.2.5. — *L'algèbre $\mathbb{C}_{Rad}[\mathbb{F}_d]$ est commutative avec unité. Elle est engendrée par $\sigma_0 = \delta_e$ et σ_1 . Plus précisément, pour tout $n \geq 1$, on a, en posant $q = 2d - 1$,*

$$(6.2.1) \quad \sigma_1 * \sigma_n = \frac{1}{q+1} \sigma_{n-1} + \frac{q}{q+1} \sigma_{n+1}.$$

Démonstration. — On a $\sigma_1 * \sigma_n(x) = \frac{1}{q+1} \sum_{w \in S} \sigma_n(wx)$. De plus, si $w \in S$ et $\sigma_n(wx) \neq 0$, alors ou bien $|x| = n+1$, ou bien $|x| = n-1$. Si $|x| = n+1$, il y a un seul mot wx de longueur n avec $w \in S$ et si $|x| = n-1$, il y a exactement q mots de longueur n de la forme wx avec $w \in S$. On en déduit la formule (6.2.1). \square

Nous utilisons la relation de récurrence (6.2.1) pour étudier les caractères de $\mathbb{C}_{Rad}[\mathbb{F}_d]$, c'est-à-dire les homomorphismes non nuls φ de l'algèbre $\mathbb{C}_{Rad}[\mathbb{F}_d]$ dans l'algèbre \mathbb{C} . Soit φ un caractère de $\mathbb{C}_{Rad}[\mathbb{F}_d]$. Comme σ_0 est l'unité de cette algèbre,

nous avons $\varphi(\sigma_0) = 1$. De plus, en appliquant φ à l'égalité (6.2.1), nous obtenons

$$(6.2.2) \quad \varphi(\sigma_1)\varphi(\sigma_n) = \frac{1}{q+1}\varphi(\sigma_{n-1}) + \frac{q}{q+1}\varphi(\sigma_{n+1}).$$

Par conséquent, φ est entièrement déterminé par la valeur de $\varphi(\sigma_1)$. En posant $\varphi(\sigma_n) = \lambda_n$, nous avons donc à déterminer les suites (λ_n) satisfaisant à la relation de récurrence

$$(6.2.3) \quad \lambda_{n+1} = \frac{q+1}{q}\lambda\lambda_n - \frac{1}{q}\lambda_{n-1},$$

avec $\lambda_0 = 1$ et $\lambda_1 = \lambda$. Il est pratique de poser $\lambda = \gamma(z)$ avec $\gamma(z) = \frac{q^z + q^{1-z}}{q+1}$, $z \in \mathbb{C}$, de sorte que le discriminant de la relation est égal à $\left(\frac{q^z - q^{1-z}}{q}\right)^2$. Les racines du trinôme associé sont q^{-z} et q^{z-1} .

Lemme 6.2.6. — *Pour tout $z \in \mathbb{C}$, il existe un et un seul caractère φ_z tel que $\varphi_z(\sigma_1) = \gamma(z)$. De plus*

(i) *ou bien $q^{2z-1} \neq 1$ et alors*

$$(6.2.4) \quad \varphi_z(\sigma_n) = c(z)q^{-nz} + c(1-z)q^{-n(1-z)}, \quad \text{pour } n \geq 0,$$

$$\text{avec } c(z) = \frac{q^{1-z} - q^{z-1}}{(q+1)(q^{-z} - q^{z-1})};$$

(ii) *ou bien $q^{2z-1} = 1$, c'est-à-dire $z = \frac{1}{2} + \frac{ij\pi}{\log q}$, $j \in \mathbb{Z}$, et alors*

$$(6.2.5) \quad \varphi_z(\sigma_n) = (-1)^{jn} \left(1 + n \frac{q-1}{q+1}\right) q^{-n/2}.$$

Démonstration. — Supposons $q^{2z-1} \neq 1$. Alors $(q^{-nz})_{n \geq 0}$ et $(q^{-n(1-z)})_{n \geq 0}$ sont deux suites indépendantes satisfaisant à la relation de récurrence

$$\gamma(z)\lambda_n = \frac{1}{q+1}\lambda_{n-1} + \frac{q}{q+1}\lambda_{n+1}.$$

On a donc $\varphi_z(\sigma_n) = aq^{-nz} + bq^{-n(1-z)}$ pour tout n . Les conditions $\varphi_z(\sigma_0) = 1$ et $\varphi_z(\sigma_1) = \gamma(z)$ donnent $a = c(z)$ et $b = c(1-z)$.

Le cas (ii) se traite de façon analogue à partir des solutions indépendantes $(q^{-nz})_{n \geq 0}$ et $(nq^{-nz})_{n \geq 0}$. \square

Remarquons maintenant qu'un caractère φ de $\mathbb{C}_{Rad}[\mathbb{F}_d]$ est continu pour la norme $\|\cdot\|_1$ si et seulement si la suite $(\varphi(\sigma_n))_{n \in \mathbb{N}}$ est bornée. Les caractères continus de $\mathbb{C}_{Rad}[\mathbb{F}_d]$ sont donc exactement les φ_z avec $0 \leq \Re z \leq 1$. De plus, ils s'identifient à ceux (automatiquement continus) de l'algèbre de Banach commutative $\ell^1_{Rad}(\mathbb{F}_d)$. Comme cette dernière est engendrée par son unité et par σ_1 , l'application $\varphi \mapsto \varphi(\sigma_1)$

est une bijection de l'ensemble des caractères de $\ell_{Rad}^1(\mathbb{F}_d)$ sur le spectre de σ_1 dans $\ell_{Rad}^1(\mathbb{F}_d)$. Ce dernier est donc égal à $\{\gamma(z); 0 \leq \Re z \leq 1\}$. C'est le domaine elliptique

$$\{x + iy; x^2 + (\frac{q+1}{q-1})^2 y^2 \leq 1\}.$$

Puisque $\sigma_1 = \sigma_1^*$, un caractère φ_z de $\ell_{Rad}^1(\mathbb{F}_d)$ est auto-adjoint si et seulement si $\gamma(z) = \varphi_z(\sigma_1) \in \mathbb{R}$. Pour $z = x + iy$, cette condition est réalisée si et seulement si l'une des deux conditions suivantes l'est : soit $x = 1/2$ avec y arbitraire, soit $-1 \leq x \leq 1$ avec $y = \frac{k\pi}{\log q}$, $k \in \mathbb{Z}$. Posons $\zeta = \frac{\pi}{\log q}$. Compte tenu des périodicités et des symétries de la fonction γ , en limitant z au sous-ensemble

$$\{1/2 + it; t \in [0, \zeta]\} \cup \{t + ij\zeta; t \in [0, 1/2], j = 0, 1\}$$

de \mathbb{C} , on trouve tous les caractères auto-adjoints de $\ell_{Rad}^1(\mathbb{F}_d)$, sans répétition. Suivant les notations de l'appendice B, $\text{Sp}_h(\ell_{Rad}^1(\mathbb{F}_d))$ désigne leur ensemble. Pour simplifier, nous le noterons $\text{Sp}_{\mathbb{R}}(\sigma_1)$. L'application $z \mapsto \varphi_z(\sigma_1) = \gamma(z)$ l'identifie à l'intervalle $[-1, 1]$. Plus précisément, γ induit une bijection de l'intervalle $\{1/2 + it; t \in [0, \zeta]\}$ sur l'intervalle

$$[-2q^{1/2}(q+1)^{-1}, 2q^{1/2}(q+1)^{-1}],$$

une bijection de $[0, 1/2]$ sur

$$[2q^{1/2}(q+1)^{-1}, 1],$$

et une bijection de $\{t + i\zeta; t \in [0, 1/2]\}$ sur

$$[-1, -2q^{1/2}(q+1)^{-1}].$$

Ainsi $\text{Sp}_{\mathbb{R}}(\sigma_1)$ s'identifie à la partie réelle du spectre de σ_1 dans $\ell_{Rad}^1(\mathbb{F}_d)$.

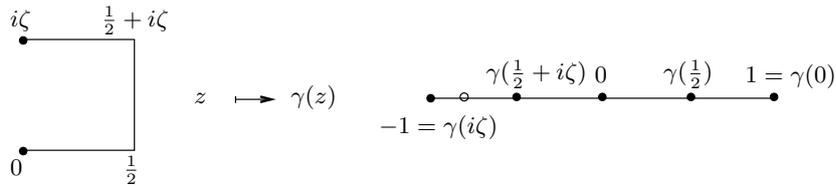


FIGURE 7. $\text{Sp}_{\mathbb{R}}(\sigma_1)$

Enfin, remarquons pour la suite que φ_0 est le caractère trivial associant 1 à tous les σ_n et que $\varphi_{i\zeta}$ est le caractère alterné $\epsilon : \sigma_n \mapsto (-1)^n$.

Remarque 6.2.7. — Si f est une fonction à valeurs complexes définie sur \mathbb{F}_d , posons $Lf = \sigma_1 * f - f$, c'est-à-dire

$$Lf(x) = \frac{1}{q+1} \sum_{w \in S} (f(wx) - f(x)).$$

L'opérateur L est appelé le *laplacien combinatoire* sur \mathbb{F}_d . Si f est une fonction radiale sur \mathbb{F}_d , on notera $f(n)$ sa valeur commune sur les mots de longueur n . Alors Lf est aussi une fonction radiale et

$$Lf(n) = \frac{1}{q+1}f(n-1) + \frac{q}{q+1}f(n+1) - f(n).$$

Identifions tout caractère φ sur $\mathbb{C}_{Rad}[\mathbb{F}_d]$ à la fonction radiale valant $\varphi(\sigma_n)$ sur les mots de longueur n . L'équation (6.2.2) exprime alors que les caractères de $\mathbb{C}_{Rad}[\mathbb{F}_d]$ coïncident avec les fonctions radiales sur \mathbb{F}_d qui sont valeurs propres du laplacien et valent 1 en e . Ce sont les fonctions sphériques sur \mathbb{F}_d . Les caractères bornés à valeur réelle en σ_1 sont les fonctions sphériques de type positif. Les représentations de \mathbb{F}_d qu'elles définissent sont irréductibles (voir [FTP83, Théorème 3.9, Chap. 5]). La série principale est formée des représentations associées aux φ_z avec $z \in \{1/2 + it; t \in [0, \zeta]\}$; la série complémentaire est formée de celles qui sont associées aux φ_z avec $z \in \{t + ij\zeta; t \in]0, 1/2[, j = 0, 1\}$.

6.2.3. Estimation des caractères. — Le lemme ci-dessous montre que les caractères auto-adjoints, sauf le caractère trivial (défini par φ_0) et le caractère alterné (défini par $\varphi_{i\zeta}$), admettent une majoration uniforme par une quantité décroissant exponentiellement vers 0. Ce fait va intervenir de façon cruciale dans la démonstration des théorèmes ergodiques par la méthode spectrale.

Lemme 6.2.8. — *Il existe une constante positive C_q telle que pour tout $z \in \text{Sp}_{\mathbb{R}}(\sigma_1)$ et tout entier n on ait*

$$(6.2.6) \quad |\varphi_z(\sigma_n)| \leq C_q(n+1)q^{-ns},$$

où $s = \Re z$.

Démonstration. — Si $z = \frac{1}{2} + ij\zeta$ avec $j = 0$ ou 1 , on a

$$|\varphi_z(\sigma_n)| = \left(1 + n \frac{q-1}{q+1}\right) q^{-n/2} \leq (n+1)q^{-n/2}.$$

Si $z = \frac{1}{2} + it$, $t \in]0, \zeta[$, on écrit

$$\varphi_z(\sigma_n) = 2\Re(c(z)q^{-nz}) = q^{-n/2} \left(\cos nt' + \frac{(q-1) \cos t' \sin nt'}{(q+1) \sin t'} \right),$$

avec $t' = t \log q$. En utilisant l'égalité $\left| \frac{\sin nt'}{\sin t'} \right| \leq n$, on obtient

$$|\varphi_z(\sigma_n)| \leq (n+1)q^{-n/2}.$$

Supposons maintenant que $z = s + ij\zeta$ avec $0 \leq s < 1/2$ et $j = 0$ ou 1 . En posant $s' = (1/2 - s) \log q$, on a

$$c(s + ij\zeta) = \frac{q^{1-s} - q^{s-1}}{(q+1)(q^{-s} - q^{s-1})} = \frac{qe^{s'} - e^{-s'}}{2(q+1)\text{sh}s'},$$

$$c(1-s - ij\zeta) = \frac{q^s - q^{-s}}{(q+1)(q^{s-1} - q^{-s})} = -\frac{qe^{-s'} - e^{s'}}{2(q+1)\text{sh}s'},$$

d'où

$$\begin{aligned} \varphi_z(\sigma_n) &= \frac{(-1)^{jn} q^{-n/2}}{2(q+1)\text{sh}s'} \left[(qe^{s'} - e^{-s'})e^{ns'} - (qe^{-s'} - e^{s'})e^{-ns'} \right] \\ &= \frac{(-1)^{jn} q^{-n/2}}{q+1} \left[\frac{q \text{sh}((n+1)s')}{\text{sh}s'} - \frac{\text{sh}((n-1)s')}{\text{sh}s'} \right]. \end{aligned}$$

Pour conclure, on utilise les inégalités

$$\frac{\text{sh}(n \pm 1)s'}{\text{sh}s'} \leq (n \pm 1)e^{(n \pm 1)s'}.$$

□

6.2.4. Démonstration du théorème 6.2.1. — Soit π une représentation unitaire du groupe \mathbb{F}_d dans un espace hilbertien H . Commençons par une observation, qui s'adapte aussi au cas des théorèmes ergodiques ponctuels.

Remarque 6.2.9. — Pour démontrer le théorème 6.2.1, il suffit en fait de s'assurer que la suite $(\pi(\sigma_{2n}))$ converge fortement vers $E_1 + E_{-1}$.

En effet, on déduit alors de la formule (6.2.1) que, pour tout $v = \pi(\sigma_1)x$ appartenant à l'image $\text{Im } \pi(\sigma_1)$ de $\pi(\sigma_1)$, on a

$$\lim_{n \rightarrow \infty} \pi(\sigma_{2n+1})v = E_1x + E_{-1}x.$$

Comme $E_1v = E_1\pi(\sigma_1)x = \pi(\sigma_1)E_1x = E_1x$ et, de même, $E_{-1}v = -E_{-1}x$, on voit que $\lim_{n \rightarrow \infty} \pi(\sigma_{2n+1})v = E_1v - E_{-1}v$.

Cette égalité est valable aussi pour tout v dans l'orthogonal $\text{Ker } \pi(\sigma_1)$ de $\text{Im } \pi(\sigma_1)$. En effet, on déduit par récurrence de la formule (6.2.1) que $\text{Ker } \pi(\sigma_1) \subset \text{Ker } \pi(\sigma_{2n+1})$ pour tout $n \geq 0$, d'où $\lim_{n \rightarrow +\infty} \pi(\sigma_{2n+1})v = 0$, ce qui permet de conclure car on a $E_1v = 0 = E_{-1}v$. La convergence forte de $(\pi(\sigma_{2n+1}))$ est ainsi établie.

L'assertion (ii) du théorème 6.2.1 se déduit immédiatement de (i). Pour démontrer (iii), on remarque d'abord que, d'après (i), on a $\lim_{n \rightarrow +\infty} \pi(\beta_n)v = 0$ si v est orthogonal à $E_1(H) \oplus E_{-1}(H)$. Si $v \in E_1(H)$, on a bien sûr $\lim_{n \rightarrow +\infty} \pi(\beta_n)v = v$. Il reste à considérer le cas où $v \in E_{-1}(H)$. En observant que

$$\beta_p = \frac{1}{|B_p|} \sum_{k=0}^p |S_k| \sigma_k,$$

pour tout entier p , on obtient dans ce cas

$$\pi(\beta_p)v = \left(\frac{1}{|B_p|} \sum_{k=0}^p |S_k|(-1)^k \right) v.$$

Comme $\sum_{k=0}^p |S_k|(-1)^k = (-q)^p$ et $|B_p| = 1 + (q+1)\frac{q^p-1}{q-1}$, on voit que

$$\lim_{n \rightarrow +\infty} \pi(\beta_{2n})v = \frac{q-1}{q+1}v \quad \text{et} \quad \lim_{n \rightarrow +\infty} \pi(\beta_{2n+1})v = -\frac{q-1}{q+1}v.$$

Pour démontrer le théorème 6.2.1 il reste donc à étudier le comportement de la suite $(\pi(\sigma_{2n}))$ et pour cela nous utilisons la théorie spectrale. Notons B la fermeture de $\pi(\ell_{Rad}^1(\mathbb{F}_d))$ dans $\mathcal{B}(H)$. C'est une C^* -algèbre commutative, et par conséquent tous ses caractères sont auto-adjoints. Notons $\text{Sp}(B)$ l'ensemble de ces caractères. L'application $\chi \mapsto \chi \circ \pi$ est une injection de $\text{Sp}(B)$ dans l'ensemble des caractères auto-adjoints de $\ell_{Rad}^1(\mathbb{F}_d)$. On peut aussi la voir comme une injection du spectre $\text{Sp}(\pi(\sigma_1)) = \{\chi \circ \pi(\sigma_1); \chi \in \text{Sp}(B)\}$ de l'opérateur auto-adjoint $\pi(\sigma_1) \in B$, à valeurs dans la partie réelle $\text{Sp}_{\mathbb{R}}(\sigma_1)$ du spectre de $\sigma_1 \in \ell_{Rad}^1(\mathbb{F}_d)$. En effet, comme la C^* -algèbre B est engendrée par l'identité et $\pi(\sigma_1)$, l'application $\chi \mapsto \chi \circ \pi$ est un homéomorphisme de $\text{Sp}(B)$ sur $\text{Sp}(\pi(\sigma_1))$, ce qui permet d'identifier ces deux espaces.

Soient v un vecteur de H et m_v la mesure spectrale associée. D'après le théorème spectral rappelé dans l'appendice B, nous avons

$$(6.2.7) \quad \|\pi(\sigma_{2n})v\|^2 = \int_{\text{Sp}_{\mathbb{R}}(\sigma_1)} |\varphi_z(\sigma_{2n})|^2 dm_v(z).$$

L'estimation (6.2.6) entraîne que $\lim_{n \rightarrow +\infty} |\varphi_z(\sigma_{2n})| = 0$ pour tout z dans $\text{Sp}_{\mathbb{R}}(\sigma_1)$ sauf aux extrémités $z = 0$ et $z = i\zeta$, pour lesquelles, comme nous l'avons vu, $|\varphi_z(\sigma_{2n})| = 1$ pour tout n . Le théorème de convergence dominée de Lebesgue donne alors

$$\lim_{n \rightarrow +\infty} \|\pi(\sigma_{2n})v\|^2 = m_v(\{0, i\zeta\}).$$

De plus, on a $m_v(\{0, i\zeta\}) = \|E_1v + E_{-1}v\|^2$ d'après le lemme B.2.3. Par conséquent, si v est orthogonal à $E_1(H) \oplus E_{-1}(H)$, nous voyons que $\lim_{n \rightarrow +\infty} \|\pi(\sigma_{2n})v\| = 0$. Pour conclure que $\lim_{n \rightarrow +\infty} \pi(\sigma_{2n}) = E_1 + E_{-1}$, il suffit de remarquer que $\pi(\sigma_{2n})v = v$ pour tout $v \in E_1(H) \oplus E_{-1}(H)$. \square

Remarque 6.2.10. — D'après l'égalité (6.2.7), la démonstration du théorème 6.2.1 nécessite seulement de savoir que la suite des fonctions $z \mapsto |\varphi_z(\sigma_{2n})|$, vues comme fonctions sur $[-1, 1]$, est uniformément bornée sur cet intervalle, tend simplement vers 0 à l'intérieur et vers 1 aux extrémités. Cela se déduit facilement de la résolution de la relation de récurrence (6.2.3) avec $\lambda \in [-1, 1]$, ainsi qu'Arnold et Krylov l'ont observé dans [AK63].

6.3. Inégalités maximales et théorèmes ergodiques ponctuels dans L^2

Dans cette section, nous nous donnons un espace mesuré (X, \mathcal{B}, m) où m est une mesure de probabilité. Nous supposons que \mathbb{F}_d agit sur (X, \mathcal{B}, m) en préservant m et nous considérons la représentation unitaire associée π de \mathbb{F}_d dans $L^2(X, m)$. Toutes les mesures considérées ici sont symétriques (*i.e.* $\mu = \check{\mu}$). Rappelons que pour une telle mesure $\mu \in \ell^1(\mathbb{F}_d)$, nous avons $\pi(\mu)f = \mu \cdot f$ pour tout $f \in L^2(X, m)$.

Nous allons établir le théorème suivant, dû à Nevo.

Théorème 6.3.1 ([Nev94a]). — Soit $f \in L^2(X, m)$.

- (i) La suite $(\pi(\sigma_{2n})f)$ converge presque partout vers l'espérance conditionnelle $\mathbb{E}(f|\mathcal{I}_2)$, où \mathcal{I}_2 désigne la tribu des invariants sous l'action de $\mathbb{F}_d^{(2)}$.
- (ii) Les suites $(\pi(\nu_n)f)$ et $(\pi(\sigma'_n)f)$ convergent presque partout vers $E_1f = \mathbb{E}(f|\mathcal{I})$, où \mathcal{I} désigne la tribu des invariants sous l'action de \mathbb{F}_d .
- (iii) La suite $(\pi(\beta_{2n})f)$ converge presque partout vers $E_1f + \frac{q-1}{q+1}E_{-1}f$.

Remarque 6.3.2. — On en déduit que la suite $(\pi(\sigma_{2n+1})f)$ converge presque partout vers $E_1f - E_{-1}f$ et que la suite $(\pi(\beta_{2n+1})f)$ converge presque partout vers $E_1f - \frac{q-1}{q+1}E_{-1}f$. On voit en particulier que la suite $(\pi(\sigma_n)f)$ converge presque partout si et seulement si $E_{-1}f = 0$.

Suivant la démarche classique, nous allons établir des inégalités maximales (voir le théorème 6.3.8). Indiquons d'abord comment nous les utilisons pour démontrer le théorème ci-dessus.

Pour toute partie borélienne U de \mathbb{R} , nous désignons par $E(U)$ le projecteur spectral de $\pi(\sigma_1)$ associé. Étant donné $\varepsilon > 0$, nous posons $U_\varepsilon = \{\gamma(z); \varepsilon \leq \Re z \leq 1/2\}$, et $H_\varepsilon = E(U_\varepsilon)H$ où $H = L^2(X, m)$. Le lemme suivant nous permettra de conclure facilement.

Lemme 6.3.3. — Soit $f \in H_\varepsilon$. On a $\sum_{n=0}^{+\infty} \|\pi(\sigma_n)f\|_2^2 < +\infty$, et par conséquent $\lim_{n \rightarrow +\infty} \pi(\sigma_n)f(x) = 0$ pour presque tout $x \in X$.

Démonstration. — En notant m_f la mesure spectrale associée à f et à l'opérateur auto-adjoint $\pi(\sigma_1)$, on a

$$\sum_{n=0}^{+\infty} \|\pi(\sigma_n)f\|_2^2 = \sum_{n=0}^{+\infty} \int_{\text{Sp}_{\mathbb{R}}(\sigma_1)} |\varphi_z(\sigma_n)|^2 dm_f(z).$$

Si $\varepsilon \leq \Re z \leq 1/2$, l'estimation (6.2.6) donne $|\varphi_z(\sigma_n)| \leq C_q(n+1)q^{-n\varepsilon}$, d'où

$$\sum_{n=0}^{+\infty} \|\pi(\sigma_n)f\|_2^2 \leq C_q^2 \sum_{n=0}^{+\infty} (n+1)^2 q^{-2n\varepsilon} < +\infty.$$

Il en résulte que $\sum_{n=0}^{+\infty} |\pi(\sigma_n)f(x)|^2$ est presque partout convergente. On a donc $\lim_{n \rightarrow +\infty} \pi(\sigma_n)f(x) = 0$ pour presque tout $x \in X$. \square

Démonstration du théorème 6.3.1 à partir du théorème maximal 6.3.8. —

Comme on l'a déjà vu dans la remarque 6.2.9 pour les théorèmes ergodiques en moyenne, il suffit de considérer le cas de la suite $(\pi(\sigma_{2n}))$. Rappelons (voir l'appendice B) que

$$H = E_1(H) \oplus E_{-1}(H) \oplus \overline{\cup_{\varepsilon>0} H_\varepsilon}.$$

Pour $f \in E_1(H) \oplus E_{-1}(H)$, on a $f = \pi(\sigma_{2n})f$. Il résulte alors du lemme 6.3.3 que pour $f \in E_1(H) + E_{-1}(H) + \cup_{\varepsilon>0} H_\varepsilon$, la suite $(\pi(\sigma_{2n})f)$ converge presque partout. D'après le théorème 1.2.1 et le théorème maximal 6.3.8, elle converge presque partout pour tout $f \in L^2(X, m)$. Enfin, le théorème 6.2.1 (i) et la remarque 6.2.2 entraînent que $\mathbb{E}(f|\mathcal{I}_2)$ est la limite de cette suite. \square

Nous allons en fait donner deux façons différentes d'établir les inégalités maximales, ce qui sera l'occasion de présenter diverses techniques classiques. La première méthode concerne uniquement le cas L^2 . Elle utilise la sous-additivité de la suite (ν_n) et l'inégalité maximale sous-additive générale. Elle présente l'avantage de s'étendre à des groupes discrets autres que les groupes libres, notamment les groupes hyperboliques.

Nous allons ensuite établir les théorèmes maximaux L^p à l'aide d'autres outils, comme l'inégalité maximale de Hopf-Dunford-Schwartz et l'interpolation complexe pour une famille analytique d'opérateurs.

6.3.1. Inégalité maximale sous-additive. — Soit (T_n) une suite d'opérateurs de $L^2(X, m)$. Si $f \in L^2(X, m)$, on pose $f^*(x) = \sup_{n \geq 0} |T_n f(x)|$.

Définition 6.3.4. — Une suite sous-additive d'opérateurs de Markov auto-adjoints est une suite (T_n) d'opérateurs sur $L^2(X, m)$ telle que

- (i) $T_n = T_n^*$ et $\|T_n\| \leq 1$ pour tout n ;
- (ii) $T_n \mathbf{1} = \mathbf{1}$ et $T_n f \geq 0$ si $f \geq 0$;
- (iii) il existe une constante $C > 0$ et un entier $k \geq 1$ tels que pour toute fonction $f \geq 0$ appartenant à $L^2(X, m)$ et tous $n, p \in \mathbb{N}$, on ait

$$T_n T_p f(x) \leq C(T_{kn} f(x) + T_{kp} f(x)) \quad \text{presque partout.}$$

Exemple 6.3.5. — Notons que d'autres suites sous-additives d'opérateurs de Markov auto-adjoints ont déjà été rencontrées dans le chapitre 3. En effet, soit G un groupe localement compact unimodulaire agissant sur (X, m) en préservant la mesure de probabilité m . Considérons une suite croissante (F_n) de parties mesurables symétriques dans G avec $0 < |F_n| < +\infty$. On suppose en outre que, pour tous n, p , on a

$$F_n F_p \subset F_{2 \max(n,p)} \quad \text{et} \quad |F_{2n}| \leq c |F_n|.$$

On note T_n l'opérateur de $L^2(X, m)$ défini par

$$T_n(f)(x) = \frac{1}{|F_n|} \int_{F_n} f(gx) \, dg.$$

On a alors, pour tous n, p et pour toute fonction positive f dans $L^2(X, m)$,

$$T_n T_p f(x) \leq c(T_{2n} f(x) + T_{2p} f(x)) \quad \text{presque partout.}$$

Théorème 6.3.6 ([Bar93],[Nev94a]). — Soit (T_n) une suite sous-additive d'opérateurs de Markov auto-adjoints. Alors pour tout $f \in L^2(X, m)$, on a $\|f^*\|_2 \leq 2C\|f\|_2$, où C est la constante qui intervient dans la définition 6.3.4.

Démonstration. — Il suffit de considérer le cas où f est positive et bornée. Pour tout entier N , posons $f_N^* = \max_{0 \leq n \leq N} T_n f$ et notons $M(f, N)$ la fonction mesurable qui prend en x la plus petite valeur n telle que le maximum $\max_{0 \leq n \leq N} T_n f(x)$ est atteint. Alors f_N^* s'écrit aussi $T_{M(f, N)} f$ où, pour toute fonction mesurable M à valeurs dans $\{0, 1, \dots, N\}$ et pour $g \in L^2(X, m)$, on pose

$$T_M g = \sum_{n=0}^N \mathbf{1}_{\{M=n\}} T_n g.$$

Ici, $\mathbf{1}_{\{M=n\}}$ est la fonction indicatrice de l'ensemble $\{x \in X; M(x) = n\}$. Si ϵ_n désigne le projecteur $g \mapsto \mathbf{1}_{\{M=n\}} g$ défini sur $L^2(X, m)$, on voit que $T_M = \sum_{n=0}^N \epsilon_n \circ T_n$ et que l'adjoint de T_M s'écrit $T_M^* = \sum_{n=0}^N T_n \circ \epsilon_n$.

La principale étape de la démonstration consiste à établir l'inégalité

$$(6.3.1) \quad \|T_M^* g\|_2^2 \leq 2C \langle g^*, g \rangle,$$

pour toute fonction mesurable positive bornée g définie sur X . Le fait que la constante ne dépend pas du choix de M est crucial. Cette inégalité s'obtient de la façon suivante, en utilisant la sous-additivité de la suite (T_n) :

$$\begin{aligned} \|T_M^* g\|_2^2 &= \langle T_M T_M^* g, g \rangle = \sum_{n=0}^N \sum_{p=0}^N \langle \epsilon_n \circ T_n \circ T_p \circ \epsilon_p g, g \rangle \\ &\leq C \sum_{n=0}^N \sum_{p=0}^N \langle \epsilon_n \circ (T_{kn} + T_{kp}) \circ \epsilon_p g, g \rangle \\ &\leq C \sum_{n=0}^N \langle \epsilon_n \circ T_{kn} (\sum_{p=0}^N \epsilon_p g), g \rangle + C \sum_{p=0}^N \langle T_{kp} \circ \epsilon_p g, \sum_{n=0}^N \epsilon_n g \rangle \\ &\leq C \sum_{n=0}^N \langle T_{kn} g, \epsilon_n g \rangle + C \sum_{p=0}^N \langle \epsilon_p g, T_{kp} g \rangle \leq 2C \langle g^*, g \rangle. \end{aligned}$$

On déduit de (6.3.1) que

$$\langle g, f_N^* \rangle^2 = \langle g, T_{M(f, N)} f \rangle^2 \leq \|T_{M(f, N)}^* g\|_2^2 \|f\|_2^2 \leq 2C \langle g^*, g \rangle \|f\|_2^2,$$

d'où

$$\langle g, f^* \rangle^2 \leq 2C \langle g^*, g \rangle \|f\|_2^2,$$

en prenant la borne supérieure dans le membre de gauche.

On a en particulier $\langle f, f^* \rangle \leq 2C\|f\|_2^2$ et donc, grâce à (6.3.1),

$$\|T_M^* f\|_2 \leq 2C\|f\|_2,$$

pour toute fonction positive bornée f .

Cette inégalité, appliquée à f^* au lieu de f et avec $M(f, N)$, donne

$$\begin{aligned} \langle f^*, f_N^* \rangle^2 &= \langle f^*, T_{M(f, N)} f \rangle^2 = \langle T_{M(f, N)}^* f^*, f \rangle^2 \\ &\leq \|T_{M(f, N)}^* f^*\|_2^2 \|f\|_2^2 \leq 4C^2 \|f^*\|_2^2 \|f\|_2^2 \end{aligned}$$

et donc $\|f^*\|_2 \leq 2C\|f\|_2$. \square

6.3.2. Sous-additivité de la suite (ν_n) . — Dans le lemme suivant, l'espace des fonctions à valeurs réelles définies sur \mathbb{F}_d est muni de l'ordre usuel.

Lemme 6.3.7. — *On a*

- (i) $\sigma_t * \sigma_s \leq \sum_{p=0}^s q^{-p} \sigma_{t+s-2p}$ pour s, t entiers avec $s \leq t$;
- (ii) $\sum_{i=0}^n \sigma_{m+i} * \sigma_s \leq \frac{1}{1 - q^{-1/2}} \sum_{k=0}^{n+2m} \sigma_k$ pour n, m et s entiers avec $s \leq m$;
- (iii) $\nu_n * \nu_m \leq \frac{4}{1 - q^{-1/2}} (\nu_{2n} + \nu_{2m})$ quels que soient les entiers n et m .

Démonstration. — (i) Pour tout mot w , on a

$$\sigma_t * \sigma_s(w) = \sum_{uv=w} \sigma_t(u) \sigma_s(v).$$

Soient u, v des mots réduits tels que $|u| = t$ et $|v| = s$, et soit p le nombre de simplifications intervenant dans le mot uv . On a $0 \leq p \leq s$ et $|uv| = t + s - 2p$. La fonction radiale $\sigma_t * \sigma_s$ est donc une combinaison de termes de la forme σ_{t+s-2p} avec $0 \leq p \leq s$. Estimons alors le coefficient de σ_{t+s-2p} . Le nombre de façons d'écrire un mot de longueur $t + s - 2p$ sous la forme uv avec $|u| = t$ et $|v| = s$ est majoré par le nombre de mots de longueur p . Il en résulte que le coefficient de σ_{t+s-2p} est majoré par $\frac{|S_{t+s-2p}| |S_p|}{|S_t| |S_s|} \leq q^{-p}$.

(ii) Fixons un entier l avec $0 \leq l \leq n + 2m$ et estimons le coefficient de σ_l dans le développement de $\sum_{i=0}^n \sigma_{m+i} * \sigma_s$. D'après (i), si σ_l a un coefficient non nul dans $\sigma_{m+i} * \sigma_s$, alors $l = m + s + i - 2p$ avec $0 \leq p \leq s$, et le coefficient est majoré par q^{-p} . En posant $j = s + m + i - l$ on voit donc que le coefficient de σ_l dans $\sum_{i=0}^n \sigma_{m+i} * \sigma_s$ est majoré par $\sum_{j=0}^{2s} q^{-j/2}$ et donc par $(1 - q^{-1/2})^{-1}$.

(iii) Supposons $n \geq m$. Écrivons

$$\nu_n * \nu_m = \frac{1}{(n+1)(m+1)} \left[\left(\sum_{k=m}^n \sigma_k \right) * \left(\sum_{k=0}^m \sigma_k \right) + \left(\sum_{k=0}^{m-1} \sigma_k \right) * \left(\sum_{k=0}^m \sigma_k \right) \right].$$

Grâce à (ii), on obtient

$$\left(\sum_{k=m}^n \sigma_k\right) * \left(\sum_{k=0}^m \sigma_k\right) = \sum_{k=0}^m \sum_{i=0}^{n-m} \sigma_{m+i} * \sigma_k \leq \frac{m+1}{1-q^{-1/2}} \sum_{k=0}^{2n} \sigma_k.$$

De même, on trouve

$$\sum_{0 \leq s, t \leq m} \sigma_t * \sigma_s \leq 2 \sum_{0 \leq s \leq t \leq m} \sigma_t * \sigma_s \leq \frac{2(m+1)}{1-q^{-1/2}} \sum_{k=0}^{2m} \sigma_k,$$

d'où, en regroupant les deux estimations,

$$\begin{aligned} \nu_n * \nu_m &\leq \frac{1}{(1-q^{-1/2})(n+1)} \left(\sum_{k=0}^{2n} \sigma_k + 2 \sum_{k=0}^{2m} \sigma_k \right) \\ &\leq \frac{4}{1-q^{-1/2}} (\nu_{2n} + \nu_{2m}). \end{aligned}$$

□

6.3.3. Inégalités maximales L^2 . — Soit $f \in L^2(X, m)$. Nous poserons comme d'habitude

$$\begin{aligned} f_\sigma^*(x) &= \sup_{n \geq 0} |\sigma_n \cdot f(x)|, & f_\nu^*(x) &= \sup_{n \geq 0} |\nu_n \cdot f(x)|, \\ f_{\sigma'}^*(x) &= \sup_{n \geq 0} |\sigma'_n \cdot f(x)|, & f_\beta^*(x) &= \sup_{n \geq 0} |\beta_n \cdot f(x)|. \end{aligned}$$

Théorème 6.3.8 ([Nev94a]). — Les opérateurs maximaux associés aux suites (ν_n) , (σ'_n) , (σ_n) et (β_n) sont bornés dans $L^2(X, m)$.

Démonstration. — L'assertion relative à la suite (ν_n) résulte du lemme 6.3.7 et du théorème 6.3.6.

Pour la suite (σ'_n) , on utilise une méthode due à Stein [Ste61]. C'est une version discrète de la méthode de Littlewood-Paley-Stein déjà rencontrée dans la section 5.1.2. On compare la fonction maximale relative à la suite (σ'_n) et la fonction maximale relative aux moyennes de Cesàro $\nu'_n = \frac{1}{n+1} \sum_{k=0}^n \sigma'_k$. La transformation d'Abel donne

$$\sigma'_n - \frac{1}{n+1} \sum_{k=0}^n \sigma'_k = \frac{1}{n+1} \sum_{k=1}^n k(\sigma'_k - \sigma'_{k-1}).$$

Grâce à l'inégalité de Cauchy-Schwarz, on obtient pour $f \in L^2(X, m)$,

$$\begin{aligned} |\sigma'_n \cdot f(x) - \nu'_n \cdot f(x)|^2 &\leq \frac{1}{(n+1)^2} \left(\sum_{k=1}^n k \right) \left(\sum_{k=1}^n k |\sigma'_k - \sigma'_{k-1} \cdot f(x)|^2 \right) \\ &\leq \sum_{k=1}^n k |\sigma'_k - \sigma'_{k-1} \cdot f(x)|^2. \end{aligned}$$

En prenant les racines carrées, on en déduit que

$$f_{\sigma'}^*(x) \leq \sup_{n \geq 0} |\nu'_n \cdot f(x)| + R(f)(x)$$

où $(R(f)(x))^2 = \sum_{k=1}^{+\infty} k |(\sigma'_k - \sigma'_{k-1}) \cdot f(x)|^2$.

Puisque $\nu'_n \leq 2\nu_{n+1}$, on voit que

$$\|f_{\sigma'}^*\|_2 \leq 2\|f_{\nu}^*\|_2 + \|R(f)\|_2.$$

En faisant intervenir la mesure spectrale m_f associée à f et $\pi(\sigma_1)$, on trouve

$$\begin{aligned} \|R(f)\|_2^2 &= \sum_{k=1}^{+\infty} k \|(\sigma'_k - \sigma'_{k-1}) \cdot f\|_2^2 \\ &= \int_{\text{Sp}_{\mathbb{R}}(\sigma_1)} \sum_{k=1}^{+\infty} k |\varphi_z(\sigma'_k) - \varphi_z(\sigma'_{k-1})|^2 dm_f(z). \end{aligned}$$

Pour conclure, il suffit de montrer que la fonction ψ définie sur $\text{Sp}_{\mathbb{R}}(\sigma_1)$ par

$$\psi(z) = \sum_{k=1}^{+\infty} k |\varphi_z(\sigma'_k) - \varphi_z(\sigma'_{k-1})|^2$$

est bornée. En effet, on aura alors $\|R(f)\|_2 \leq \|\psi\|_{\infty}^{1/2} \|f\|_2$, d'où

$$\|f_{\sigma'}^*\|_2 \leq \|f_{\nu}^*\|_2 + \|\psi\|_{\infty}^{1/2} \|f\|_2.$$

Le fait que l'opérateur maximal $f \mapsto f_{\nu}^*$ est borné dans $L^2(X, m)$ implique alors la même propriété pour $f \mapsto f_{\sigma'}^*$.

Pour démontrer que ψ est bornée sur $\text{Sp}_{\mathbb{R}}(\sigma_1)$ nous observons d'abord que

$$(6.3.2) \quad \psi(z) = \frac{1}{4} \sum_{k=1}^{+\infty} k |\varphi_z(\sigma_{k+1}) - \varphi_z(\sigma_{k-1})|^2.$$

Aux extrémités $z = 0$ et $z = i\zeta$, on voit que $\psi(z) = 0$.

Si $1/4 \leq \Re z = s \leq 1/2$, on a d'après (6.3.2) et (6.2.6),

$$\begin{aligned} \psi(z) &\leq \frac{1}{4} \sum_{k=1}^{+\infty} k (|\varphi_z(\sigma_{k+1})| + |\varphi_z(\sigma_{k-1})|)^2 \\ &\leq C_q^2 \sum_{k=1}^{+\infty} (k+1)^3 q^{-(k-1)/2} < +\infty. \end{aligned}$$

Enfin, supposons que $z = s + ij\zeta$, avec $s \in]0, 1/4[$ et $j = 0$ ou 1 . En utilisant l'égalité (6.2.4),

$$\varphi_{s+ij\zeta}(\sigma_n) = (-1)^{jn} \left(c(s + ij\zeta) q^{-ns} + c(1 - s - ij\zeta) q^{-n(1-s)} \right),$$

on obtient

$$\begin{aligned} \psi(z) &= \frac{1}{4} \sum_{k=1}^{+\infty} k \left| c(s + ij\zeta)(q^{-s} - q^s)q^{-ks} + c(1 - s - ij\zeta)(q^{-(1-s)} - q^{(1-s)})q^{-k(1-s)} \right|^2 \\ &\leq \frac{a(q)^2}{4} \sum_{k=1}^{+\infty} k \left| (q^s - q^{-s})q^{-ks} + (q+1)q^{-k(1-s)} \right|^2, \end{aligned}$$

où

$$a(q) = \max\{|c(s + ij\zeta)|, |c(1 - s - ij\zeta)|; s \in [0, 1/4], j = 0, 1\}.$$

Comme $s \in [0, 1/4]$, on majore $q^{-k(1-s)}$ par $q^{-3k/4}$, d'où

$$\psi(z)^{1/2} \leq \frac{a(q)}{2} \left[\left(\sum_{k=1}^{+\infty} k(q^s - q^{-s})^2 q^{-2ks} \right)^{1/2} + (q+1) \left(\sum_{k=1}^{+\infty} kq^{-3k/2} \right)^{1/2} \right]$$

grâce à l'inégalité triangulaire ℓ^2 . En remarquant que $(q^s - q^{-s})^2 \sum_{k=1}^{+\infty} kq^{-2ks} = 1$, on conclut que ψ est bornée sur $\{s + ij\zeta; s \in [0, 1/4], j = 0, 1\}$ et donc sur $\text{Sp}_{\mathbb{R}}(\sigma_1)$ tout entier.

Finalement, on déduit que $f \mapsto f_{\sigma}^*$ est borné dans $L^2(X, m)$, à partir de l'inégalité $\sigma_n \leq 2\sigma'_n$ et de ce qui précède. Pour traiter le cas de $f \mapsto f_{\beta}^*$ on remarque que pour tout $f \in L^2(X, m)$, on a

$$|\beta_n \cdot f(x)| \leq \sup_{0 \leq k \leq n} |\sigma_k \cdot f(x)|.$$

□

6.4. Inégalités maximales et théorèmes ergodiques dans L^p .

Notre but est de démontrer les deux théorèmes suivants :

Théorème 6.4.1 ([NS94]). — *L'opérateur maximal $f \mapsto f_{\nu}^*$ est de type faible $(1, 1)$. Autrement dit, il existe une constante $C > 0$ (dépendant de d) telle que, pour tout $f \in L^1(X, m)$ et tout $\alpha > 0$, on ait*

$$m(\{x \in X; f_{\nu}^*(x) > \alpha\}) \leq \frac{C}{\alpha} \|f\|_1.$$

Par suite, pour tout $p > 1$, cet opérateur maximal est de type fort (p, p) .

Théorème 6.4.2 ([NS94]). — *Pour tout $p > 1$, l'opérateur maximal $f \mapsto f_{\sigma}^*$ est de type fort (p, p) . Autrement dit, il existe une constante $C > 0$ (dépendant de p et de d) telle que, pour tout $f \in L^p(X, m)$, on ait l'inégalité*

$$(6.4.1) \quad \|f_{\sigma}^*\|_p \leq C \|f\|_p.$$

Remarque 6.4.3. — Par interpolation avec le cas évident $p = +\infty$, les résultats ci-dessus se déduisent du théorème 6.3.8 lorsque $p \geq 2$. Toutefois, nous établissons maintenant directement les inégalités maximales pour tout $p > 1$.

L'existence d'une inégalité L^1 faible pour la fonction maximale relative aux sphères est un problème ouvert.

Suivant la méthode usuelle, utilisant les inégalités maximales et les résultats de convergence ponctuelle démontrés dans la section précédente pour les éléments de $L^2(X, m)$ et donc de la partie $L^2(X, m) \cap L^p(X, m)$, dense dans $L^p(X, m)$, on en déduit le théorème ci-dessous.

Théorème 6.4.4 ([NS94]). — Soit $f \in L^p(X, m)$. Si $p \geq 1$, la suite $(\nu_n \cdot f)$ converge presque partout vers $\mathbb{E}(f|\mathcal{I})$. Si $p > 1$, la suite $(\sigma_{2n} \cdot f)$ converge presque partout vers $\mathbb{E}(f|\mathcal{I}_2)$.

Comme dans le cas L^2 , on a aussi la convergence ponctuelle de la suite $(\sigma_{2n+1} \cdot f)$ avec, en général, une limite différente de celle de la suite $(\sigma_{2n} \cdot f)$. Rappelons par ailleurs que les résultats de convergence en moyenne L^2 établis dans le théorème 6.2.1 entraînent la convergence dans L^p pour tout $p \in [1, +\infty[$ (voir la remarque 1.3.5).

Les sections suivantes vont être consacrées aux démonstrations des théorèmes 6.4.1 et 6.4.2. Pour le premier, nous allons utiliser l'inégalité maximale de Hopf-Dunford-Schwartz (voir la proposition 1.3.4), en majorant la somme des $\sigma_k \cdot f$, pour $k \leq n$, par une somme de puissances de l'opérateur de convolution par σ_1 appliqué à f . Le second théorème découlera d'une méthode d'interpolation complexe pour une famille analytique d'opérateurs.

6.4.1. Majoration des moyennes de Cesàro. — L'objet de cette section est de démontrer le lemme suivant :

Lemme 6.4.5. — Il existe une constante C telle que, pour tout entier positif n , on ait l'inégalité

$$\sum_{k \leq n} \sigma_k \leq C \sum_{k \leq 2n} \sigma_1^{*k},$$

où σ_1^{*k} désigne la puissance k -ième de convolution de la mesure σ_1 .

Admettons un instant le lemme, et montrons le théorème 6.4.1.

Démonstration du théorème 6.4.1. — Notons T l'opérateur défini sur $L^1(X, m)$ par $Tf = \sigma_1 \cdot f$. Puisque π est une représentation sur $L^1(X, m)$, la puissance k -ième de T est donnée par une puissance k -ième de convolution :

$$T^k f = \sigma_1^{*k} \cdot f.$$

D'autre part, il résulte du lemme que, si f est une fonction positive,

$$\frac{1}{n+1} \sum_{k \leq n} \sigma_k \cdot f \leq \frac{2C}{2n+1} \sum_{k \leq 2n} T^k f.$$

Le théorème maximal pour les moyennes de Cesàro découle alors directement de l'inégalité maximale faible (1, 1) pour T . \square

Démonstration du lemme 6.4.5. — Notons, comme précédemment, $q = 2d - 1$. Nous allons utiliser l'identité (6.2.1) donnée dans le lemme 6.2.5, qui est valable pour $n \geq 1$,

$$\sigma_1 * \sigma_n = \frac{1}{q+1} \sigma_{n-1} + \frac{q}{q+1} \sigma_{n+1}.$$

Il en découle, par récurrence, que σ_1^{*k} s'écrit sous la forme

$$\sigma_1^{*k} = \sum_{j=0}^k a_k(j) \sigma_j,$$

formule dans laquelle les coefficients $a_k(j)$ sont tous positifs ou nuls et s'obtiennent par récurrence à partir des trois règles suivantes

$$\begin{aligned} a_k(0) &= \frac{1}{q+1} a_{k-1}(1) \\ a_k(1) &= \frac{1}{q+1} a_{k-1}(2) + a_{k-1}(0) \\ a_k(j) &= \frac{1}{q+1} a_{k-1}(j+1) + \frac{q}{q+1} a_{k-1}(j-1), \quad 2 \leq j \leq k, \end{aligned}$$

de la valeur initiale $a_0(0) = 1$ et, pour tout $k \geq 0$, de la condition $a_k(j) = 0$ si $j > k$.

Comme

$$\sum_{k=0}^{2n} \sigma_1^{*k} = \sum_{j=0}^{2n} \left(\sum_{k=j}^{2n} a_k(j) \right) \sigma_j,$$

il suffit de montrer l'existence d'une constante $C > 0$ telle que, pour tout entier positif

n et tout $j \leq n$, on ait $\sum_{k=j}^{2n} a_k(j) \geq C$.

On reconnaît dans la relation de récurrence qui donne la suite a_k à partir de la suite a_{k-1} la relation qui, pour une marche aléatoire de Bernoulli asymétrique sur \mathbb{N} avec réflexion en 0, donne la probabilité d'être au point j à l'instant k . Pour d'autres conditions initiales cette relation permet le calcul d'autres probabilités classiques (cf. [Fel68, XIV.4]). Cela incite à minorer les coefficients $a_k(j)$ par les nombres $b_k(j)$, d'expression connue, définis par récurrence par les règles

$$\begin{aligned} b_k(0) &= 0, \quad \text{pour tout } k \geq 0, \\ b_k(j) &= \frac{1}{q+1} b_{k-1}(j+1) + \frac{q}{q+1} b_{k-1}(j-1), \quad \text{pour } 1 \leq j \leq k, \end{aligned}$$

par la valeur initiale $b_1(1) = q/(q+1)$ et par la condition $b_k(j) = 0$ si $j > k$. On retrouve aisément par récurrence qu'on a $b_k(j) = 0$ si k et j ne sont pas de même parité et, dans le cas contraire,

$$(6.4.2) \quad b_k(j) = \frac{j}{k} \binom{k}{\frac{k+j}{2}} \left(\frac{q}{q+1}\right)^{\frac{k+j}{2}} \left(\frac{1}{q+1}\right)^{\frac{k-j}{2}}.$$

Les nombres $a_k(j)$, pour $0 \leq j \leq k$, sont strictement positifs. Pour achever la démonstration du lemme 6.4.5 il suffit de montrer qu'il existe un indice $j_0 > 0$ et une constante $C > 0$ tels qu'on ait $\sum_{k=j}^{2n} b_k(j) \geq C$ pour tout $j \geq j_0$ et tout $n \geq j$. Posons $\tau = \frac{q+1}{q-1}$. Pour j assez grand et $n \geq j$ on a $[\tau j - \sqrt{j}, \tau j] \subset [j, 2n]$. Il suffit donc d'établir une minoration de la forme

$$\sum_{k \in [\tau j - \sqrt{j}, \tau j]} b_k(j) \geq C.$$

Une telle minoration résulte du lemme suivant :

Lemme 6.4.6. — *Il existe une constante strictement positive c telle que, pour tout couple (j, k) d'entiers de même parité où k appartient à l'intervalle $[\tau j - \sqrt{j}, \tau j]$, on ait l'inégalité*

$$b_k(j) \geq \frac{c}{\sqrt{j}}.$$

Démonstration. — On peut écrire les entiers k de l'intervalle $[\tau j - \sqrt{j}, \tau j]$ sous la forme $k = \tau j(1 - \alpha)$ avec $0 \leq \alpha \leq \frac{1}{\tau\sqrt{j}}$. Quand j tend vers l'infini, on a, uniformément par rapport à un entier k appartenant à $[\tau j - \sqrt{j}, \tau j]$, $k \approx \frac{k+j}{2} \approx \frac{k-j}{2} \approx j$. La formule de Stirling ($x! \sim x^x e^{-x} \sqrt{2\pi x}$) donne alors, pour j tendant vers l'infini, k dans $[\tau j - \sqrt{j}, \tau j]$ et de même parité que j ,

$$b_k(j) \sim \frac{c(\tau)}{\sqrt{j}} \left(\frac{q}{q+1} \times \frac{2k}{k+j}\right)^{\frac{k+j}{2}} \left(\frac{1}{q+1} \times \frac{2k}{k-j}\right)^{\frac{k-j}{2}}, \quad c(\tau) > 0,$$

ou encore, en remarquant que $\frac{2q}{q+1} = \frac{\tau+1}{\tau}$ et que $\frac{2}{q+1} = \frac{\tau-1}{\tau}$,

$$b_k(j) \sim \frac{c(\tau)}{\sqrt{j}} \left(\frac{1-\alpha}{1-\frac{\tau}{\tau+1}\alpha}\right)^{\frac{k+j}{2}} \left(\frac{1-\alpha}{1-\frac{\tau}{\tau-1}\alpha}\right)^{\frac{k-j}{2}}.$$

Un développement limité par rapport à l'infiniment petit α donne :

$$\begin{aligned} & \log \left[\left(\frac{1-\alpha}{1-\frac{\tau}{\tau+1}\alpha}\right)^{\frac{k+j}{2}} \left(\frac{1-\alpha}{1-\frac{\tau}{\tau-1}\alpha}\right)^{\frac{k-j}{2}} \right] \\ &= \frac{j}{2}(\tau+1-\tau\alpha) \log \left(\frac{1-\alpha}{1-\frac{\tau}{\tau+1}\alpha}\right) + \frac{j}{2}(\tau-1-\tau\alpha) \log \left(\frac{1-\alpha}{1-\frac{\tau}{\tau-1}\alpha}\right) \\ &= \frac{j}{2} \left[(\tau+1)(-\alpha + \frac{\tau}{\tau+1}\alpha) + (\tau-1)(-\alpha + \frac{\tau}{\tau-1}\alpha) + O(\alpha^2) \right] \\ &= jO(\alpha^2) = O(1). \end{aligned}$$

Dans l'équivalent de $b_k(j)$, le facteur $\left(\frac{1-\alpha}{1-\frac{\tau}{\tau+1}\alpha}\right)^{\frac{k+j}{2}} \left(\frac{1-\alpha}{1-\frac{\tau}{\tau-1}\alpha}\right)^{\frac{k-j}{2}}$ est donc minoré par une constante strictement positive. Cela fournit la minoration de $b_k(j)$ cherchée. \square

Ceci termine la démonstration du lemme 6.4.5, et donc du théorème 6.4.1. \square

Remarquons que nous ne nous sommes pas servis des résultats L^2 de la section 6.3. Nous allons toutefois utiliser maintenant la méthode L^2 pour démontrer le théorème 6.4.2.

6.4.2. Le théorème d'interpolation pour une famille analytique d'opérateurs. — La démonstration du théorème 6.4.2 repose sur le théorème d'interpolation pour une famille analytique d'opérateurs $T(\lambda)$ (voir [SW71, chapitre V] par exemple). Rappelons d'abord la version de ce résultat dont nous aurons besoin.

Soient $1 \leq p_0, p_1 \leq +\infty$ deux nombres réels et p'_0, p'_1 les exposants conjugués. Pour chaque λ dans une bande du plan complexe $a_0 \leq \Re\lambda \leq a_1$, on se donne un opérateur $T(\lambda)$ qui est borné dans $L^p(X, m)$, pour $p_0 \leq p \leq p_1$. On suppose également que, pour tous $f \in L^{p_0}(X, m) \cap L^{p_1}(X, m)$ et $g \in L^{p'_0}(X, m) \cap L^{p'_1}(X, m)$,

$$\lambda \mapsto \int_X T(\lambda)(f) g \, dm$$

est une fonction holomorphe dans la bande $a_0 < \Re\lambda < a_1$, continue jusqu'au bord de la bande.

Le théorème suivant donne une estimation des normes d'opérateurs dans la bande.

Théorème 6.4.7. — *On suppose de plus qu'il existe des constantes $C, \kappa > 0$ telles que pour tout $\beta \in \mathbb{R}$,*

$$\|T(a_0 + i\beta)\|_{L^{p_0} \rightarrow L^{p_0}} \leq C e^{\kappa\beta^2} \quad \text{et} \quad \|T(a_1 + i\beta)\|_{L^{p_1} \rightarrow L^{p_1}} \leq C e^{\kappa\beta^2}.$$

Il existe une constante C' , indépendante de T , telle que si $t \in]0, 1[$ et si p_t est défini par

$$\frac{1}{p_t} = \frac{1-t}{p_0} + \frac{t}{p_1},$$

alors

$$\|T((1-t)a_0 + ta_1)\|_{L^{p_t} \rightarrow L^{p_t}} \leq C'.$$

Nous allons utiliser ce théorème d'interpolation pour l'opérateur maximal $f \mapsto f_\sigma^*$ de la façon suivante. Pour des raisons qui seront expliquées à la fin de la sous-section 6.4.5, nous étudions plutôt l'opérateur $f \mapsto \sup_n |\sigma_{2n} \cdot f|$. Comme, pour $f \geq 0$, on a

$$\sigma_{2n-1} \cdot f \leq \sigma_{2n-1} \cdot f + q\sigma_{2n+1} \cdot f = (q+1)\sigma_{2n}(\sigma_1 \cdot f),$$

pour démontrer le théorème 6.4.2, il suffit en fait de savoir qu'il existe une constante $C > 0$, indépendante de N , telle que $\|f_N^*\|_p \leq C\|f\|_p$ pour tout $f \in L^p(X, m)$, où

$$f_N^* = \max_{n \leq N} |\sigma_{2n} \cdot f|.$$

Reprenant une idée déjà apparue dans la démonstration du théorème 6.3.6, il suffit même de majorer uniformément l'application qui à f fait correspondre

$$x \mapsto \sigma_{2n(x)} \cdot f(x) = \frac{1}{|S_{2n(x)}|} \sum_{|w|=2n(x)} f(wx),$$

où $x \mapsto n(x)$ est n'importe quelle application mesurable de X dans $\{0, 1, \dots, N\}$. Pour f donné, cela s'appliquera en particulier à une fonction $x \mapsto n(x)$ telle que

$$\max_{n \leq N} |\sigma_{2n} \cdot f(x)| = |\sigma_{2n(x)} \cdot f(x)|.$$

Ayant fixé une telle application mesurable, nous allons insérer l'application linéaire qui associe à f la fonction $x \mapsto \sigma_{2n(x)} \cdot f(x)$ dans une famille analytique d'opérateurs de moyennes.

6.4.3. Les sommes de Cesàro de paramètre complexe. — Pour $\lambda \in \mathbb{C}$, on définit les sommes de Cesàro complexes par

$$S_n^\lambda = \sum_{k=0}^n A_{n-k}^\lambda \sigma_{2k}, \quad \text{avec } A_n^\lambda = \begin{cases} \frac{(\lambda+1)\cdots(\lambda+n)}{n!} & \text{si } n \neq 0 \\ 1 & \text{si } n = 0 \end{cases},$$

et on leur associe les opérateurs maximaux

$$S_\lambda^* f = \sup_{n \geq 1} \left| \frac{1}{(n+1)^{\lambda+1}} S_n^\lambda \cdot f \right|.$$

Ces sommes de Cesàro ont été introduites depuis longtemps pour étudier la sommabilité des séries de Fourier (voir par exemple [Zyg02, Chapitre III]). L'observation fondamentale pour nous est que $S_n^{-1} = \sigma_{2n}$ et $S_n^0 = \sum_{k=0}^n \sigma_{2k}$. Ayant fixé un entier m et un réel $r > 1$, on va voir -1 comme barycentre de $-m$ et de 0 . On établira une estimation L^2 pour $\Re \lambda = -m$:

$$(6.4.3) \quad \|S_{-m+i\beta}^* f\|_2 \leq C_m e^{3\beta^2} \|f\|_2,$$

et une estimation L^r pour $\Re \lambda = 0$:

$$(6.4.4) \quad \|S_{i\beta}^* f\|_r \leq C_r e^{3\beta^2} \|f\|_r,$$

où les constantes C_m et C_r ne dépendent que de m et r respectivement.

Le théorème d'interpolation 6.4.7 va alors être appliqué à la famille analytique d'opérateurs $T(\lambda)$ avec

$$T(\lambda)f(x) = (n(x) + 1)^{-\lambda-1} S_{n(x)}^\lambda \cdot f(x).$$

Il est clair que l'opérateur $T(\lambda)$ est borné dans tout L^p (avec une borne dépendant *a priori* de $\max\{n(x), x \in X\} \leq N$). Les inégalités (6.4.3) et (6.4.4) donnent respectivement

$$\|T(-m + i\beta)\|_{L^2 \rightarrow L^2} \leq C_m e^{3\beta^2} \quad \text{et} \quad \|T(i\beta)\|_{L^r \rightarrow L^r} \leq C_r e^{3\beta^2}.$$

En écrivant ensuite $-1 = \lambda = -m(1-t) + 0t$ et $\frac{1}{p} = \frac{1-t}{2} + \frac{t}{r}$, le théorème 6.4.7 montre qu'il existe une constante $C'_{m,r}$, ne dépendant que de m et r , telle que $\|T(-1)\|_{L^p \rightarrow L^p} \leq C'_{m,r}$, d'où

$$\|f_N^*\|_p = \left\| \max_{1 \leq n \leq N} |S_n^{-1} \cdot f| \right\|_p \leq C'_{m,r} \|f\|_p,$$

pour tout $f \in L^p(X)$. Comme m peut être pris arbitrairement grand et r peut être pris arbitrairement proche de 1, p est aussi proche de 1 qu'on le souhaite.

Avant de passer à la démonstration des inégalités (6.4.3) et (6.4.4), rappelons quelques formules classiques. On a $S_0^\lambda = \delta_e$ pour tout $\lambda \in \mathbb{C}$ et, pour tout $n \geq 1$,

$$(6.4.5) \quad A_n^\lambda - A_{n-1}^\lambda = A_n^{\lambda-1} \quad \text{et} \quad S_n^\lambda - S_{n-1}^\lambda = S_n^{\lambda-1}.$$

Par ailleurs, en remarquant que

$$(1-x)^{-\lambda-1} = \sum_{k=0}^{\infty} A_k^\lambda x^k,$$

on obtient immédiatement la formule de convolution

$$(6.4.6) \quad A_n^{\lambda+\lambda'+1} = \sum_{k=0}^n A_{n-k}^{\lambda'} A_k^\lambda$$

qui entraîne

$$(6.4.7) \quad S_n^{\lambda+\lambda'+1} = \sum_{k=0}^n A_{n-k}^{\lambda'} S_k^\lambda.$$

Pour établir l'inégalité (6.4.4), observons d'abord que

$$(6.4.8) \quad |A_n^{i\beta}|^2 = (1+\beta^2) \left(1 + \frac{\beta^2}{2^2}\right) \cdots \left(1 + \frac{\beta^2}{n^2}\right) \leq e^{C\beta^2},$$

avec $C = \sum_{k \geq 1} \frac{1}{k^2} = \frac{\pi^2}{6} \leq 4$. Il en résulte que

$$\left| \frac{1}{(n+1)^{1+i\beta}} S_n^{i\beta} \cdot f \right| \leq e^{2\beta^2} \left(\frac{1}{n+1} \sum_{k=0}^n \sigma_{2k} \cdot |f| \right) \leq 2e^{2\beta^2} |f|_\nu^*.$$

L'inégalité (6.4.4) découle donc du théorème 6.4.1.

L'inégalité (6.4.3) est plus difficile à établir. On réduit son étude à celle de la proposition 6.4.8, comme nous allons l'expliquer maintenant.

6.4.4. Réduction du problème. — Nous allons montrer qu'il suffit d'établir l'inégalité (6.4.3) pour $\beta = 0$.

A partir de (6.4.7) on commence par écrire

$$S_n^{-m+i\beta} = \sum_{k=0}^n A_{n-k}^{i\beta} S_k^{-m-1} = \sum_{k=0}^{n'} A_{n-k}^{i\beta} S_k^{-m-1} + \sum_{k=n'+1}^n A_{n-k}^{i\beta} S_k^{-m-1}$$

où n' est la partie entière de $n/2$. Pour la seconde somme, on déduit de l'estimation $|A_k^{i\beta}| \leq e^{2\beta^2}$ que

$$(6.4.9) \quad (n+1)^{m-1} \left| \sum_{k=n'+1}^n A_{n-k}^{i\beta} S_k^{-m-1} \cdot f \right| \leq 2^m e^{2\beta^2} \sup_{n'+1 \leq k \leq n} (k+1)^m |S_k^{-m-1} \cdot f| \\ \leq 2^m e^{2\beta^2} |S_{-m-1}^* f|$$

puisque $(n+1)^m \leq 2^m (k+1)^m$ pour $n'+1 \leq k \leq n$.

Pour la première somme, on utilise l'identité

$$\sum_{k=0}^{n'} A_{n-k}^\gamma S_k^\delta = \sum_{k=0}^{n'} A_{n-k}^{\gamma-1} S_k^{\delta+1} + A_{n-n'-1}^\gamma S_{n'}^{\delta+1}$$

obtenue grâce à (6.4.5) et à une transformation d'Abel. On répète cette identité jusqu'à obtenir la somme

$$\sum_{k=0}^{n'} A_{n-k}^{i\beta} S_k^{-m-1} = \sum_{k=0}^{n'} A_{n-k}^{-m+i\beta} S_k^{-1} + \sum_{j=0}^{m-1} A_{n-n'-1}^{-j+i\beta} S_{n'}^{j-m}.$$

Pour $0 < j \leq \ell$, en écrivant à l'aide de la définition de A_ℓ^λ ,

$$A_\ell^{-j+i\beta} = A_{\ell-j}^{i\beta} \frac{(-j+1+i\beta) \cdots (i\beta)}{(\ell-j+1) \cdots \ell},$$

on obtient, grâce à l'inégalité (6.4.8), la majoration

$$\left| A_\ell^{-j+i\beta} \right| \leq e^{2\beta^2} \frac{(|j-1| + |\beta|) \cdots |\beta|}{(\ell-j+1) \cdots \ell} \leq e^{2\beta^2} \frac{(j+|\beta|)^j}{(\ell-j+1)^j}.$$

Par conséquent, il existe d'une constante $C(j)$ telle que pour tout $\beta \in \mathbb{R}$ et tout entier $\ell \geq j$, on ait

$$(\ell+1)^j \left| A_\ell^{-j+i\beta} \right| \leq C(j) e^{3\beta^2}.$$

Quitte à modifier la constante $C(j)$, on peut supposer que cette majoration est valable aussi pour $0 \leq \ell \leq j$.

On en déduit la majoration

$$\begin{aligned}
& \left\| \sup_n (n+1)^{m-1} \left| \sum_{k=0}^{n'} A_{n-k}^{-m+i\beta} S_k^{-1} \cdot f \right| \right\|_2 \\
& \leq C(m) e^{3\beta^2} \sup_n \left((n+1)^{m-1} \sum_{k=0}^{n'} \frac{1}{(n-k+1)^m} \right) \times \|S_{-1}^* f\|_2 \\
(6.4.10) \quad & \leq C'(m) e^{3\beta^2} \|S_{-1}^* f\|_2.
\end{aligned}$$

Par ailleurs, on a

$$\begin{aligned}
& \left\| \sup_n (n+1)^{m-1} \left| \sum_{j=0}^{m-1} A_{n-n'-1}^{-j+i\beta} S_{n'}^{j-m} \cdot f \right| \right\|_2 \\
& \leq e^{3\beta^2} \sum_{j=0}^{m-1} \left\| \sup_n (n+1)^{m-1} \frac{C(j)}{(n-n')^j} \left| S_{n'}^{j-m} \cdot f \right| \right\|_2 \\
& \leq C''(m) e^{3\beta^2} \sum_{j=0}^{m-1} \left\| \sup_n \frac{1}{(n'+1)^{j-m+1}} \left| S_{n'}^{j-m} \cdot f \right| \right\|_2 \\
(6.4.11) \quad & \leq C''(m) e^{3\beta^2} \sum_{j=0}^{m-1} \|S_{j-m}^* f\|_2.
\end{aligned}$$

On déduit alors des estimations (6.4.9), (6.4.10) et (6.4.11) que l'inégalité maximale (6.4.3) pour $S_{-m+i\beta}^*$ résulte de l'inégalité maximale L^2 pour S_{-m}^* , avec $m \geq 1$.

6.4.5. La fonction de Littlewood-Paley et la théorie L^2 . — Il nous reste donc à établir la proposition suivante.

Proposition 6.4.8. — *Quel que soit l'entier $m \geq 1$, il existe une constante C_m telle que, pour toute fonction $f \in L^2(X)$,*

$$\|S_{-m}^* f\|_2 \leq C_m \|f\|_2.$$

Nous allons montrer que la fonction maximale est dominée par une fonction de Littlewood-Paley pour laquelle on pourra utiliser la théorie L^2 . Pour cela, on écrit à l'aide de (6.4.5),

$$\begin{aligned}
(n+1)^m S_n^{-m} \cdot f &= \sum_{j=0}^n ((j+1)^m S_j^{-m} \cdot f - j^m S_{j-1}^{-m} \cdot f) \\
&= \sum_{j=0}^n [((j+1)^m - j^m) S_j^{-m} \cdot f + j^m S_j^{-m-1} \cdot f].
\end{aligned}$$

On en déduit la majoration

$$(6.4.12) \quad (n+1)^m |S_n^{-m} \cdot f| \leq m \sum_{j=0}^n (j+1)^{m-1} |S_j^{-m} \cdot f| + \sum_{j=0}^n (j+1)^m |S_j^{-m-1} \cdot f|.$$

Posons

$$R_m f = \left(\sum_{j=0}^{+\infty} (j+1)^{2m-1} |S_j^{-m-1} \cdot f|^2 \right)^{1/2}.$$

L'inégalité de Cauchy-Schwarz et l'inégalité $\sum_{j=0}^n (j+1) \leq (n+1)^2$ donnent

$$\frac{1}{n+1} \sum_{j=0}^n |a_j| \leq \frac{1}{n+1} \left(\sum_{j=0}^n (j+1) \right)^{1/2} \left(\sum_{j=0}^{+\infty} \frac{|a_j|^2}{j+1} \right)^{1/2} \leq \left(\sum_{j=0}^{+\infty} \frac{|a_j|^2}{j+1} \right)^{1/2}.$$

On déduit alors de (6.4.12) que

$$S_{-m}^* f \leq m R_{m-1} f + R_m f.$$

Il reste donc à établir une majoration de $\|R_m f\|_2$. Remarquons qu'une variante d'une telle majoration a été obtenue dans la démonstration du théorème 6.3.8 pour $m = 1$. La méthode employée ici est semblable et repose sur le théorème spectral qui permet d'écrire

$$\begin{aligned} \|R_m f\|_2^2 &= \sum_{j=0}^{+\infty} (j+1)^{2m-1} \int_X |S_j^{-m-1} \cdot f|^2 dm \\ &= \int_{\text{Sp}_{\mathbb{R}}(\sigma_1)} \sum_{j=0}^{+\infty} (j+1)^{2m-1} |\varphi_z(S_j^{-m-1})|^2 dm_f(z) \end{aligned}$$

où m_f est la mesure spectrale associée à f . Il s'agit alors de majorer uniformément le terme sous l'intégrale. C'est ici qu'est utilisé le fait que seules les mesures σ_k , avec k pair apparaissent dans la définition des S_n^λ . En effet, comme cela a déjà été remarqué dans la démonstration des théorèmes ergodiques L^2 , pour $f \in L^2(X)$ orthogonale aux sous-espaces $\text{Ker}(\sigma_1 + I)$ et $\text{Ker}(\sigma_1 - I)$, les caractères φ_0 et $\varphi_{i\zeta}$ de l'algèbre des fonctions radiales $\ell_{Rad}^1(\mathbb{F}_d)$ n'interviennent pas. On obtient une majoration uniforme en z de $\sum_{j=0}^{+\infty} (j+1)^{2m-1} |\varphi_z(S_j^{-m-1})|^2$ sur $\text{Sp}_{\mathbb{R}}(\sigma_1) \setminus \{0, i\zeta\}$ en utilisant notamment l'estimation (6.2.6). Pour les détails nous renvoyons à l'article [NS94]. Si f appartient à $\text{Ker}(\sigma_1 + I)$ ou à $\text{Ker}(\sigma_1 - I)$, on a $\sigma_{2k} \cdot f = f$ pour tout k , et dans ce cas on obtient directement l'inégalité maximale pour $S_{-m}^* f$.

6.5. Le théorème de Bufetov

L'approche du théorème de Nevo-Stein, due à Bufetov [Buf02], consiste à relier l'étude de la suite des opérateurs de moyenne $\pi(\sigma_{2n})$ à celle des puissances d'un unique

opérateur de Markov auxquelles s'applique le théorème de convergence de Rota. (Une démonstration de ce théorème est rappelée dans l'appendice C).

Le théorème de Bufetov s'énonce ainsi.

Théorème 6.5.1. — Soit (X, \mathcal{B}, m) un espace probabilisé sur lequel \mathbb{F}_d agit en préservant la mesure m . Pour toute fonction $f \in L^1(X, m)$ la suite $\sigma_{2n} \cdot f$ converge dans $L^1(X, m)$ vers $\mathbb{E}(f|\mathcal{I}_2)$, où \mathcal{I}_2 est la tribu des invariants de X sous l'action de $\mathbb{F}_d^{(2)}$. Si de plus $f \in L \log L(X, m)$, alors la convergence a lieu m -presque partout.

Démonstration. — Rappelons que $S = A \cup A^{-1}$ désigne le système générateur de \mathbb{F}_d et s'identifie à la sphère unité de \mathbb{F}_d . En particulier pour une fonction $f \in L^1(X, m)$, on a

$$\sigma_1 \cdot f(x) = \frac{1}{q+1} \sum_{s \in S} f(sx),$$

où $q = 2d - 1$. Bufetov introduit alors l'espace auxiliaire $(\tilde{X}, \tilde{\mathcal{B}}, \tilde{m})$ où \tilde{X} est l'espace produit $X \times S$, muni de la tribu produit et de la mesure $\tilde{m} = m \otimes p$, où p est la mesure de probabilité uniforme sur S . Il considère ensuite l'opérateur \tilde{P} sur (\tilde{X}, \tilde{m}) défini, pour toute fonction F de \tilde{X} dans \mathbb{R} , par

$$\tilde{P}F(x, s) = \frac{1}{q} \sum_{s' \in S; s' \neq e} F(sx, s').$$

Cet opérateur, qui préserve la mesure \tilde{m} , définit un opérateur de Markov sur $L^1(\tilde{X}, \tilde{m})$, au sens de la définition C.1.8 de l'appendice C. Le lien entre l'opérateur \tilde{P} et les opérateurs de moyenne sphérique σ_n est réalisé par le résultat suivant : si on associe à chaque fonction $f : X \rightarrow \mathbb{R}$ la fonction \tilde{f} définie sur \tilde{X} par $\tilde{f}(x, s) = f(x)$, alors

$$\sigma_1 \cdot f(x) = \frac{1}{q+1} \sum_{s \in S} \tilde{P}\tilde{f}(x, s).$$

De plus, l'itération de \tilde{P} conduit à

$$\tilde{P}^n F(x, s) = \frac{1}{q^n} \sum_{w \in S_n; w_n = s} \sum_{s' \in S; s' w_1 \neq e} F(wx, s'),$$

où le mot w est écrit $w = w_1 \cdots w_n$ en fonction de l'alphabet S . On en déduit la relation, pour $\tilde{f}(x, s) = f(x)$,

$$\sigma_n \cdot f(x) = \frac{1}{q+1} \sum_{s \in S} \tilde{P}^n \tilde{f}(x, s).$$

Si $f \in L \log L(X, m)$, la fonction \tilde{f} appartient à $L \log L(\tilde{X}, \tilde{m})$. L'étude de la convergence de la suite $\sigma_{2n} \cdot f$ se ramène donc à celle de la suite $\tilde{P}^{2n} F$, pour une fonction F de $L \log L(\tilde{X}, \tilde{m})$. Or, le théorème de Rota (voir C.1.9) montre que, si \tilde{P}^* désigne l'opérateur de $L^1(\tilde{X}, \tilde{m})$ adjoint de \tilde{P} , alors, pour toute fonction $F \in L^1(X, m)$, la suite de fonctions $(\tilde{P}^*)^n \tilde{P}^n F$ converge dans L^1 . De plus, la convergence

a aussi lieu presque sûrement lorsque $F \in L \log L(\tilde{X}, \tilde{m})$. Si l'opérateur \tilde{P} était auto-adjoint, on aurait la convergence souhaitée. Ce n'est pas le cas, mais nous allons voir qu'il existe une relation simple entre \tilde{P}^* et \tilde{P} . Notons tout d'abord que l'opérateur \tilde{P}^* , adjoint de \tilde{P} , vérifie, pour tout $F \in L^1(\tilde{X}, \tilde{m})$,

$$\tilde{P}^* F(x, s) = \frac{1}{q} \sum_{s' \in S; s' s \neq e} F(s'^{-1}x, s').$$

Soit U l'opérateur défini par

$$UF(x, s) = F(sx, s^{-1}).$$

Alors $U^2 = Id$ et on vérifie que U conjugue \tilde{P} à \tilde{P}^* :

$$\tilde{P} = U\tilde{P}^*U.$$

Lemme 6.5.2. — *On a de plus la relation :*

$$\tilde{P}^*\tilde{P} = \frac{q-1}{q}U\tilde{P} + \frac{1}{q}Id.$$

Démonstration. — En effet, on a d'une part

$$U\tilde{P}F(x, s) = \tilde{P}F(sx, s^{-1}) = \frac{1}{q} \sum_{s' \neq s} F(x, s'),$$

et d'autre part,

$$\begin{aligned} \tilde{P}^*\tilde{P}F(x, s) &= \frac{1}{q} \sum_{s' s \neq e} \tilde{P}F(s'^{-1}x, s') \\ &= \frac{1}{q^2} \sum_{s' \in S; s' s \neq e} \sum_{s'' \in S; s' s'' \neq e} F(x, s'') \\ &= \frac{q-1}{q} \sum_{s'' \neq s} F(x, s'') + \frac{1}{q}F(x, s). \end{aligned}$$

□

En utilisant ce lemme et la relation de conjugaison, on obtient alors par récurrence sur n ,

$$(\tilde{P}^*)^n \tilde{P}^n = \frac{q-1}{q^2} U \tilde{P}^{2n-1} + \frac{1}{q} (\tilde{P}^*)^{n-1} \tilde{P}^{n-1}$$

soit encore

$$\tilde{P}^{2n-1} = \frac{q}{q-1} U (\tilde{P}^*)^n \tilde{P}^n - \frac{1}{q-1} U (\tilde{P}^*)^{n-1} \tilde{P}^{n-1}.$$

Ainsi, la convergence de $\tilde{P}^{2n}F$ est une conséquence du théorème de Rota.

L'invariance de la limite par l'action de $\mathbb{F}_d^{(2)}$ résulte du théorème 6.2.1 et de la remarque 6.2.2, ce qui conclut la démonstration du théorème de Bufetov. □

Notice bibliographique. — Comme cela a déjà été dit, les résultats exposés dans ce chapitre sont essentiellement dus à Nevo [Nev94a] et Nevo et Stein [NS94]. L'approche proposée par Bufetov se situe dans le prolongement de travaux de Grigorchuk [Gri99, Gri00] et trouve ses racines dans la démonstration du théorème ergodique aléatoire de Kakutani [Kak51].

La méthode de Bufetov est très simple et s'adapte bien au cas des actions de semi-groupes libres (et plus généralement de semi-groupes strictement markoviens [Buf00]).

En revanche, l'approche de Nevo et Stein présente l'avantage d'établir des inégalités maximales. De plus l'inégalité maximale sous-additive énoncée dans le théorème 6.3.6 s'applique aux groupes hyperboliques de Gromov. À côté des groupes libres, on trouve dans cette classe les groupes fondamentaux de variétés compactes à courbure strictement négative et les réseaux uniformes de groupes de Lie simples de rang 1. La géométrie des groupes hyperboliques est suffisamment proche de celle des groupes libres pour que soit conservée l'inégalité maximale L^2 relative aux moyennes de Cesàro ν_n sur les sphères. Cependant les inégalités maximales pour les moyennes sur les sphères ou les boules, ainsi que les théorèmes ergodiques en moyenne ou ponctuels établis dans ce chapitre pour les groupes libres, requièrent une connaissance précise de la théorie spectrale de l'algèbre des fonctions radiales du groupe considéré. Elles ne s'étendent pas facilement aux groupes hyperboliques.

L'existence en toute généralité de théorèmes ergodiques en moyenne et ponctuels pour ces groupes est une question largement ouverte. Pour des résultats partiels nous renvoyons aux travaux de Fujiwara et Nevo [FN98] et de Nevo [Nev98]. Dans ce dernier article, un principe de transfert spectral [Cow79], combiné à des estimations de norme d'opérateurs de convolution [Jol90] permet notamment d'obtenir un théorème ergodique ponctuel avec vitesse de décroissance exponentielle pour les actions des réseaux uniformes du groupe $Sp(n, 1)$, $n \geq 2$.

CHAPITRE 7

THÉORÈMES ERGODIQUES POUR $SO^o(d, 1)$.

Le chapitre 6 était consacré à un exemple typique de structure hyperbolique discrète : le groupe libre \mathbb{F}_d à $d > 1$ générateurs, qui s'identifie à l'arbre homogène \mathbb{T}_q à $q+1 = 2d$ arêtes.

Dans le présent chapitre, nous considérons un exemple typique de structure hyperbolique continue : l'espace hyperbolique réel \mathbb{H}^d de dimension $d > 1$, sur lequel opère le groupe de Lorentz, plus précisément sa composante connexe $G = SO^o(d, 1)$. Nous avons rassemblé dans l'appendice D des informations sur \mathbb{H}^d et sur G et nous ne mentionnerons ici que la réalisation de \mathbb{H}^d comme espace homogène G/K avec $K = SO(d)$.

La parenté entre le cadre discret du chapitre 6 et le cadre continu du chapitre 7 est bien connue. Tout d'abord, il s'agit de structures hyperboliques au sens de Gromov (voir par exemple [Gro87], [GdlH90], [BH99]). Mais leurs liens sont plus étroits. D'un point de vue géométrique, les arbres (semi-)homogènes sont les immeubles affines de rang 1 tandis que les espaces hyperboliques (réels, complexes, quaternioniens, octonien) sont les espaces symétriques riemanniens de type non compact et de rang 1. D'un point de vue algébrique, $\mathbb{T}_q = PGL(2, \mathbb{Q}_q)/PGL(2, \mathbb{Z}_q)$ si q est un nombre premier tandis que $\mathbb{H}^2 = PGL(2, \mathbb{R})/PO(2)$. D'un point de vue analytique, la réalisation de ces espaces comme quotients G/K de paires de Gelfand (G, K) explique la commutativité du produit de convolution pour les fonctions (et les mesures ...) radiales. De plus, les analyses de Fourier radiales qui en résultent sont des cas limites d'une même théorie quantique de fonctions spéciales, élaborée par Cherednik et liée aux algèbres de Hecke affines doubles de rang 1 (voir par exemple [Che05, chap. 2], [Mac03]).

Si les arguments sont similaires à ceux du chapitre 6, les aspects techniques diffèrent suffisamment pour mériter un nouveau chapitre. En particulier, aux problèmes analytiques globaux s'ajoutent maintenant des problèmes analytiques locaux semblables à ceux du cas euclidien traité au chapitre 5.

Dans la suite, $G = SO^o(d, 1)$ et $K = SO(d)$. Un peu abusivement, nous appellerons sphères et boules centrées à l'origine e dans G les ensembles bi- K -invariants correspondant aux sphères et boules géodésiques centrées à l'origine $o = eK$ dans $\mathbb{H}^d = G/K$. Plus précisément, si $A = \{a_t; t \in \mathbb{R}\}$ est le groupe à un paramètre dans G tel que $t \mapsto a_t K$ soit une géodésique dans G/K alors, pour $t > 0$, la sphère \mathcal{S}_t de G , de rayon t centrée en e , est

$$\mathcal{S}_t = Ka_tK = \{g \in G; d(eK, gK) = t\},$$

et la boule \mathcal{B}_t de rayon t centrée en e est

$$\mathcal{B}_t = K \{a_s; 0 \leq s \leq t\} K = \{g \in G; d(eK, gK) \leq t\},$$

où d est la distance riemannienne dans \mathbb{H}^d . La sphère \mathcal{S}_t porte la mesure de probabilité σ_t définie par

$$\sigma_t = m_K * \delta_{a_t} * m_K,$$

où m_K est la mesure de Haar de K , et δ_{a_t} la masse de Dirac au point a_t . La boule \mathcal{B}_t porte la mesure de probabilité β_t correspondant au volume riemannien normalisé :

$$\beta_t = \frac{c_d}{|\mathcal{B}_t|} \int_0^t (\operatorname{sh} r)^{d-1} \sigma_r \, dr,$$

où $c_d = 2\pi^{d/2} / \Gamma(d/2)$ et $|\mathcal{B}_t| = c_d \int_0^t (\operatorname{sh} r)^{d-1} \, dr$. Nous introduisons également les moyennes $\nu_t = \frac{1}{t} \int_0^t \sigma_r \, dr$. Ces trois familles sont analogues aux suites de mesures σ_n , β_n et ν_n considérées au chapitre 6.

Dans ce chapitre, nous établissons plusieurs théorèmes ergodiques pour les moyennes sphériques en suivant les articles [Nev94b], [Nev97] de Nevo et [NS97] de Nevo-Stein.

L'algèbre $\mathcal{M}(G, K)$ des mesures complexes bi- K -invariantes sur G , de variation totale finie, est engendrée par les mesures de probabilité σ_t . Elle est l'analogue pour $SO^o(d, 1)$, de l'algèbre $l_{rad}^1(\mathbb{F}_d)$ des fonctions radiales. Comme cette dernière, c'est une algèbre de Banach involutive *commutative*, ce qui rend possible l'utilisation de méthodes spectrales.

Nous établissons d'abord (section 7.1) le théorème ergodique en moyenne L^2 dû à Howe-Moore [HM79], en suivant [Nev94b], pour les moyennes sur les sphères de $SO^o(d, 1)$, $d \geq 2$. L'évaluation de normes L^2 à l'aide du théorème spectral nécessite la détermination des caractères auto-adjoints de $\mathcal{M}(G, K)$. Ils sont donnés par des fonctions sphériques dont le comportement à l'infini permet de conclure.

Dans la section 7.2, nous établissons le théorème ergodique ponctuel L^2 de Nevo [Nev94b] pour les moyennes sur les sphères de $SO^o(d, 1)$, lorsque $d \geq 3$. Comme dans le chapitre précédent, la méthode de Littlewood-Paley-Stein sert, avec la théorie spectrale, à démontrer l'inégalité maximale L^2 . Ici, on a besoin d'estimer la décroissance de la première dérivée des fonctions sphériques. L'analyse spectrale de $\mathcal{M}(G, K)$

sert également pour démontrer que la convergence ponctuelle a lieu sur un sous-espace dense de $L^2(X, m)$.

La section 7.3 est consacrée au théorème ergodique ponctuel dans L^p pour les moyennes sphériques. Il suffit de démontrer l'inégalité maximale. Elle est vraie dans L^p pour tout $d \geq 3$ avec $p > \frac{d}{d-1}$. Sous cette forme optimale, elle est due à Nevo-Stein [NS97]. Nous nous limitons aux cas obtenus par Nevo dans [Nev97] : $p > \frac{d}{d-1}$ lorsque la dimension d est impaire et $p > \frac{d-1}{d-2}$ lorsque d est paire. La démonstration utilise la méthode d'interpolation complexe vue au chapitre 6 et l'intégration fractionnaire de Riemann-Liouville, qui est à rapprocher de l'utilisation des moyennes de Cesàro complexes de ce chapitre 6. Des estimations suffisamment précises et uniformes du comportement des fonctions sphériques et de leurs dérivées sont nécessaires.

Ces fonctions sphériques sont les fonctions propres radiales du laplacien sur \mathbb{H}^d . Nous terminons ce chapitre par un paragraphe technique (section 7.4), où nous donnons les estimations de ces fonctions propres et de leurs dérivées utilisées pour établir les théorèmes ergodiques. C'est l'une des difficultés des travaux de Nevo et Nevo-Stein, pour lesquels les estimations classiques d'Harish-Chandra sont insuffisantes. Nous démontrons ces estimations uniquement en dimension $d = 3$, où elles sont élémentaires.

7.1. Théorème ergodique en moyenne dans L^2

7.1.1. Énoncé. — Comme nous l'avons déjà fait dans la section 6.2, nous traitons plus généralement le cas d'une représentation unitaire π de G dans un espace hilbertien H . Cette représentation se prolonge en une représentation de l'algèbre de Banach involutive $\mathcal{M}(G)$ des mesures bornées sur G en posant $\pi(\mu) = \int_G \pi(g) d\mu(g)$ pour tout $\mu \in \mathcal{M}(G)$ (voir l'appendice A). Nous notons $E_1 \in \mathcal{B}(H)$ le projecteur orthogonal sur l'espace des vecteurs G -invariants de H .

Le théorème suivant est démontré dans [Nev94b]. Il peut aussi se déduire du théorème de Howe-Moore [HM79] relatif au comportement asymptotique des coefficients de π .

Théorème 7.1.1. — *Soit π une représentation unitaire de $G = SO^o(d, 1)$, $d \geq 2$. Alors $\pi(\sigma_t)$ converge fortement vers le projecteur E_1 quand t tend vers $+\infty$.*

La démonstration de ce théorème est analogue à celle du théorème 6.2.1. Elle repose sur l'étude des caractères auto-adjoints de la sous-algèbre $\mathcal{M}(G, K)$ de $\mathcal{M}(G)$ formée des mesures bi- K -invariantes, *i.e.* invariantes par les actions par multiplication à gauche et à droite de K sur G . Nous verrons que ces caractères s'expriment en termes de certaines fonctions propres sphériques du laplacien sur l'espace hyperbolique \mathbb{H}^d . Le théorème ergodique en moyenne découle d'une application du théorème spectral et des propriétés de décroissance de ces fonctions radiales. Notons que ce théorème est valable pour tout $d \geq 2$.

7.1.2. Détermination des caractères auto-adjoints de l'algèbre $\mathcal{M}(G, K)$ des mesures bi- K -invariantes sur G . — Concernant l'algèbre $\mathcal{M}(G)$ nous adoptons les notations introduites dans l'appendice A.

Lemme 7.1.2. — *L'algèbre $\mathcal{M}(G, K)$ est une algèbre de Banach commutative. Elle admet $\sigma_0 = m_K$ comme unité.*

Démonstration. — Soit $D : G \rightarrow \mathbb{R}_+$ l'application distance à l'origine définie par $D(g) = d(o, go)$. Alors toute mesure μ de $\mathcal{M}(G, K)$ s'écrit $\mu = \int_0^\infty \sigma_t d(D_*\mu)(t)$. Autrement dit, les mesures de $\mathcal{M}(G, K)$ sont exactement les moyennes des mesures σ_t . Comme $\sigma_t * \sigma_0 = \sigma_0 * \sigma_t$, on en déduit que σ_0 est l'unité de $\mathcal{M}(G, K)$.

Pour montrer que $\mathcal{M}(G, K)$ est une algèbre commutative, il suffit de vérifier que les mesures σ_t , pour $t \geq 0$, commutent entre elles. Or l'involution $\mu \mapsto \mu^*$ fixe chaque σ_t . Sa restriction à $\mathcal{M}(G, K)$ est donc l'involution complexe $\mu \mapsto \bar{\mu}$. Comme $(\sigma_t * \sigma_{t'})^* = \sigma_{t'}^* * \sigma_t^*$, on en déduit que $\bar{\sigma}_t * \bar{\sigma}_{t'} = \bar{\sigma}_{t'} * \bar{\sigma}_t = \bar{\sigma}_{t'} * \sigma_t$, d'où $\sigma_t * \sigma_{t'} = \sigma_{t'} * \sigma_t$ pour tous $t, t' \geq 0$. L'algèbre $\mathcal{M}(G, K)$ est donc bien commutative. \square

Nous aurons besoin de la propriété suivante [Nev94b, Lemma 2] :

Lemme 7.1.3. — *Soient t et t' deux nombres réels strictement positifs. Alors $\sigma_t * \sigma_{t'}$ est une mesure de probabilité absolument continue par rapport à la mesure de Haar sur G et elle a un support d'intérieur non vide.*

Nous allons maintenant déterminer les caractères auto-adjoints de $\mathcal{M}(G, K)$. Rappelons qu'ils sont automatiquement continus et de norme 1 (voir l'appendice B). Le groupe G n'ayant pas la combinatoire simple du groupe \mathbb{F}_d , on ne peut pas procéder "à la main" comme au chapitre 6. Suivant la méthode évoquée dans la remarque 6.2.7, l'étude se ramène à celle des fonctions propres du laplacien hyperbolique sur \mathbb{H}^d . Nous suivrons [Hel84, p. 399 et seq.].

Soit $L^1(G, K)$ la sous-algèbre de Banach involutive de $L^1(G)$ formée des fonctions à valeurs complexes intégrables pour la mesure de Haar et bi- K -invariantes. L'application $f \mapsto f dg$, où dg désigne la mesure de Haar de G , définit un homomorphisme isométrique de $L^1(G, K)$ dans $\mathcal{M}(G, K)$.

Tout caractère χ de $\mathcal{M}(G, K)$ se restreint donc naturellement en un caractère de $L^1(G, K)$. De plus, s'il existe $t > 0$ tel que $\chi(\sigma_t) \neq 0$, on a alors $\chi(\sigma_t * \sigma_t) = \chi(\sigma_t)^2 \neq 0$, et le lemme 7.1.3 assure que la restriction de χ à $L^1(G, K)$ est une forme linéaire non nulle et de norme inférieure ou égale à 1.

Définissons le *projecteur de moyennisation* $f \mapsto f^\sharp$ de $L^1(G)$ dans $L^1(G, K)$ par

$$f^\sharp(g) = \iint_{K \times K} f(k_1 g k_2) dk_1 dk_2.$$

Il est de norme inférieure ou égale à 1. À partir de la restriction de χ à $L^1(G, K)$, nous pouvons définir une forme linéaire continue L_χ de norme inférieure ou égale à 1 sur $L^1(G)$ par $L_\chi(f) = \chi(f^\sharp)$.

Lemme 7.1.4. — Soit χ un caractère auto-adjoint de $\mathcal{M}(G, K)$ dont la restriction à $L^1(G, K)$ est non nulle. Alors la forme linéaire L_χ associée sur $L^1(G)$ s'écrit $L_\chi(f) = \int_G f(g)\psi(g) dg$, où $\psi \in L^\infty(G)$ est une fonction à valeurs réelles, de norme inférieure ou égale à 1, bi- K -invariante, qui vérifie pour presque tous $x, y \in G$,

$$(7.1.1) \quad \psi(x)\psi(y) = \int_K \psi(xky) dk.$$

Démonstration. — Comme la forme linéaire L_χ est continue et de norme ≤ 1 , il existe $\psi \in L^\infty(G)$ telle que $\|\psi\|_\infty \leq 1$ et $L_\chi(f) = \int_G f(x)\psi(x) dx$ pour tout $f \in L^1(G)$. L'application L_χ est bi- K -invariante, donc ψ l'est aussi. Par ailleurs, ψ est à valeurs réelles car χ est auto-adjoint.

Enfin, pour f et h dans $L^1(G)$ on vérifie que $(f^\# * h)^\# = f^\# * h^\#$, d'où

$$\int_G (f^\# * h)(x)\psi(x) dx = L_\chi(f^\# * h) = \chi((f^\# * h)^\#) = \chi(f^\#)\chi(h^\#) = L_\chi(f)L_\chi(h).$$

Comme ψ est bi- K -invariante, le terme de gauche vaut

$$\int_{G \times G} f(x)h(y) \left(\int_K \psi(xky) dk \right) dx dy.$$

Celui de droite vaut

$$\iint_{G \times G} f(x)h(y)\psi(x)\psi(y) dx dy.$$

Donc ψ vérifie (7.1.1) pour presque tous x et y . □

Lemme 7.1.5. — (a) Soit $\psi \in L^\infty(G)$ une fonction non nulle vérifiant (7.1.1). Alors il existe une fonction propre K -invariante du laplacien sur \mathbb{H}^d , notée h , telle que $\psi(g) = h(gK)$ presque partout.

(b) Réciproquement, si h est une fonction propre K -invariante du laplacien sur \mathbb{H}^d qui vaut 1 en o , alors la fonction ψ définie sur G par $\psi(g) = h(gK)$ est K -invariante et vérifie (7.1.1) pour tous x, y de G .

Démonstration. — (a) L'équation (7.1.1) montre tout d'abord que ψ coïncide presque partout avec une fonction \mathcal{C}^∞ . En effet, choisissons une fonction $\rho : G \rightarrow \mathbb{R}$, de classe \mathcal{C}^∞ , à support compact, telle que $\int_G \rho(y)\psi(y) dy \neq 0$. Alors, pour presque tout $x \in G$, on a

$$\psi(x) \int_G \psi(y)\rho(y) dy = \int_G \rho(y) \left(\int_K \psi(xky) dk \right) dy = \int_K \left(\int_G \rho(k^{-1}x^{-1}z)\psi(z) dz \right) dk.$$

Le terme de droite est bien \mathcal{C}^∞ en x , et donc ψ l'est aussi.

L'équation (7.1.1) montre aussi que ψ est bi- K -invariante. Donc la fonction définie sur G/K par $h(yK) = \psi(y)$ est \mathcal{C}^∞ et K -invariante. Le laplacien Δ de l'espace hyperbolique $\mathbb{H}^d \simeq G/K$ est un opérateur invariant par l'action à gauche par isométries

de G sur G/K : pour tout $g \in G$, on a $\Delta h(gyK) = \Delta(h \circ g)(yK)$. En appliquant Δ aux deux membres de (7.1.1) considérés comme fonctions de y , il vient

$$h(xK)\Delta h(yK) = \int_K \Delta h(xkyK) dk,$$

d'où $h(xK)\Delta h(o) = \Delta h(xK)$ en prenant $y = e$. Donc h est une fonction propre de Δ pour la valeur propre $\Delta h(o)$.

(b) Soit h une fonction propre K -invariante du laplacien sur \mathbb{H}^d pour la valeur propre λ , telle que $h(o) = 1$. Comme nous le rappelons dans la section 7.4, il existe une unique fonction h avec ces propriétés.

Soit ψ la fonction définie sur G par $\psi(g) = h(gK)$. Pour tout $x \in G$, la fonction $y \mapsto \int_K \psi(xky) dk$ est bi- K -invariante. De plus, c'est une moyenne de fonctions propres pour la valeur propre λ ; elle induit donc une fonction K -invariante sur \mathbb{H}^d qui est propre pour la valeur propre λ , et qui vaut $\psi(x) = h(xK)$ en $y = e$. Par unicité de h , on en déduit que pour tout $y \in G$, on a $\psi(x)\psi(y) = \int_K \psi(xky) dk$. \square

En restriction à $L^1(G, K)$, un caractère auto-adjoint de $\mathcal{M}(G, K)$ s'écrit donc

$$\chi : f \mapsto \chi(f) = \int_G f(g)h(gK) dg,$$

où h est soit la fonction nulle, soit une fonction propre du laplacien, K -invariante, bornée, à valeurs réelles et telle que $h(o) = 1$. Nous verrons dans la section 7.4 que pour $\lambda \in \mathbb{C}$ fixé, il existe une unique fonction K -invariante, propre pour la valeur propre λ , qui vaut 1 à l'origine o : c'est la *fonction λ -propre sphérique*, notée h_s avec $\lambda = s(s-d+1)$. On a $h_s = h_{d-1-s}$. De plus la fonction h_s est réelle si et seulement si $s \in \mathbb{R}$ ou $\Re s = \frac{d-1}{2}$ et elle est bornée si et seulement si $\Re s \in [0, d-1]$ (voir proposition 7.4.3) .

L'ensemble $S = [0, \frac{d-1}{2}] \cup \{\frac{d-1}{2} + it ; t \geq 0\}$ paramètre donc exactement l'ensemble des fonctions sphériques bornées réelles.

En résumé, la restriction à $L^1(G, K)$ de tout caractère auto-adjoint de l'algèbre $\mathcal{M}(G, K)$ non identiquement nul sur les σ_t , $t > 0$, s'écrit sous la forme

$$\chi_s(f) = \int_G h_s(gK)f(g) dg,$$

avec $s \in S$. Notons que h_0 est la fonction constante égale à 1 ; par conséquent χ_0 est le caractère trivial $f \mapsto \int_G f(g) dg$.

Lemme 7.1.6. — Soient χ un caractère auto-adjoint de $\mathcal{M}(G, K)$ non identiquement nul sur les σ_t , $t > 0$, et $s \in S$ tel que $\chi(f) = \int_G h_s(gK)f(g) dg$, pour $f \in L^1(G, K)$. Alors, pour toute mesure $\mu \in \mathcal{M}(G, K)$, on a $\chi(\mu) = \int_G h_s(gK) d\mu(g)$.

Démonstration. — Commençons par remarquer que $\lim_{t \rightarrow 0} \chi(\sigma_t) = 1$. En effet, d'après le lemme 7.1.3, la mesure $\sigma_t * \sigma_t * \sigma_t$ est absolument continue, donc

$$(\chi(\sigma_t))^3 = \chi(\sigma_t * \sigma_t * \sigma_t) = \int_G \int_G \int_G h_s(ghkK) d\sigma_t(g) d\sigma_t(h) d\sigma_t(k)$$

tend vers 1 quand $t \rightarrow 0$. On a la même chose avec le carré $(\chi(\sigma_t))^2$, d'où le résultat pour $\chi(\sigma_t)$.

Ensuite, du lemme 7.1.3 et du fait que toute mesure μ dans $\mathcal{M}(G, K)$ s'écrit comme une intégrale des σ_t , on déduit que pour tout $\mu \in \mathcal{M}(G, K)$ tel que $\mu(K) = 0$ et pour tout $t > 0$ la mesure $\mu * \sigma_t$ est absolument continue. On a donc dans ce cas

$$\chi(\mu)\chi(\sigma_t) = \chi(\mu * \sigma_t) = \int_G \int_G h_s(ghK) d\mu(g) d\sigma_t(h).$$

En faisant tendre t vers 0, on obtient $\chi(\mu) = \int_G h_s(gK) d\mu(g)$.

Dans le cas général, il suffit de décomposer μ en la somme d'une mesure nulle sur K et d'un multiple de $\sigma_0 = m_K$. \square

Il existe en outre un seul caractère auto-adjoint de l'algèbre $\mathcal{M}(G, K)$ qui est identiquement nul sur les σ_t , pour $t > 0$. Il est défini par $\chi_\infty(\mu) = \mu(K)$ d'où $\chi_\infty(\sigma_0) = 1$ et $\chi_\infty(\sigma_t) = 0$ si $t > 0$. Le spectre auto-adjoint Sp_h de $\mathcal{M}(G, K)$ s'identifie donc à $S \cup \{\chi_\infty\}$: d'après le lemme 7.1.6, à tout $s \in S$ est associé le caractère auto-adjoint χ_s sur $\mathcal{M}(G, K)$ défini par la formule $\chi_s(\mu) = \int_G h_s(gK) d\mu(g)$. En particulier, $\chi_s(\sigma_t) = h_s(t)$, où $t \geq 0$ est la distance à l'origine o dans \mathbb{H}^d . La *topologie spectrale* (i.e. la topologie faible) sur l'ensemble Sp_h des caractères est la compactification de la topologie usuelle sur S (d'après la continuité de la fonction $s \mapsto h_s$ et son comportement à l'infini étudié dans la section 7.4).

7.1.3. Démonstration du théorème ergodique en moyenne. — Soit π une représentation unitaire de G sur un espace hilbertien H .

Comme dans le cas du groupe libre, pour démontrer le théorème ergodique en moyenne, nous appliquerons à l'algèbre $\mathcal{M}(G, K)$ les résultats de théorie spectrale rappelés dans l'appendice B. Or ces résultats ne sont valides que si l'élément unité σ_0 est envoyé par π sur l'élément unité de l'algèbre $\mathcal{B}(H)$ des opérateurs bornés sur H . Ceci n'est pas le cas ici, et nous considérerons donc plutôt l'espace $H^K = \pi(\sigma_0)H = \pi(m_K)H$ des éléments K -invariants de H .

Remarquons d'ailleurs que l'orthogonal de H^K ne joue aucun rôle dans les théorèmes ergodiques : en effet, d'une part $\pi(\mu)v = 0$ pour toute mesure μ bi- K -invariante et tout vecteur v orthogonal à H^K et d'autre part l'espace $E_1(H)$ des vecteurs G -invariants est bien sûr contenu dans H^K . Notons aussi que χ_∞ n'intervient pas car le projecteur spectral $E(\{\chi_\infty\})$ est nul : si v est dans son image, on a $\pi(f)v = 0$ pour tout $f \in L^1(G, K)$ d'après le lemme B.2.3, d'où $v = 0$. Ces observations seront souvent utilisées implicitement dans la suite.

Outre la théorie spectrale rappelée dans l'appendice B, la démonstration ci-dessous utilise des résultats sur les fonctions propres sphériques du laplacien qui sont présentés dans la section 7.4.

Notons B l'adhérence de l'algèbre $\pi(\mathcal{M}(G, K))$ dans l'algèbre $\mathcal{B}(H^K)$ pour la topologie de la norme d'opérateurs. C'est une C^* -algèbre commutative avec unité et par conséquent tous ses caractères sont auto-adjoints (voir l'appendice B.1). Le spectre de B s'injecte donc continûment dans le spectre auto-adjoint Sp_h de $\mathcal{M}(G, K)$ par l'application $\chi \mapsto \chi \circ \pi$; on identifiera ainsi ce spectre avec un sous-ensemble de l'espace compact $\text{Sp}_h = S \cup \{\chi_\infty\}$.

Nous procédons comme dans la démonstration du théorème 6.2.1. Soit v un vecteur de H^K , et m_v la mesure spectrale associée sur Sp_h . D'après l'appendice B, et puisque $\{\chi_\infty\}$ est de mesure spectrale nulle, nous avons

$$\|\pi(\sigma_t)v\|^2 = \int_S |h_s(t)|^2 dm_v(s).$$

Le corollaire 7.4.6 de la section 7.4 assure que $\lim_{t \rightarrow +\infty} h_s(t) = 0$ pour tout $s \in S \setminus \{0\}$. Le théorème de convergence dominée donne alors

$$\lim_{t \rightarrow \infty} \|\pi(\sigma_t)v\|^2 = m_v(\{0\}) = \|E_1(v)\|^2.$$

On en déduit immédiatement le théorème ergodique en moyenne. \square

7.2. Inégalités maximales et théorème ergodique ponctuel dans L^2

7.2.1. Énoncé. — Nous considérons une action de $G = SO^o(d, 1)$ sur un espace mesuré (X, \mathcal{B}, m) , où m est une mesure de probabilité G -invariante. Maintenant π désigne la représentation unitaire associée de G dans $L^2(X, m)$. Rappelons que si μ est une mesure bornée symétrique sur G on a $\pi(\mu)f = \mu \cdot f$ pour tout $f \in L^2(X, m)$.

Théorème 7.2.1 ([Nev94b]). — *On suppose $d \geq 3$. Alors, pour tout $f \in L^2(X, m)$, on a $\lim_{t \rightarrow +\infty} \sigma_t \cdot f = E_1 f = \mathbb{E}(f|\mathcal{I})$, où la convergence a lieu presque partout et dans L^2 et \mathcal{I} est la tribu des invariants.*

Pour démontrer ce théorème nous utiliserons la méthode désormais classique : nous montrerons une inégalité maximale dans $L^2(X, m)$ (théorème 7.2.2) pour les moyennes sur les sphères de G , puis nous exhiberons (paragraphe 7.2.3) un sous-espace dense \mathcal{D}' de fonctions de $L^2(X, m)$ qui vérifient le théorème ergodique ponctuel. Ceci suffira d'après le théorème 1.2.1 du chapitre 1.

7.2.2. Inégalité maximale L^2 . — Nous avons besoin dans ce paragraphe de la famille des mesures $\nu_t = \frac{1}{t} \int_0^t \sigma_r dr$. Nous considérons seulement les fonctions maximales à grande échelle $f_\sigma^*(x) = \sup_{t \geq 1} |\sigma_t \cdot f(x)|$ et $f_\nu^*(x) = \sup_{t \geq 1} |\nu_t \cdot f(x)|$ pour $f \in L^2(X, m)$. Leur étude suffit pour établir des théorèmes ergodiques. Le cas des

fonctions maximales locales est évoqué dans la remarque 7.3.9 (1). La fonction maximale f_ν^* est bien définie et mesurable pour toute fonction $f \in L^2(X, m)$. Mais les mesures σ_t étant singulières, la fonction maximale f_σ^* n'a aucune raison en général d'être définie (et encore moins mesurable). En pratique, on étudie cette fonction maximale pour un ensemble \mathcal{D} de fonctions f régulières et on démontre l'inégalité maximale du théorème 7.2.2 ci-dessous pour ces fonctions. Cette inégalité maximale permet alors de montrer que la fonction maximale f_σ^* est bien définie et mesurable pour tout $f \in L^2(X, m)$, et qu'elle vérifie l'inégalité maximale. Nous renvoyons à l'appendice A.2 pour des précisions sur cette question de la mesurabilité de f_σ^* lorsque $f \in L^2(X, m)$ (voir en particulier la proposition A.2.2, le théorème A.2.4 et la remarque A.2.3).

Théorème 7.2.2. — Soit $d \geq 3$ et soit $G = SO^o(d, 1)$ agissant sur (X, \mathcal{B}, m) . Il existe $C > 0$ tel que pour toute fonction $f \in L^2(X, m)$ on ait $\|f_\sigma^*\|_2 \leq C\|f\|_2$.

Pour démontrer ce théorème, nous établissons d'abord une inégalité maximale dans L^p , $p > 1$, pour la famille $\{\nu_t\}$. Ce résultat découle de l'inégalité maximale L^p pour une action de \mathbb{R} :

Proposition 7.2.3. — Soient $p > 1$ et $d \geq 2$. Il existe $C_p > 0$ tel que pour toute fonction $f \in L^p(X, m)$ on ait $\|f_\nu^*\|_p \leq C_p\|f\|_p$.

Démonstration. — Posons $M_t = \frac{1}{t} \int_0^t \delta_{a_s} ds$. Rappelons que π désigne la représentation de G associée à son action sur X . Par définition, on a

$$\nu_t = \frac{1}{t} \int_0^t \sigma_s ds = m_K * M_t * m_K,$$

d'où

$$|\sup_{t \geq 1} (\nu_t \cdot f(x))| \leq \pi(m_K) \left(\sup_{t \geq 1} \pi(M_t) (\pi(m_K)|f|)(x) \right).$$

Comme $\|\pi(m_K)\|_p \leq 1$, on voit que $\|f_\nu^*\|_p \leq \|\sup_{t \geq 1} \pi(M_t) (\pi(m_K)|f|)\|_p$. Le groupe $A = \{a_s; s \in \mathbb{R}\}$ étant isomorphe à \mathbb{R} , le théorème maximal de Hardy-Littlewood s'applique (voir section 3.1). Il existe donc $C_p > 0$ tel que

$$\|\sup_{t \geq 1} \pi(M_t) (\pi(m_K)|f|)\|_p \leq C_p \|\pi(m_K)|f|\|_p.$$

Puisque $\|\pi(m_K)|f|\|_p \leq \|f\|_p$, la proposition 7.2.3 est démontrée. \square

Nous allons maintenant comparer les deux fonctions maximales f_ν^* et f_σ^* , en utilisant comme au chapitre précédent la méthode de Littlewood-Paley-Stein. Nous aurons besoin d'une majoration précise de la première dérivée des fonctions sphériques. Cette estimation découle de la proposition 7.4.10, qui permet de conclure seulement lorsque $d \geq 3$.

Rappelons (théorème A.2.4) qu'on peut supposer que X est un espace métrique compact, l'action de G étant continue. Cela a donc un sens de considérer des fonctions continues sur X . Soit \mathcal{D} l'ensemble des fonctions sur X de la forme

$$f(x) = \pi(h)u(x) = \int_G h(g)u(g^{-1}x) dg$$

où u est une fonction continue sur X et h une fonction \mathcal{C}^∞ à support compact sur G . L'ensemble \mathcal{D} est inclus et dense dans $C(X)$, et donc dans $L^2(X)$. De plus, si $f \in \mathcal{D}$, la fonction $g \mapsto \pi(g)f$ est \mathcal{C}^∞ de G dans $C(X)$, et donc dans $L^2(X)$.

Si $f = \pi(h)u \in \mathcal{D}$, alors $t \mapsto \sigma_t \cdot f$ est \mathcal{C}^∞ de \mathbb{R}_+ dans $C(X)$ (et donc dans $L^2(X, m)$). En effet, puisque $m_K * h$ est une fonction \mathcal{C}^∞ à support compact sur G , on a $\pi(m_K)f \in \mathcal{D}$; ainsi l'application $t \mapsto \pi(a_t)(\pi(m_K)f)$ est \mathcal{C}^∞ et par composition avec $\pi(m_K)$ il en est donc de même pour $t \mapsto \sigma_t \cdot f$. On notera $\frac{d^k}{dt^k} \sigma_t \cdot f$ sa dérivée d'ordre k au point t .

Plus généralement, si $f \in L^2(X, m)$ est un vecteur \mathcal{C}^∞ pour la représentation unitaire de G dans $L^2(X, m)$ (voir l'appendice A), la fonction $t \mapsto \sigma_t \cdot f$ est \mathcal{C}^∞ de \mathbb{R}_+ dans $L^2(X, m)$.

D'après la proposition A.2.2 et la remarque A.2.3, le théorème 7.2.2 résulte de la proposition suivante.

Proposition 7.2.4. — *Soit $d \geq 3$. Il existe $C > 0$ tel que pour tout $f \in \mathcal{D}$, on ait $\|f_\sigma^*\|_2 \leq C\|f\|_2$.*

Démonstration. — Nous reprenons des idées déjà utilisées dans le chapitre 5 (lemme 5.1.5). Sachant que $t \mapsto \sigma_t \cdot f$ est \mathcal{C}^∞ de \mathbb{R}_+ dans $C(X)$, une intégration par parties donne, pour tout $r \geq 1$ et tout $x \in X$,

$$\int_1^r t \frac{d}{dt} \sigma_t \cdot f(x) dt = r \sigma_r \cdot f(x) - \sigma_1 \cdot f(x) - \int_1^r \sigma_t \cdot f(x) dt,$$

d'où

$$\frac{1}{r} \int_1^r t \frac{d}{dt} \sigma_t \cdot f(x) dt = \sigma_r \cdot f(x) - \frac{1}{r} \sigma_1 \cdot f(x) - \nu_r \cdot f(x) + \frac{1}{r} \nu_1 \cdot f(x).$$

Grâce à l'inégalité de Cauchy-Schwarz, on obtient

$$\left| \frac{1}{r} \int_1^r t \frac{d}{dt} \sigma_t \cdot f(x) dt \right|^2 \leq \frac{1}{r^2} \left(\int_1^r t dt \right) \left(\int_1^r t \left| \frac{d}{dt} \sigma_t \cdot f(x) \right|^2 dt \right).$$

Donc, si $x \in X$ et $r \geq 1$, on a

$$|\sigma_r \cdot f(x)| \leq |\nu_r \cdot f(x)| + g(f)(x) + |\sigma_1 \cdot f(x)| + |\nu_1 \cdot f(x)|,$$

où $g(f)$ est la fonction de Littlewood-Paley associée à f définie par

$$g(f)(x) = \left(\int_1^\infty t \left| \frac{d}{dt} \sigma_t \cdot f(x) \right|^2 dt \right)^{\frac{1}{2}}.$$

La majoration ci-dessus implique que pour tout $x \in X$, on a

$$|f_\sigma^*(x)| \leq 2|f_\nu^*(x)| + g(f)(x) + |\sigma_1 \cdot f(x)|.$$

La proposition 7.2.4 découle alors des propositions 7.2.3 et 7.2.5 et de ce que $f \mapsto \sigma_1 \cdot f$ est une contraction L^2 . \square

Proposition 7.2.5. — Soit $d \geq 3$. Il existe $C > 0$ tel que $\|g(f)\|_2 \leq C\|f\|_2$ pour tout $f \in \mathcal{D}$.

Démonstration. — On a

$$\begin{aligned} \|g(f)\|_2^2 &= \int_X \left(\int_1^\infty t \left| \frac{d}{dt} \sigma_t \cdot f(x) \right|^2 dt \right) dm(x) \\ &= \int_1^\infty t \left(\int_X \left| \frac{d}{dt} \sigma_t \cdot f(x) \right|^2 dm(x) \right) dt \\ &= \int_1^\infty t \left\| \frac{d}{dt} \sigma_t \cdot f \right\|_2^2 dt. \end{aligned}$$

Cette quantité se calcule en utilisant la théorie spectrale. Nous conservons les notations de la section 7.1.3 avec maintenant $H = L^2(X, m)$. Ainsi $H^K = \pi(\sigma_0)L^2(X, m)$ est l'espace des fonctions K -invariantes et B désigne l'adhérence en norme de l'image de $\mathcal{M}(G, K)$ par π dans l'algèbre $\mathcal{B}(H^K)$. Rappelons que le spectre de B est naturellement contenu dans le spectre auto-adjoint $\text{Sp}_h = S \cup \{\chi_\infty\}$ de l'algèbre $\mathcal{M}(G, K)$.

Nous n'utilisons ici le lemme suivant que pour $k = 1$, mais nous aurons besoin du cas $k > 1$ dans la démonstration du théorème ergodique ponctuel dans L^p (section 7.3).

Lemme 7.2.6 ([Nev94b],[Nev97]). — Soit $k \geq 1$. Si $f \in L^2(X, m)$ est un vecteur C^∞ , pour tout $t > 0$, on a

$$\left\| \frac{d^k}{dt^k} \sigma_t \cdot f \right\|_2^2 = \int_S |h_s^{(k)}(t)|^2 dm_f(s).$$

On rappelle que h_s est la fonction propre sphérique du laplacien définie au paragraphe 7.1.2.

Démonstration. — Observons que c'est seulement la projection orthogonale de f sur H^K qui intervient. On peut donc supposer que f appartient à H^K . Pour $\varepsilon > 0$, posons $U_\varepsilon = [\varepsilon, \frac{d-1}{2}] \cup \{\frac{d-1}{2} + i\tau; 0 \leq \tau \leq \frac{1}{\varepsilon}\}$ et notons $P_\varepsilon = E(U_\varepsilon)$ le projecteur spectral correspondant. Rappelons que P_ε commute avec B et que la mesure spectrale du vecteur $P_\varepsilon f$ est la restriction de m_f à U_ε .

Considérons d'abord le cas $k = 1$. Soit $\delta > 0$. D'après le théorème spectral, pour $t > 0$, on a

$$\frac{1}{\delta^2} \|P_\varepsilon(\sigma_{t+\delta} \cdot f - \sigma_t \cdot f)\|_2^2 = \frac{1}{\delta^2} \int_{U_\varepsilon} |h_s(t+\delta) - h_s(t)|^2 dm_f(s).$$

Puisque f est un vecteur C^∞ , le terme de gauche tend vers $\|P_\varepsilon(\frac{d}{dt}\sigma_t \cdot f)\|_2^2$, lorsque $\delta \rightarrow 0$.

Pour tout $s \in U_\varepsilon$, la quantité $\frac{1}{\delta}|h_s(t + \delta) - h_s(t)|$ tend vers $|h'_s(t)|$ quand $\delta \rightarrow 0$. Comme s varie dans U_ε , son module $|s|$ est borné (par une constante dépendant de ε), et le corollaire 7.4.9 donne l'existence d'une constante $C_{\varepsilon, d}$ ne dépendant que de ε et de d telle que pour tout $t \geq 0$,

$$|h'_s(t)| \leq C_{\varepsilon, d} e^{-\varepsilon t} \leq C_{\varepsilon, d}.$$

Ceci permet d'appliquer le théorème de convergence dominée, qui donne

$$\left\| P_\varepsilon\left(\frac{d}{dt}\sigma_t \cdot f\right) \right\|_2^2 = \int_{U_\varepsilon} |h'_s(t)|^2 dm_f(s).$$

Le résultat pour $k = 1$ s'obtient en faisant tendre ε vers 0. On a en effet

$$\lim_{\varepsilon \rightarrow 0} P_\varepsilon\left(\frac{d}{dt}\sigma_t \cdot f\right) = E(S \setminus \{0\})\frac{d}{dt}\sigma_t \cdot f = \frac{d}{dt}\sigma_t \cdot f - E(\{0\})\frac{d}{dt}\sigma_t \cdot f$$

et

$$E(\{0\})\frac{d}{dt}\sigma_t \cdot f = \frac{d}{dt}E(\{0\})\sigma_t \cdot f = \frac{d}{dt}\sigma_t \cdot E(\{0\})f = 0$$

puisque $E(\{0\})f$ est G -invariante.

Lorsque $k \geq 1$, la démonstration est analogue. On écrit

$$\frac{1}{\delta^{2k}} \left\| P_\varepsilon\left(\sum_{j=0}^k (-1)^{k-j} \binom{j}{k} \sigma_{t+j\delta} \cdot f\right) \right\|_2^2 = \frac{1}{\delta^{2k}} \int_{U_\varepsilon} \left| \sum_{j=0}^k (-1)^{k-j} \binom{j}{k} h_s(t+j\delta) \right|^2 dm_f(s).$$

Quand δ tend vers 0, le terme de gauche tend vers $\|P_\varepsilon(\frac{d^k}{dt^k}\sigma_t \cdot f)\|_2^2$, et on a

$$\lim_{\delta \rightarrow 0} \frac{1}{\delta^k} \left| \sum_{j=0}^k (-1)^{k-j} \binom{j}{k} h_s(t+j\delta) \right| = |h_s^{(k)}(t)|.$$

Le corollaire 7.4.9 assure qu'il existe une constante $C_{\varepsilon, d, k}$ telle que pour tout $s \in U_\varepsilon$ et tout $t \geq 0$,

$$|h_s^{(k)}(t)| \leq C_{\varepsilon, d, k} e^{-\varepsilon t} \leq C_{\varepsilon, d, k}.$$

Le théorème de convergence dominée s'applique donc, et le terme de droite tend donc vers $\int_{U_\varepsilon} |h_s^{(k)}(t)|^2 dm_f(s)$. On en déduit le résultat en faisant tendre ε vers 0. \square

Revenons à la démonstration de la proposition 7.2.5. On déduit en particulier du lemme ci-dessus que

$$\|g(f)\|_2^2 = \int_1^\infty t \int_S |h'_s(t)|^2 dm_f(s) dt = \int_S \left(\int_1^\infty t |h'_s(t)|^2 dt \right) dm_f(s).$$

Comme $m_f(S) = \|f\|_2^2$, il suffit donc de vérifier que la fonction $s \mapsto \int_1^\infty t |h'_s(t)|^2 dt$ est bornée sur S . Pour cela nous utilisons le corollaire 7.4.12 qui donne

$$\int_1^\infty t |h'_s(t)|^2 dt \leq C(1 + |s|)^{2(1 - \frac{d-1}{2})}.$$

Le terme de droite est borné sur S si et seulement si $d \geq 3$. Ceci permet donc de conclure la démonstration de la proposition 7.2.5 seulement dans ce cas. \square

7.2.3. Démonstration du théorème ergodique ponctuel dans L^2 . — Comme conséquence du théorème 7.2.2, il suffit, d'après le principe de Banach, de trouver un sous-ensemble dense dans $L^2(X, m)$ formé de fonctions qui vérifient le théorème 7.2.1.

Nous avons déjà remarqué qu'il suffit de se placer dans l'espace $H^K = \pi(\sigma_0)L^2(X, m)$ des fonctions K -invariantes. Notons $\mathcal{D}_\varepsilon \subset P_\varepsilon(H^K)$ l'ensemble des vecteurs K -invariants \mathcal{C}^∞ de $L^2(X, m)$ dont la mesure spectrale est à support dans l'ensemble $U_\varepsilon = [\varepsilon, \frac{d-1}{2}] \cup \{\frac{d-1}{2} + i\tau; 0 \leq \tau \leq \frac{1}{\varepsilon}\}$. Voici le dernier point crucial :

Lemme 7.2.7. — Soit $f \in \mathcal{D}_\varepsilon$. Alors $\lim_{t \rightarrow +\infty} \sigma_t \cdot f(x) = 0$ pour presque tout $x \in X$.

Démonstration. — Si $f \in \mathcal{D}_\varepsilon$, alors f est un vecteur \mathcal{C}^∞ pour la représentation de G dans $L^2(X, m)$, donc l'application $t \mapsto \sigma_t \cdot f$ est dérivable. Pour presque tout $x \in X$, et pour tous réels $t' \geq t$, on a alors

$$|\sigma_{t'} \cdot f(x) - \sigma_t \cdot f(x)| = \left| \int_t^{t'} \frac{d}{du} \sigma_u \cdot f(x) du \right| \leq \int_t^\infty \left| \frac{d}{du} \sigma_u \cdot f(x) \right| du.$$

Soit $\delta > 0$. La mesure de l'ensemble des $x \in X$ tels que $\int_t^\infty \left| \frac{d}{du} \sigma_u \cdot f(x) \right| du \geq \delta$ est majorée par

$$\frac{1}{\delta} \int_X \int_t^\infty \left| \frac{d}{du} \sigma_u \cdot f(x) \right| du dm(x) \leq \frac{1}{\delta} \int_t^\infty \left\| \frac{d}{du} \sigma_u \cdot f \right\|_2 du.$$

De plus, le lemme 7.2.6 donne $\left\| \frac{d}{du} \sigma_u \cdot f \right\|_2^2 = \int_S \left| \frac{\partial}{\partial u} h_s(u) \right|^2 dm_f(s)$.

D'après le corollaire 7.4.9, comme le module de s est borné sur U_ε , il existe une constante C_ε telle que si $s \in U_\varepsilon$ et $u \geq 0$, alors $\left| \frac{\partial}{\partial u} h_s(u) \right| \leq C_\varepsilon e^{-\varepsilon u}$, d'où

$$\left\| \frac{d}{du} \sigma_u \cdot f \right\|_2 \leq C_\varepsilon e^{-\varepsilon u} \|f\|_2.$$

On en déduit aisément que pour presque tout $x \in X$, l'intégrale $\int_t^\infty \left| \frac{d}{du} \sigma_u \cdot f(x) \right| du$ tend vers 0 quand $t \rightarrow +\infty$. Ceci implique alors que pour presque tout x , $\sigma_t \cdot f(x)$ a une limite lorsque $t \rightarrow +\infty$. D'après le théorème ergodique en moyenne, cette limite est égale à $E_1 f = E(\{0\})f$. Comme $f \in \mathcal{D}_\varepsilon$ et comme 0 n'appartient pas à U_ε , on voit que $E_1 f = 0$. \square

Posons $\mathcal{D}' = \cup_{\varepsilon > 0} \mathcal{D}_\varepsilon \oplus E_1 L^2(X, m) \subset H^K$. Cet ensemble est dense dans H^K et comme $\lim_{t \rightarrow +\infty} \sigma_t \cdot f = E_1 f$ presque partout pour tout $f \in \mathcal{D}'$, la démonstration du théorème 7.2.1 est terminée.

7.3. Inégalités maximales et théorème ergodique dans L^p

Nous supposons toujours que $G = SO^o(d, 1)$ agit sur (X, m) en laissant la mesure de probabilité m invariante. Nous étendons ici le théorème ergodique 7.2.1 au cas des espaces L^p . Ce résultat est dû à Nevo [Nev97] pour certaines valeurs de p , et à Nevo et Stein [NS97] en toute généralité.

Théorème 7.3.1 (Théorème ergodique ponctuel L^p , [Nev97] [NS97])

Soient un entier $d \geq 3$ et un réel $p > \frac{d}{d-1}$. Alors, pour tout $f \in L^p(X, m)$, on a $\lim_{t \rightarrow +\infty} \sigma_t \cdot f = \mathbb{E}(f|\mathcal{I})$, où la convergence a lieu presque partout et dans L^p et \mathcal{I} est la tribu des invariants.

À nouveau, la démonstration suit la méthode usuelle : comme on a convergence ponctuelle sur le sous-espace dense $L^2(X, m) \cap L^p(X, m)$ de $L^p(X, m)$, le théorème ergodique ponctuel L^p découle de l'inégalité maximale suivante.

Théorème 7.3.2 (Inégalité maximale pour les fonctions L^p , [Nev97] [NS97])

Soient un entier $d \geq 3$ et un réel $p > \frac{d}{d-1}$. Il existe $C_p > 0$ tel que pour toute fonction $f \in L^p(X)$, on ait $\|f_\sigma^*\|_p \leq C_p \|f\|_p$.

Dans cette section, $f \mapsto f_\sigma^*$ reste l'opérateur maximal relatif aux grandes sphères introduit dans la section 7.2. Nous ne présentons que la démonstration de Nevo [Nev97] qui donne le résultat pour tout $p > \frac{d}{d-1}$ quand $d \geq 3$ est impair et pour tout $p > \frac{d-1}{d-2}$ quand $d \geq 4$ est pair. Nous renvoyons à [NS97] pour le cas où d est pair et $\frac{d}{d-2} < p \leq \frac{d-1}{d-2}$.

L'idée de la démonstration est, comme dans le chapitre précédent, de faire intervenir une famille d'opérateurs qui dépendent analytiquement d'un paramètre complexe et d'utiliser le théorème d'interpolation de Stein. Le rôle des moyennes de Césàro fractionnaires est maintenant joué par les intégrales fractionnaires de Riemann-Liouville. Leur définition est rappelée dans le paragraphe suivant, dans lequel on fixe les notations utilisées.

7.3.1. Intégration fractionnaire de Riemann-Liouville. — Soit F une fonction sur \mathbb{R}_+ , de classe C^∞ , nulle dans un voisinage de 0 (plus tard, ce sera le produit de la fonction $t \mapsto \sigma_t \cdot f(x)$ et d'une fonction "cut-off" en t).

L'opérateur d'intégration fractionnaire d'ordre α est défini, pour $\alpha = a + ib$, $a > 0$, $t \geq 0$, par

$$I^\alpha F(t) = \frac{1}{\Gamma(\alpha)} \int_0^t (t-u)^{\alpha-1} F(u) du.$$

L'opérateur normalisé d'intégration fractionnaire d'ordre α est

$$M^\alpha F(t) = \frac{I^\alpha F(t)}{t^\alpha} = \frac{1}{\Gamma(\alpha)} \int_0^1 (1-u)^{\alpha-1} F(ut) du.$$

C'est une fonction analytique de α dans le demi-plan $\Re \alpha > 0$. Une intégration par parties permet d'écrire, du fait que la fonction à intégrer est identiquement nulle au voisinage de 0,

$$M^\alpha F(t) = \frac{t}{\Gamma(\alpha + 1)} \int_0^1 (1-u)^\alpha F'(ut) \, du.$$

Il en résulte que $M^\alpha F(t) = tM^{\alpha+1}F'(t)$. Par récurrence, on en déduit que $M^\alpha F(t)$ est, comme fonction de α , analytique dans tout le plan complexe. Remarquons que $I^0 F = F$ et que $I^{-k} F = F^{(k)}$ pour $k \in \mathbb{N}^*$. On vérifie que la dérivée d'ordre k de $I^k F$ vaut F , ce qui justifie, pour I^α , l'appellation d'intégrale fractionnaire d'ordre α . Plus généralement, $I^{\alpha+\beta} = I^\alpha \circ I^\beta$: on le vérifie aisément pour $\Re \alpha > 0$ et $\Re \beta > 0$, et on étend le résultat à tout α et tout β dans \mathbb{C} par prolongement analytique (voir aussi [Ste70b]).

Comme dans la sous-section 7.2.2, nous supposons que X est un espace métrique compact et que l'action est continue. Rappelons que \mathcal{D} est l'ensemble des fonctions sur X de la forme $x \mapsto \int_G h(g)u(g^{-1}x) \, dg$, où u est une fonction continue sur X et h une fonction \mathcal{C}^∞ à support compact sur G . Dans la suite, pour $x \in X$ fixé, on prendra $F_x(t) = \eta(t)\sigma_t \cdot f(x)$, avec $f \in \mathcal{D}$ et pour η une fonction de classe \mathcal{C}^∞ sur \mathbb{R} , à valeurs dans $[0, 1]$, telle que $\eta(t) = 0$ si $t < \frac{1}{4}$ et $\eta(t) = 1$ si $t > \frac{1}{2}$ (fonction "cut-off"). Posons

$$S_\alpha^* f(x) = \sup_{t \geq 1} |M^\alpha F_x(t)|.$$

Lorsque $\alpha = 0$ il est clair que cette fonction est égale à la fonction f_σ^* qui est notre objet d'étude. Lorsque $\alpha = 1$, elle est majorée par $|f|_\nu^*$.

7.3.2. Inégalités maximales pour les opérateurs S_α^* . — Pour utiliser le théorème d'interpolation de Stein, nous démontrons une inégalité maximale L^q , $q > 1$, pour l'opérateur S_{1+ib}^* (lemme 7.3.4), et une inégalité maximale L^2 pour l'opérateur S_{a+ib}^* , avec a bien choisi (proposition 7.3.5).

Les calculs nécessitent à plusieurs reprises l'estimation suivante.

Lemme 7.3.3. — Soit $a \in \mathbb{R}$. Il existe une constante $C_a > 0$ telle que pour tout $b \in \mathbb{R}$ on ait

$$(7.3.1) \quad \frac{1}{|\Gamma(a+ib)|} \leq C_a \exp(\pi|b|).$$

Démonstration. — La formule de Stirling donne, pour tout a réel, $|\Gamma(a+ib)| \sim \sqrt{2\pi} e^{-\frac{\pi}{2}|b|} |b|^{a-\frac{1}{2}}$ quand $|b| \rightarrow +\infty$ (voir [EMOT81] par exemple). Comme $\frac{1}{\Gamma(z)}$ est une fonction continue sur \mathbb{C} , cela permet d'obtenir, pour tout $a \in \mathbb{R}$, l'inégalité voulue. \square

Lemme 7.3.4. — Pour tout réel $q > 1$, il existe $C_q > 0$ tel que pour toute fonction $f \in \mathcal{D}$, on ait $\|S_{1+ib}^* f\|_q \leq C_q \exp(\pi|b|) \|f\|_q$.

Démonstration. — On a

$$|M^1 F_x(t) - \nu_t \cdot f(x)| = \frac{1}{t} \left| \int_0^{1/2} (\eta(u) - 1) \sigma_u \cdot f(x) du \right| \leq \nu_{1/2} \cdot |f|(x).$$

pour tout $t \geq 1$.

On déduit de la proposition 7.2.3 que S_1^* est continu dans L^q pour tout $q > 1$. C'est le lemme 7.3.4 lorsque $b = 0$. Pour b différent de 0, on utilise les inégalités

$$S_{1+ib}^* f(x) \leq \frac{1}{|\Gamma(1+ib)|} S_1^* |f|(x),$$

et (7.3.1). Le résultat du lemme découle alors du cas $b = 0$. \square

Nous nous intéressons maintenant à l'inégalité maximale L^2 .

Proposition 7.3.5. — Soient des entiers $d \geq 3$ et $1 \leq m \leq (d-1)/2$. Soit $\alpha = a + ib$, avec $-m + \frac{1}{2} < a$. Il existe $C_a > 0$ tel que, pour toute fonction $f \in \mathcal{D}$, on ait $\|S_\alpha^* f\|_2 \leq C_a \exp(\pi|b|) \|f\|_2$.

Nous établissons d'abord cette proposition dans le cas où α est le nombre entier $-k + 1$ avec $1 \leq k \leq m$ (lemme 7.3.8). Pour cela, nous utilisons encore la méthode de Littlewood-Paley-Stein. Si $f \in \mathcal{D}$, et $k \geq 1$, nous considérons la fonction

$$g_k(f)(x) = \left(\int_1^\infty t^{2k-1} |F_x^{(k)}(t)|^2 dt \right)^{1/2}.$$

Lemme 7.3.6. — Soient des entiers $d \geq 3$ et $1 \leq k \leq (d-1)/2$. Il existe $C_k > 0$ tel que pour tout $f \in \mathcal{D}$ on ait $\|g_k(f)\|_2 \leq C_k \|f\|_2$.

Démonstration. — Comme dans la démonstration de la proposition 7.2.5, on montre, grâce au lemme 7.2.6, que

$$\|g_k(f)\|_2^2 = \int_S \int_1^\infty t^{2k-1} |h_s^{(k)}(t)|^2 dt dm_f(s).$$

On utilise le corollaire 7.4.12 et le fait que $m_f(S) = \|f\|_2^2$ pour conclure. \square

Remarque 7.3.7. — Remarquons que la méthode précédente permet de même d'obtenir la continuité L^2 de l'application $f \mapsto G_k(f)(x) = \left(\int_0^\infty t^{2k-1} |F_x^{(k)}(t)|^2 dt \right)^{1/2}$ sur \mathcal{D} .

On a en effet

$$\left(G_k(f)(x) \right)^2 - \left(g_k(f)(x) \right)^2 = \int_{1/4}^1 t^{2k-1} |F_x^{(k)}(t)|^2 dt,$$

et on montre que la racine carrée de cette quantité est bornée dans L^2 en utilisant les mêmes estimées.

Nous énonçons le lemme suivant avec l'opérateur maximal *global* $f \mapsto M_{k+1}^* f(x) = \sup_{t>0} |M^{k+1} F_x(t)|$, car cette étape est nécessaire pour traiter ensuite le cas des opérateurs S_α^* .

Lemme 7.3.8. — *Soient des entiers $d \geq 3$ et $1 \leq k \leq (d-1)/2$. Il existe $C_k > 0$ tel que pour tout $f \in \mathcal{D}$, on ait $\|M_{-k+1}^* f\|_2 \leq C_k \|f\|_2$, d'où aussi $\|S_{-k+1}^* f\|_2 \leq C_k \|f\|_2$.*

Démonstration. — Si $k = 1$, comme $M^0 F_x(t) = \eta(t) \sigma_t \cdot f(x)$, on a $|M_0^* f(x)| \leq \sup_{t \geq 1/4} \sigma_t \cdot |f|(x)$ et le résultat se déduit de la proposition 7.2.4 dont la démonstration s'adapte facilement pour traiter le cas de $\sup_{t \geq 1/4} \sigma_t \cdot |f|(x)$ au lieu de $|f|_\sigma^*(x) = \sup_{t \geq 1} \sigma_t \cdot |f|(x)$. Si $k \geq 2$, on utilise que, pour $t > 0$ et $x \in X$,

$$\begin{aligned} M^{-k+1} F_x(t) &= t^{k-1} F_x^{(k-1)}(t) = \frac{1}{t} \int_0^t (u^k F_x^{(k-1)}(u))' du \\ &= \frac{1}{t} \int_0^t u^k F_x^{(k)}(u) du + \frac{k}{t} \int_0^t u^{k-1} F_x^{(k-1)}(u) du. \end{aligned}$$

Après application de l'inégalité de Cauchy-Schwarz, on en déduit que pour $t > 0$,

$$|M^{-k+1} F_x(t)| \leq G_k(f)(x) + k G_{k-1}(f)(x).$$

d'où $|M_{-k+1}^* f| \leq G_k(f) + k G_{k-1}(f)$. On conclut grâce à la remarque 7.3.7. \square

Nous pouvons maintenant démontrer la proposition 7.3.5.

Démonstration de la proposition 7.3.5. — Pour $\Re \alpha > -m$ on a

$$\begin{aligned} M^\alpha F_x(t) &= \frac{1}{\Gamma(\alpha+m)} \int_0^1 (1-u)^{\alpha+m-1} t^m F_x^{(m)}(ut) du \\ (7.3.2) \quad &= \frac{1}{\Gamma(\alpha+m)} \int_0^{\frac{1}{2}} (1-u)^{\alpha+m-1} t^m F_x^{(m)}(ut) du \end{aligned}$$

$$(7.3.3) \quad + \frac{1}{\Gamma(\alpha+m)} \int_{\frac{1}{2}}^1 (1-u)^{\alpha+m-1} t^m F_x^{(m)}(ut) du$$

Le terme (7.3.3) est estimé en utilisant l'inégalité de Cauchy-Schwarz. L'hypothèse $\alpha > -m + \frac{1}{2}$ intervient ici pour assurer la convergence de l'intégrale $\int_{\frac{1}{2}}^1 (1-u)^{2(\alpha+m-1)} du$. Plus précisément, on a

$$\begin{aligned}
& \left| \int_{\frac{1}{2}}^1 (1-u)^{\alpha+m-1} t^m F_x^{(m)}(ut) \, du \right| \\
& \leq \left(\int_{\frac{1}{2}}^1 (1-u)^{2(\alpha+m-1)} \, du \right)^{\frac{1}{2}} \left(\int_{\frac{1}{2}}^1 t^{2m-1} |F_x^{(m)}(ut)|^2 t \, du \right)^{\frac{1}{2}} \\
& \leq C_a'' \left(\int_{t/2}^t v^{2m-1} |F_x^{(m)}(v)|^2 \, dv \right)^{\frac{1}{2}} \\
& \leq C_a'' \left(\int_0^\infty v^{2m-1} |F_x^{(m)}(v)|^2 \, dv \right)^{\frac{1}{2}} = C_a'' G_m(f)(x),
\end{aligned}$$

en tenant compte du fait que $u \geq 1/2$ pour la deuxième inégalité.

Pour le terme (7.3.2), une suite d'intégrations par parties donne

$$\begin{aligned}
\frac{1}{\Gamma(\alpha+m)} \int_0^{\frac{1}{2}} (1-u)^{\alpha+m-1} t^m F_x^{(m)}(ut) \, du &= \frac{1}{\Gamma(\alpha)} \int_0^{1/2} (1-u)^{\alpha-1} F_x(ut) \, du \\
&+ \sum_{k=1}^m \frac{t^{k-1} F_x^{(k-1)}\left(\frac{t}{2}\right)}{2^{\alpha+k-1} \Gamma(\alpha+k)}
\end{aligned}$$

On observe que

$$\left| \int_0^{\frac{1}{2}} (1-u)^{\alpha-1} F_x(ut) \, du \right| \leq \left(\int_0^{\frac{1}{2}} (1-u)^{\alpha-1} \, du \right) \sup_{v>0} |F_x(v)| = C_a' M_0^* f(x).$$

Le deuxième terme est majoré en module par

$$\sum_{k=1}^m \frac{|M^{-k+1} F_x(\frac{t}{2})|}{2^{\alpha+k-1} |\Gamma(\alpha+k)|} \leq \sum_{k=1}^m \frac{M_{-k+1}^* f(x)}{2^{\alpha+k-1} |\Gamma(\alpha+k)|}.$$

On voit donc que

$$\sup_{t>0} |M^{a+ib} F_x(t)| \leq \frac{C_a' M_0^* f(x)}{|\Gamma(a+ib)|} + \sum_{k=1}^m \frac{M_{-k+1}^* f(x)}{2^{\alpha+k-1} |\Gamma(a+k+ib)|} + \frac{C_a'' G_m(f)(x)}{|\Gamma(a+m+ib)|}.$$

La proposition 7.3.5 découle maintenant du lemme 7.3.8, de la remarque 7.3.7 et de l'inégalité (7.3.1).

Observons qu'on a en fait obtenu le résultat pour l'opérateur maximal *global* M_{a+ib}^* . \square

7.3.3. Démonstration de l'inégalité maximale L^p . — On applique comme dans le cas du groupe libre le théorème d'interpolation de Stein (théorème 6.4.7) pour les familles analytiques d'opérateurs linéaires. Dans notre cas, on a $a_0 = a > -m + \frac{1}{2}$, $a_1 = 1$, $0 = ta + (1-t)$ et $\frac{1}{p} = \frac{t}{2} + \frac{1-t}{q}$, avec $q > 1$. Ceci donne pour tout $p > 1 + \frac{1}{2m}$ l'existence d'une constante $C_p > 0$ telle que pour tout $f \in L^p(X)$, on ait $\|S_0^* f\|_p \leq C_p \|f\|_p$. La condition $d \geq 2m + 1$ donne la restriction $p > \frac{d}{d-1}$ lorsque d est impair et la restriction $p > \frac{d-1}{d-2}$ lorsque d est pair.

Il y a deux différences entre les hypothèses habituelles du théorème d'interpolation et nos données. D'abord, les opérateurs d'intégration fractionnaire M^α sont définis pour les fonctions de \mathcal{D} , et non sur *les fonctions simples*, c'est-à-dire les combinaisons linéaires finies de fonctions indicatrices d'ensembles mesurables. D'autre part, les S_α^* ne sont pas des opérateurs linéaires, mais des opérateurs maximaux. Pour les détails de la méthode utilisée pour surmonter ces difficultés, nous renvoyons à [Nev97, p. 254]

Remarques 7.3.9. — (1) Les fonctions maximales locales

$$f_\sigma^\#(x) = \sup_{0 \leq t \leq 1} |\sigma_t \cdot f(x)|$$

sont étudiées dans [Nev97]. Nevo y démontre l'inégalité

$$\|f_\sigma^\#\|_p \leq C_p \|f\|_p$$

pour tout $f \in L^p(X, m)$, tout $d \geq 2$ et tout $p > d/(d-1)$ (comme dans le cas euclidien). Notons que ce résultat, ainsi que l'inégalité maximale du théorème 7.3.2, pour $d \geq 3$, sont optimaux (voir [Nev97, 5.4]).

(2) Dans [Nev94b, Appendix B], Nevo donne un exemple d'action de $PSL_2(\mathbb{R}) \simeq SO^o(2, 1)$ pour laquelle l'inégalité maximale L^2 est fautive. Dans ce cas $d = 2$, on ne sait pas si l'inégalité maximale à grande échelle et le théorème ergodique ponctuel sont vrais dans L^p pour $p > 2$.

(3) L'analogie du théorème de Wiener 3.1.3, c'est-à-dire le cas des moyennes sur les boules de $SO^o(d, 1)$ est étudié dans [NS97]. Soit $\mathcal{B}_t = \{g \in G; d(o, go) \leq t\}$ la boule du groupe $SO^o(d, 1)$, centrée en l'identité, de rayon $t \geq 0$. Nous notons β_t la restriction à \mathcal{B}_t de la mesure de Haar de $SO^o(d, 1)$, normalisée en une mesure de probabilité. La mesure β_t est un barycentre des mesures sur les sphères σ_r pour $0 \leq r \leq t$. Plus précisément, le lemme D.1.1 donne

$$\beta_t = \frac{\int_0^t (\operatorname{sh} r)^{d-1} \sigma_r \, dr}{\int_0^t (\operatorname{sh} r)^{d-1} \, dr}.$$

Nevo et Stein démontrent, avec les techniques exposées dans ce chapitre, que pour $d \geq 2$ et $p > 1$, l'inégalité maximale et le théorème ergodique ponctuel sont vrais dans L^p . Ainsi, pour $f \in L^p(X, m)$, on a

$$\lim_{t \rightarrow +\infty} \beta_t \cdot f = \mathbb{E}(f|\mathcal{I}), \quad \text{presque partout.}$$

Le cas $p = 1$ est ouvert.

On notera la différence de comportement avec le cas discret traité dans le chapitre précédent, où on a vu que les moyennes d'indice pair et celles d'indice impair convergent séparément.

7.4. Estimations des fonctions sphériques

La plupart des inégalités établies dans ce chapitre reposent sur le comportement des fonctions sphériques. Nous rassemblons dans ce paragraphe des informations utiles à leur sujet, qui sont extraites ou déduites de [Koo84] par exemple.

Entre autres caractérisations, les *fonctions sphériques* de la paire de Gelfand $(G, K) = (SO^o(d, 1), SO(d))$ sont les fonctions propres h du laplacien Δ sur l'espace hyperbolique $\mathbb{H}^d \simeq G/K$ qui sont K -invariantes, *i.e.* radiales, et qui valent 1 à l'origine $o = eK$. En coordonnées polaires, on aboutit ainsi à l'équation différentielle suivante sur $[0, +\infty[$:

$$(7.4.1) \quad \frac{d^2 h}{dr^2} + (d-1) \coth r \frac{dh}{dr} = \lambda h, \quad \text{avec } h(0) = 1,$$

qui possède exactement une solution, pour tout $\lambda \in \mathbb{C}$.

Passons à l'expression intégrale des fonctions sphériques (due à Harish-Chandra dans le cas général d'un espace symétrique riemannien). Dans le modèle du demi-espace, on vérifie facilement que $e_s(x) = x_d^s$ est une fonction propre du laplacien

$$\Delta = (x_d)^d \sum_{j=1}^d \frac{\partial}{\partial x_j} \left(x_d^{2-d} \frac{\partial}{\partial x_j} \right) = (x_d)^2 \sum_{j=1}^d \frac{\partial^2}{\partial x_j^2} - (d-2) x_d \frac{\partial}{\partial x_d}$$

pour la valeur propre $\lambda = s(s-d+1)$. En utilisant l'invariance du laplacien par isométries, on voit que la moyenne

$$h_s(x) = \int_K e_s(kx) \, dk$$

de e_s sur les sphères centrées à l'origine est une fonction sphérique pour la valeur propre $\lambda = s(s-d+1)$. En introduisant des coordonnées adéquates, on aboutit à l'expression explicite

$$(7.4.2) \quad h_s(r) = \frac{1}{c_d} \int_0^\pi (\operatorname{ch} r - \operatorname{sh} r \cos \theta)^{-s} (\sin \theta)^{d-2} \, d\theta,$$

où $c_d = \int_0^\pi (\sin \theta)^{d-2} \, d\theta = \sqrt{\pi} \Gamma(\frac{d-1}{2}) / \Gamma(\frac{d}{2})$ est une constante de normalisation. En résumé, on a le résultat suivant.

Proposition 7.4.1. — (a) *Les fonctions sphériques sur \mathbb{H}^d sont données par (7.4.2), où s est un paramètre complexe (et r est la distance à l'origine).*

(b) *On a $h_s = h_{d-1-s}$ et c'est le seul cas d'égalité.*

Remarques 7.4.2. — (1) Dans le modèle de la boule, $P(x, \xi) = (\operatorname{ch} r - \operatorname{sh} r \cos \theta)^{1-d}$ est le noyau de Poisson, pour le problème de Dirichlet hyperbolique. Ici x est un point intérieur, à distance hyperbolique r de l'origine, et ξ est un point frontière, faisant un angle θ avec x , vu de l'origine. En général h_s est une transformée de Poisson, pour la valeur propre $\lambda = s(s-d+1)$, de la fonction égale à 1 au bord.

(2) D'après la théorie générale, les fonctions sphériques sont élémentaires pour les espaces symétriques G/K avec G complexe, ce qui est le cas de $\mathbb{H}^3 \simeq \mathrm{SL}(2, \mathbb{C})/\mathrm{SU}(2)$. En dimension $d = 3$, l'expression (7.4.2) se réduit effectivement à

$$(7.4.3) \quad h_s(r) = \frac{\mathrm{sh}(s-1)r}{(s-1)\mathrm{sh}r}.$$

(3) Plus généralement, on a

$$h_s(r) = \frac{\Gamma(d-1)}{\Gamma(\frac{d-1}{2})} \frac{\Gamma(s-d+2)}{(s-\frac{d-1}{2})\Gamma(s)} \left(\frac{\partial}{\partial(2\mathrm{ch}r)} \right)^{\frac{d-1}{2}} \left(2\mathrm{ch}\left(s - \frac{d-1}{2}\right)r \right)$$

en dimension d impaire. Cette expression est également valable en dimension d paire, moyennant la définition suivante de la dérivée fractionnaire :

$$\left(\frac{\partial}{\partial(2\mathrm{ch}r)} \right)^{\frac{1}{2}} f(r) = \frac{1}{\sqrt{2\pi}} \int_0^r f'(s) \frac{ds}{\sqrt{\mathrm{ch}r - \mathrm{ch}s}}.$$

(4) D'autre part, en effectuant le changement de variables $z = -\mathrm{sh}^2 r$, l'équation (7.4.1) devient

$$z(1-z) \frac{d^2 h}{dz^2} + \left(\frac{1}{2}z - \frac{d}{2}(1-z) \right) \frac{dh}{dz} - \frac{\lambda}{4} h = 0.$$

Ceci conduit à une autre expression des fonctions sphériques, au moyen de la fonction hypergéométrique de Gauss :

$$h_s(r) = {}_2F_1 \left(\frac{s}{2}, \frac{d-1-s}{2}; \frac{d}{2}; -\mathrm{sh}^2 r \right).$$

Le reste de ce paragraphe est consacré à l'étude du comportement des fonctions sphériques. Nous énoncerons les résultats en dimension d quelconque, mais nous ne les démontrerons qu'en dimension $d = 3$, où l'expression des fonctions sphériques est la plus simple. Le cas général n'est qu'une affaire de technique. Commençons par quelques propriétés générales, qui se déduisent facilement de (7.4.3) en dimension $d = 3$.

Proposition 7.4.3. — (1) $h_s(r)$ est une fonction holomorphe en $s \in \mathbb{C}$ et analytique en $r \in \mathbb{R}$, avec les symétries $h_s(r) = h_{d-1-s}(r)$ et $h_s(r) = h_s(-r)$.

(2) h_s est réelle si et seulement si $s \in \mathbb{R} \cup \left(\frac{d-1}{2} + i\mathbb{R}\right)$.

(3) $h_s \geq 0$ si et seulement si $s \in \mathbb{R}$. Dans ce cas, on a en fait $h_s > 0$.

(4) h_s est bornée si et seulement si $\Re s \in [0, d-1]$. Dans ce cas, on a en fait $|h_s| \leq 1$.

Passons au comportement asymptotique des fonctions sphériques.

Proposition 7.4.4. — Fixons $s \in \mathbb{C}$ avec $\Re s \leq \frac{d-1}{2}$ et faisons tendre $r \rightarrow +\infty$. Alors

$$h_s(r) \sim \begin{cases} \mathbf{c}(-s) e^{-sr} & \text{si } \Re s < \frac{d-1}{2} \\ \mathbf{c}(-s) e^{-sr} + \mathbf{c}(s-d+1) e^{(s-d+1)r} & \text{si } s \in \frac{d-1}{2} + i\mathbb{R}^* \\ \mathbf{c}_0 r e^{-\frac{d-1}{2}r} & \text{si } s = \frac{d-1}{2} \end{cases}$$

où $\mathbf{c}(s) = \frac{\Gamma(d-1)}{\Gamma(\frac{d-1}{2})} \frac{\Gamma(s+\frac{d-1}{2})}{\Gamma(s+d-1)}$ et $\mathbf{c}_0 = 2 \operatorname{Res}_{s=-\frac{d-1}{2}} \mathbf{c}(s) = \frac{2\Gamma(d-1)}{\Gamma(\frac{d-1}{2})^2}$.

Démonstration. — En dimension $d = 3$, cet énoncé se réduit au comportement suivant, qui se déduit facilement de (7.4.3):

$$h_s(r) \sim \begin{cases} \frac{1}{1-s} e^{-sr} & \text{si } \Re s < 1 \\ \frac{1}{1-s} e^{-sr} + \frac{1}{s-1} e^{(s-2)r} & \text{si } s \in 1 + i\mathbb{R}^* \\ 2r e^{-r} & \text{si } s = 1 \end{cases}$$

□

Les fonctions sphériques peuvent être estimées globalement comme suit.

Proposition 7.4.5. — Pour tout $r \geq 0$ et pour tout $s \in \mathbb{C}$ avec $\Re s \leq \frac{d-1}{2}$, on a

$$|h_s(r)| \leq h_{\Re s}(r) \leq h_{\frac{d-1}{2}}(r) e^{(\frac{d-1}{2} - \Re s)r} \asymp (1+r) e^{-\Re s r}.$$

Démonstration. — En dimension $d = 3$, la première inégalité se réduit à

$$\left| \frac{\operatorname{sh}(x+iy)}{x+iy} \right| \leq \frac{\operatorname{sh} x}{x} \quad \text{pour tout } x+iy \in \mathbb{C},$$

qui résulte de l'inégalité élémentaire $\frac{\operatorname{sh} y}{y} \leq 1 \leq \frac{\operatorname{sh} x}{x}$. Comme $h_{\frac{d-1}{2}}(r) = \frac{r}{\operatorname{sh} r}$, la seconde inégalité se réduit à $\frac{\operatorname{sh} x}{x} \leq e^x$ pour tout $x \geq 0$. On conclut à l'aide de l'estimation $\frac{r}{\operatorname{sh} r} \asymp (1+r) e^{-r}$. □

On en déduit l'estimation suivante de h_s dans la bande verticale $0 \leq \Re s \leq \frac{d-1}{2}$.

Corollaire 7.4.6. — Pour tout $0 < \varepsilon < \frac{d-1}{2}$, il existe $C > 0$ tel que, pour tout $r \geq 0$,

$$|h_s(r)| \leq \begin{cases} C e^{-\varepsilon r} & \text{si } \Re s \in [\varepsilon, \frac{d-1}{2}], \\ C e^{-\Re s r} & \text{si } \Re s \in [0, \varepsilon]. \end{cases}$$

Ces estimations sont suffisantes pour nos besoins, mais elles ne sont pas optimales. En particulier, elles ne font pas apparaître la décroissance polynomiale en s que laisse escompter la proposition 7.4.4. À défaut d'un énoncé général, mentionnons le résultat particulier suivant, qui est une conséquence immédiate de (7.4.3) en dimension $d = 3$.

Proposition 7.4.7. — Pour tout $\varepsilon > 0$, il existe $C > 0$ tel que, pour tout $r \geq \varepsilon$ et pour tout $s \in \frac{d-1}{2} + i\mathbb{R}$ avec $\operatorname{Im} s \geq \varepsilon$,

$$|h_s(r)| \leq C |s|^{-\frac{d-1}{2}} e^{-\frac{d-1}{2}r}.$$

Les estimations précédentes des fonctions sphériques se transmettent à leurs dérivées à l'aide de la formule de Cauchy. Plus précisément, on utilise l'expression

$$(7.4.4) \quad h_s^{(k)}(r) = \frac{k!}{2\pi i} \int_{S(r_0)} \frac{h_s(z)}{(z-r)^{k+1}} dz$$

avec $S(r_0) = \{z \in \mathbb{C} \mid |z - r_0| = \frac{2}{1+|s-(d-1)/2|}\}$ et $|r - r_0| \leq \frac{1}{1+|s-(d-1)/2|}$. On obtient ainsi les résultats suivants.

Proposition 7.4.8. — Pour tout $k \in \mathbb{N}$, il existe $C > 0$ tel que, pour tout $r \geq 0$ et pour tout $s \in \mathbb{C}$ avec $\Re s \leq \frac{d-1}{2}$,

$$|h_s^{(k)}(r)| \leq C(1+|s|)^k h_{\Re s}(r).$$

Corollaire 7.4.9. — Pour tout $k \in \mathbb{N}$ et pour tout $0 < \varepsilon < \frac{d-1}{2}$, il existe $C > 0$ tel que, pour tout $r \geq 0$,

$$|h_s^{(k)}(r)| \leq \begin{cases} C(1+|s|)^k e^{-\varepsilon r} & \text{si } \Re s \in [\varepsilon, \frac{d-1}{2}], \\ C(1+|s|)^k e^{-\Re s r} & \text{si } \Re s \in [0, \varepsilon]. \end{cases}$$

Ces estimations sont insuffisantes pour nos besoins. D'une part, la croissance polynomiale en s y est trop élevée. D'autre part, elles ne rendent pas compte de l'annulation des dérivées de $h_0 = 1$. Voici deux résultats plus précis pour $s \in \frac{d-1}{2} + i\mathbb{R}$ grand et pour $s \geq 0$ petit.

Proposition 7.4.10. — Pour tout $k \in \mathbb{N}$ et pour tout $\varepsilon > 0$, il existe $C > 0$ tel que, pour tout $r \geq \varepsilon$ et pour tout $s \in \frac{d-1}{2} + i\mathbb{R}$ avec $|\Im s| \geq \varepsilon$,

$$|h_s^{(k)}(r)| \leq C|s|^{k-\frac{d-1}{2}} e^{-\frac{d-1}{2}r}.$$

Démonstration. — À l'aide de la formule de Cauchy (7.4.4), on peut déduire cette estimation de la proposition 7.4.7 étendue au voisinage de la droite $\Re s = \frac{d-1}{2}$. En dimension $d = 3$, on peut aussi l'établir à partir de (7.4.3), après l'avoir mise sous la forme suivante :

$$h_s(r) = \frac{1}{i \Im s} (e^{i \Im s r} - e^{-i \Im s r}) \sum_{j=0}^{+\infty} e^{-(2j+1)r}.$$

□

Proposition 7.4.11. — Pour tout $k \in \mathbb{N}^*$ et $\varepsilon > 0$, il existe $C_1 > 0$ et $C_2 > 0$ tels que, pour tout $s \in [0, \frac{1}{3}]$ et pour tout $r \geq \varepsilon$,

$$|h_s^{(k)}(r)| \leq C_1 s^k e^{-sr} + C_2 e^{-\frac{r}{3}}.$$

Démonstration. — En dimension $d = 3$, on obtient ce résultat en dérivant l'expression (7.4.3), après l'avoir mise sous la forme suivante :

$$h_s(r) = \frac{1}{1-s} e^{-sr} (1 - e^{-2(1-s)r}) \sum_{j=0}^{+\infty} e^{-2jr}.$$

□

En combinant les estimations ponctuelles précédentes, on obtient finalement l'estimation intégrale suivante, qui est cruciale pour l'inégalité maximale. Rappelons que $S = [0, \frac{d-1}{2}] \cup (\frac{d-1}{2} + i\mathbb{R}_+)$.

Corollaire 7.4.12. — Pour tout $k \in \mathbb{N}^*$, il existe $C > 0$ tel que, pour tout $s \in S$,

$$\left(\int_1^{+\infty} |h_s^{(k)}(r)|^2 r^{2k} \frac{dr}{r} \right)^{\frac{1}{2}} \leq C (1 + |s|)^{k - \frac{d-1}{2}}.$$

Démonstration. — Pour les grandes valeurs de $s \in \frac{d-1}{2} + i\mathbb{R}_+$, on utilise la proposition 7.4.10. Pour s voisin de 0, on utilise la proposition 7.4.11. Le corollaire 7.4.9 permet de traiter facilement les valeurs de s restantes. \square

Notice bibliographique. — Les résultats exposés dans ce chapitre sont extraits des articles [Nev94b, Nev97, NS97]. Dans [NS97], l'étude est étendue au cas de paires de Gelfand (G, K) où G est un groupe de Lie connexe simple, de centre fini, de rang réel 1, et K est un sous-groupe compact maximal de G . L'espace hyperbolique réel \mathbb{H}^d est remplacé par l'espace homogène G/K , muni de sa métrique riemannienne G -invariante naturelle. Les résultats pour les moyennes sphériques sont exactement les mêmes que pour $(G, K) = (SO^o(d, 1), SO(d))$, où maintenant $d \geq 3$ est la dimension de l'espace symétrique G/K . Le cas $d = 2$, qui concerne $SO^o(2, 1)$ et ses revêtements de centre fini, reste exclu de l'étude.

Margulis, Nevo et Stein [MNS00] ont par la suite considéré le cas général d'un groupe de Lie connexe semi-simple, de centre fini, sans facteur compact. À nouveau, K est un sous-groupe compact maximal et G/K est muni de sa métrique riemannienne canonique. Ici, les analogues du théorème ergodique ponctuel de Wiener et de l'inégalité maximale L^p , $p > 1$, sont établis pour les moyennes uniformes sur les boules $\mathcal{B}_t = \{g \in G; d(K, gK) \leq t\}$ de G . Mais alors que le théorème ergodique ponctuel de Wiener sur les boules de \mathbb{R}^d est également vrai pour $p = 1$, le cas $p = 1$ reste ouvert pour les groupes semi-simples. Notons que dans le cas particulier de l'action de G sur G/K , l'inégalité maximale L^p , $p > 1$, avait été obtenue par Clerc et Stein [CS74], puis ensuite améliorée en une inégalité de type faible $(1, 1)$ par Strömberg. La démonstration de Margulis, Nevo et Stein utilise d'une part les inégalités maximales établies dans [NS97] pour les groupes de rang 1 et, d'autre part, le fait que les groupes de Lie simples de rang supérieur possèdent la propriété (T) de Kazhdan.

En ce qui concerne les moyennes sphériques, à part le cas des groupes de rang 1 mentionné plus haut et celui des groupes de Lie connexes, complexes, semi-simples (voir [CN01]), la question générale des théorèmes ergodiques pour les moyennes sphériques sur les groupes de Lie semi-simples n'est pas résolue.

POSTFACE

Cet ouvrage ne traite que certains aspects de la théorie ergodique des actions de groupes. Nous avons principalement considéré des moyennes relatives aux boules et sphères associées à des exemples particuliers (bien que fondamentaux) de distances invariantes à gauche sur des groupes moyennables, libres ou semi-simples (spécialement $SO^o(d, 1)$).

Rappelons quelques-uns des résultats présentés dans ce livre, en ce qui concerne les boules. Pour la métrique définie par le choix d'un voisinage compact de l'élément neutre qui engendre le groupe (métrique des mots), nous avons vu au chapitre 3 que le théorème ergodique ponctuel dans L^1 est vrai lorsque le groupe est à croissance polynomiale et que l'on effectue les moyennes uniformes sur les boules. Pour les groupes moyennables à croissance exponentielle, ce problème est entièrement ouvert, même pour les boules hyperboliques du groupe des affinités du chapitre 4.

Sur un groupe de Lie G semi-simple connexe, sans facteur compact, de centre fini, nous n'avons considéré que la métrique provenant de la distance riemannienne naturelle sur l'espace symétrique G/K associé. Dans ce cas, le théorème ergodique ponctuel est vrai dans L^p pour $p > 1$, le cas $p = 1$ étant ouvert (voir la notice du chapitre 7).

Or, on peut s'intéresser à d'autres classes de groupes, par exemples les groupes algébriques semi-simples sur des corps locaux arbitraires et leurs réseaux. On peut aussi s'intéresser à des moyennes sur des boules ou sphères relatives à différents types de distances invariantes, ou même sur des familles de domaines plus généraux, dont les couronnes.

En outre, la question très importante de l'estimation de la vitesse de convergence dans les théorèmes ergodiques n'a pas été abordée dans ce livre. Sur ce sujet, tout un ensemble de résultats remarquables, liés à la notion de trou spectral, a été obtenu (voir [Nev98, MNS00, GNre]). On dit qu'une action ergodique d'un groupe localement compact G sur un espace probabilisé (X, m) , qui préserve m , a un trou spectral si

la représentation unitaire correspondante π_0 de G dans le sous-espace $L_0^2(X, m)$ des fonctions de carré intégrable orthogonales aux fonctions constantes, présente un trou spectral. Cela signifie que pour toute mesure de probabilité μ absolument continue et symétrique sur G , dont le support engendre un sous-groupe dense dans G , on a $\|\pi_0(\mu)\| < 1$ (voir la définition 2.1.10). C'est un phénomène qui ne se produit jamais pour les groupes moyennables. En revanche, toutes les actions ergodiques des groupes ayant la propriété (T) de Kazhdan ont un trou spectral, et dans ce cas on a même une majoration uniforme de $\|\pi_0(\mu)\|$ par un réel $\delta < 1$ qui ne dépend que de μ , et non pas de l'action (voir la proposition 2.1.13). Rappelons enfin que les groupes de Lie semi-simples connexes, de centre fini, sans facteur de rang 1, ont la propriété (T).

Dans [Nev98] et [MNS00], Margulis, Nevo et Stein ont tiré parti de ce phénomène de trou spectral pour obtenir des vitesses de convergence exponentielles. En particulier, la méthode de transfert spectral est un moyen très élégant d'estimer $\|\pi(\mu)\|$ pour une représentation unitaire π qui possède un trou spectral. Ce principe de transfert s'appuie notamment sur des travaux de Cowling [Cow79] et Howe-Moore [HM79]. Nevo en déduit par exemple que si G est un groupe de Lie simple, connexe, non compact, de centre fini, alors pour toute représentation unitaire π de G présentant un trou spectral et pour toute mesure de probabilité μ sur G , on a

$$\|\pi(\mu)\| \leq \|\lambda_{reg}(\mu)\|^{1/n},$$

où n est un entier ne dépendant que de π , et où λ_{reg} est la représentation régulière à gauche de G . Dans les cas considérés, si \mathcal{B}_t désigne la boule riemannienne de rayon t centrée en l'élément neutre de G et β_t la mesure de probabilité uniforme sur \mathcal{B}_t , il existe $c > 0$ tel que $\|\lambda_{reg}(\beta_t)\| \leq \exp(-ct)$ pour $t > 0$. On obtient donc ainsi la décroissance exponentiellement rapide de $t \mapsto \|\pi(\beta_t)\|$.

Voici un exemple typique de résultat qui découle de ces estimations en norme, à décroissance exponentiellement rapide.

Théorème. — *Soit G un groupe de Lie semi-simple, connexe, sans facteur compact, de centre fini, dont l'action ergodique sur l'espace probabilisé (X, m) a un trou spectral. Alors pour tout $f \in L^p(X, m)$, $p > 1$ et pour presque tout $x \in X$, on a*

$$\left| \beta_t \cdot f(x) - \int_X f \, dx \right| \leq B_p(f, x) \exp(-\theta_p t)$$

où θ_p est un réel > 0 qui dépend de p et de l'action.

Nous renvoyons à [Nev98], [MNS00] et à l'article d'exposition [Nev06] pour les détails et de nombreux commentaires.

La monographie récente de Gorodnik et Nevo [GNre] étudie de façon exhaustive les théorèmes ergodiques ponctuels et en moyenne pour des familles générales de moyennes uniformes sur les groupes algébriques semi-simples et leurs réseaux. Pour connaître l'essentiel sur les questions résolues et sur les problèmes ouverts dans le domaine de la théorie ergodique des actions de groupe, cet ouvrage [GNre] ainsi que

l'article d'exposition [Nev06] sont des lectures indispensables. On y trouve quantité d'informations et de résultats non présentés ici. Nous mentionnons en particulier les applications des théorèmes ergodiques à l'étude de l'équidistribution des orbites des actions isométriques ainsi qu'à l'estimation asymptotique de la proportion de points d'un réseau qui appartiennent à des domaines suffisamment réguliers du groupe ambiant, les boules par exemple, problèmes fondamentaux auxquels des solutions complètes sont apportées dans [GNre].

APPENDICE A

RÉGULARITÉ DES REPRÉSENTATIONS ET MESURABILITÉ DES FONCTIONS MAXIMALES

Pour éviter des complications inutiles, notre étude se place dans le cadre suivant (suffisamment général en pratique) :

- (X, \mathcal{B}) est toujours un *espace mesurable standard*, c'est-à-dire isomorphe à un borélien d'un espace métrique complet à base dénombrable d'ouverts, muni de la structure borélienne induite ;
- m est une mesure σ -finie sur (X, \mathcal{B}) ;
- G est un groupe localement compact à base dénombrable d'ouverts. On le munit de sa mesure de Haar à gauche, notée λ ou dg .

On suppose que ce groupe agit à gauche sur X de sorte que l'application $(g, x) \mapsto gx$ soit mesurable de $G \times X$ dans X ; on dit que (X, \mathcal{B}) est un *G -espace mesurable (standard)*. De plus, on suppose toujours que G *préserve la mesure* m : pour tous $g \in G$ et $E \in \mathcal{B}$, on a $m(gE) = m(E)$. Sous ces hypothèses, on dit que (X, \mathcal{B}, m) est un *G -espace mesuré (standard)*.

Nous introduisons d'abord la représentation de G dans $L^p(X, m)$ associée à cette action et nous en donnons quelques propriétés.

Comme nous l'avons remarqué dans les chapitres précédents, l'étude des fonctions maximales sur X définies par des familles non dénombrables de mesures de probabilité sur G pose des problèmes de mesurabilité. Dans la deuxième section de cet appendice, nous donnons quelques éléments permettant de répondre à ces questions.

A.1. Représentations de G et de $\mathcal{M}(G)$

Soient f une fonction mesurable sur X et $g \in G$. On note $\pi(g)f$ la fonction mesurable définie sur X par $x \mapsto f(g^{-1}x)$. L'application $\pi(g)$ induit une isométrie linéaire de $L^p(X, m)$, $1 \leq p \leq \infty$. On définit ainsi une *représentation* $\pi : g \mapsto \pi(g)$ de G dans $L^p(X, m)$. Autrement dit, π est un homomorphisme du groupe G dans le groupe des automorphismes isométriques de $L^p(X, m)$. Si $p < \infty$, cet homomorphisme

est fortement continu comme le montre le lemme suivant. Ce résultat est faux pour $p = \infty$.

Lemme A.1.1. — *Soit $1 \leq p < \infty$. Alors π est une représentation continue de G dans $L^p(X, m)$: pour tout $f \in L^p(X, m)$, l'application $g \mapsto \pi(g)f$ est continue de G dans $L^p(X, m)$.*

Démonstration. — Nous admettrons la mesurabilité de la fonction $g \mapsto \pi(g)f$ à valeurs dans l'espace de Banach $L^p(X, m)$ (voir [DS88, Lemma III.11.16 (b)]). Pour la continuité, il suffit de la montrer en l'élément neutre e de G . Fixons un compact K de G de mesure de Haar $|K| > 0$, ainsi qu'un voisinage compact V de e . On écrit, pour $g \in G$,

$$\|\pi(g)f - f\|_p = \frac{1}{|K|} \int_K \|\pi(h^{-1}g)f - \pi(h^{-1})f\|_p \, dh.$$

Soit $\varepsilon > 0$ donné. Comme $g \mapsto \pi(g)f$ est mesurable bornée, il existe une fonction continue ψ de $(VK)^{-1}$ dans $L^p(X, m)$ telle que $\int_{VK} \|\pi(h^{-1})f - \psi(h^{-1})\|_p \, dh \leq \varepsilon$. En outre, comme ψ est uniformément continue sur $(VK)^{-1}$, on peut trouver un voisinage symétrique W de e , contenu dans V , tel que $\|\psi(h^{-1}g) - \psi(h^{-1})\|_p \leq \varepsilon$ pour tout $h \in K$ et $g \in W$.

Enfin, en utilisant l'invariance à gauche de la mesure de Haar, on voit que

$$\int_K \|\pi(h^{-1}g)f - \psi(h^{-1}g)\|_p \, dh \leq \int_{VK} \|\pi(h^{-1})f - \psi(h^{-1})\|_p \, dh \leq \varepsilon,$$

si $g \in W$. On en déduit immédiatement que $\|\pi(g)f - f\|_p \leq (1+2/|K|)\varepsilon$ si $g \in W$. \square

Nous allons maintenant étendre π aux mesures bornées sur G . Notons $\mathcal{M}(G)$ l'espace de Banach formé des mesures complexes μ sur G de variation totale $\|\mu\|_1$ finie. On munit $\mathcal{M}(G)$ du produit de convolution défini par

$$\mu_1 * \mu_2(\varphi) = \int_{G \times G} \varphi(g_1 g_2) \, d\mu_1(g_1) \, d\mu_2(g_2)$$

pour toute fonction $\varphi : G \rightarrow \mathbb{C}$ continue à support compact sur G , et de l'involution définie par

$$\mu^*(\varphi) = \int_G \overline{\varphi(g^{-1})} \, d\mu(g).$$

Ces opérations font de $\mathcal{M}(G)$ une algèbre de Banach involutive.

La représentation π se prolonge en un homomorphisme de l'algèbre $\mathcal{M}(G)$ dans l'algèbre $\mathcal{B}(L^p(X, m))$ des opérateurs bornés de $L^p(X, m)$. Ce prolongement est défini de la manière suivante. Soient $\mu \in \mathcal{M}(G)$, $f \in L^p(X, m)$, $f_1 \in L^q(X, m)$ avec $1/p + 1/q = 1$. On observe que

$$\int_{X \times G} |f(g^{-1}x)| |f_1(x)| \, dm(x) \, d|\mu|(g) \leq \|f\|_p \|f_1\|_q \|\mu\|_1 < +\infty.$$

Alors, grâce au théorème de Fubini, on voit que pour m -presque tout x , la fonction $g \mapsto f(g^{-1}x)$ est μ -intégrable. On pose alors $\pi(\mu)f(x) = \int f(g^{-1}x) d\mu(g)$ et on déduit aisément de ce qui précède que $\pi(\mu)f \in L^p(X, m)$ avec $\|\pi(\mu)f\|_p \leq \|\mu\|_1 \|f\|_p$.

Lorsque μ admet la densité φ par rapport à la mesure de Haar, on écrira $\pi(\varphi)$ au lieu de $\pi(\mu)$.

Un calcul simple montre que $\pi(\mu * \nu) = \pi(\mu)\pi(\nu)$ si $\mu, \nu \in \mathcal{M}(G)$. Enfin, si $p = 2$, on vérifie que $\pi(g^{-1}) = \pi(g)^*$ et que $\pi(\mu^*) = \pi(\mu)^*$, où $\pi(\cdot)^*$ désigne l'adjoint de $\pi(\cdot)$.

Notations A.1.2. — Dans les théorèmes ergodiques on introduit plutôt l'antireprésentation $g \mapsto T(g) = \pi(g^{-1})$ (aussi notée T_g). En intégrant par rapport à $\mu \in \mathcal{M}(G)$ on obtient l'opérateur $T(\mu) = \pi(\check{\mu})$ où $\check{\mu}$ est la mesure $E \mapsto \mu(E^{-1})$. Pour alléger les notations, on écrit souvent $\mu \cdot f$ à la place de $T(\mu)f$, si $f \in L^p(X, m)$. Ainsi on a

$$\mu \cdot f(x) = \int_G f(gx) d\mu(g).$$

Tout ceci peut s'étendre à d'autres situations. Ainsi par exemple, une *représentation unitaire continue* de G dans un espace de Hilbert H est un homomorphisme π du groupe G dans le groupe unitaire de H tel que pour tout $v \in H$, l'application $g \mapsto \pi(g)v$ soit continue. Pour $\mu \in \mathcal{M}(G)$, on définit $\pi(\mu) \in \mathcal{B}(H)$ par la formule $\pi(\mu)v = \int_G \pi(g)v d\mu(g)$. Comme précédemment, π est un homomorphisme continu (préservant l'involution) de l'algèbre involutive $\mathcal{M}(G)$ dans $\mathcal{B}(H)$.

Si G est un groupe de Lie, il est souvent utile de travailler avec des vecteurs $v \in H$ tels que la fonction $g \mapsto \pi(g)v$ soit de classe \mathcal{C}^∞ . De tels v sont dits de classe \mathcal{C}^∞ (par rapport à π). Rappelons le lemme classique suivant dû à Gårding.

Lemme A.1.3. — Soit π une représentation unitaire continue d'un groupe de Lie G dans un espace hilbertien H .

- (i) Si φ est une fonction de classe \mathcal{C}^∞ à support compact sur G et si $v \in H$, alors le vecteur $\pi(\varphi)v$ est de classe \mathcal{C}^∞ .
- (ii) l'ensemble $\mathcal{D}(H)$ des vecteurs de la forme $\pi(\varphi)v$, où v décrit H et φ l'espace des fonctions \mathcal{C}^∞ à support compact, est dense dans H .

Démonstration. — Pour démontrer (i), on remarque d'abord qu'il suffit de vérifier que, pour tout $w \in H$, la fonction $g \mapsto \langle \pi(g)\pi(\varphi)v, w \rangle$ est de classe \mathcal{C}^∞ (voir [God03, Chap. XII, Théorème 40]). Mais

$$\langle \pi(g)\pi(\varphi)v, w \rangle = \int_G \varphi(g^{-1}h) \langle \pi(h)v, w \rangle dh,$$

et on conclut en dérivant sous le signe intégrale.

Pour démontrer (ii), donnons-nous $v \in H$ et $\varepsilon > 0$. On choisit un voisinage V de e tel que $\|\pi(g)v - v\| \leq \varepsilon$ pour tout $g \in V$. Soit alors φ une fonction positive, de classe \mathcal{C}^∞ sur G , d'intégrale 1, à support compact dans V . On vérifie immédiatement que $\|\pi(\varphi)v - v\| \leq \varepsilon$. \square

A.2. Mesurabilité des fonctions maximales

On revient à la donnée générale du G -espace (X, \mathcal{B}, m) . On considère en outre une famille continue à un paramètre $(\mu_r)_{r>0}$ de mesures de probabilité sur G . La continuité signifie ici que pour toute fonction φ continue bornée sur G , la fonction $r \mapsto \mu_r(\varphi)$ est continue. Comme on travaille avec des mesures qui peuvent être singulières (comme par exemple les mesures sur les sphères de \mathbb{R}^d) et avec des fonctions définies presque partout, l'étude des fonctions maximales f_μ^* soulève quelques difficultés.

Pour éviter les ambiguïtés, introduisons l'espace $\mathcal{L}^p(X, m)$ des fonctions mesurables sur X , de puissance p -ième intégrable (où, contrairement à $L^p(X, m)$, les fonctions égales presque partout ne sont pas identifiées).

Observons d'abord que si f est une fonction mesurable positive sur X , alors $\mu_r \cdot f(x)$ est bien défini pour tout $x \in X$, puisque la fonction $g \mapsto f(gx)$ est borélienne positive. Ainsi $f_\mu^*(x) = \sup_{r>0} \mu_r \cdot f(x)$ a un sens en tout point x , mais il n'est pas évident que f_μ^* soit mesurable. Si maintenant $f \in \mathcal{L}^p(X, m)$, on a déjà remarqué que, pour tout $r > 0$, $\mu_r \cdot f(x)$ est défini m -presque partout, mais il n'est pas clair qu'il existe un ensemble $E \subset X$ négligeable tel que si $x \notin E$, $\mu_r \cdot f(x)$ soit définie pour tout $r > 0$. C'est le premier obstacle à la définition dans ce cas de f_μ^* par la formule

$$f_\mu^*(x) = \sup_{r>0} |\mu_r \cdot f(x)|.$$

En outre, pour obtenir $f \mapsto f_\mu^*$ en tant qu'opérateur sur $L^p(X, m)$, il faudra s'assurer que la classe de f_μ^* ne dépend que de celle de f . Nous allons voir comment ces questions peuvent être résolues si l'on dispose *a priori* d'une inégalité maximale sur une partie dense convenable de $L^p(X, m)$. Les idées des démonstrations qui suivent sont essentiellement dues à Stein et Wainger [SW78, Section 8] (voir aussi [Ste93, Chap. XI, Section 5]).

Commençons par un résultat facile.

Lemme A.2.1. — Soit (μ_r) une famille continue à un paramètre de mesures de probabilité sur G . Soit f une fonction sur X qui est la limite simple d'une suite croissante $(f_n)_{n \geq 1}$ de fonctions mesurables positives pour lesquelles l'opérateur maximal donne une fonction mesurable. Alors :

- (i) $f_\mu^* = \lim_n (f_n)_\mu^*$ et donc f_μ^* est mesurable;
- (ii) si $p \in [1, +\infty[$ et s'il existe $C > 0$ tel que $\|(f_n)_\mu^*\|_p \leq C \|f_n\|_p$ pour tout n , on a $\|f_\mu^*\|_p \leq C \|f\|_p$.

Démonstration. — On a bien sûr $(f_n)_\mu^* \leq f_\mu^*$ pour tout n , d'où $\lim_n (f_n)_\mu^* \leq f_\mu^*$. D'autre part, grâce au théorème de Beppo Levi, on a $\mu_r \cdot f = \lim_n \mu_r \cdot f_n$ pour tout $r > 0$. Comme $(f_n)_\mu^* \geq \mu_r \cdot f_n$, on voit alors que $\lim_n (f_n)_\mu^* \geq \mu_r \cdot f$, d'où $\lim_n (f_n)_\mu^* \geq f_\mu^*$.

Ainsi, on a $f_\mu^* = \lim_n (f_n)_\mu^*$. Ceci prouve la mesurabilité de f_μ^* . L'assertion (ii) résulte alors facilement du théorème de Beppo Levi. \square

On suppose dans la proposition suivante que X est un espace métrique localement compact sur lequel G agit continûment (*i.e.*, l'application $(g, x) \mapsto gx$ est continue de $G \times X$ dans X). De plus, m est ici une mesure de Radon préservée par l'action de G . Dans ce cas, si f est continue et bornée sur X , la définition de f_μ^* ne pose pas de problème. En effet, pour tout $x \in X$, la fonction $g \mapsto f(gx)$ est continue et bornée sur G et donc μ_r -intégrable quel que soit $r > 0$. De plus, puisque $r \mapsto \mu_r \cdot f(x)$ est continue, la fonction f_μ^* est mesurable car

$$f_\mu^*(x) = \sup_{r \in \mathbb{Q}^+} |\mu_r \cdot f(x)|.$$

Proposition A.2.2. — Gardons les hypothèses ci-dessus et soit $1 \leq p < \infty$. On suppose l'existence d'une constante $C > 0$ telle que $\|f_\mu^*\|_p \leq C\|f\|_p$ pour toute fonction f continue à support compact sur X . Soit $f \in \mathcal{L}^p(X, m)$. Il existe un ensemble $E \subset X$, négligeable, tel que si $x \notin E$:

- (i) pour tout $r > 0$, la fonction $g \mapsto f(gx)$ est μ_r -intégrable sur G ;
- (ii) la fonction $r \mapsto \mu_r \cdot f(x) = \int_G f(gx) d\mu_r(g)$ est continue sur \mathbb{R}_+^* .

De plus, en posant $f_\mu^*(x) = \sup_{r > 0} |\mu_r \cdot f(x)|$ si $x \notin E$ et $f_\mu^*(x) = 0$ sinon, la fonction f_μ^* est mesurable et vérifie la même inégalité maximale $\|f_\mu^*\|_p \leq C\|f\|_p$. En particulier, la classe de f_μ^* ne dépend que de celle de f .

Démonstration. — Pour simplifier, on se limite au cas où X est compact. On va procéder en trois étapes.

Étape 1. C'est la clé de la démonstration. Si $B \subset X$ est un ensemble négligeable, on montre l'existence d'un ensemble négligeable $E \subset X$ tel que si $x \notin E$, l'ensemble $B_x = \{g \in G : gx \in B\}$ est μ_r -négligeable pour tout $r > 0$.

Tout d'abord, si U est un ouvert de X , la fonction caractéristique $\mathbf{1}_U$ est la limite de la suite croissante $(f_n)_{n \geq 1}$ où $f_n(x) = d(x, X \setminus U)^{1/n}$ (en choisissant la distance d telle que le diamètre de X soit plus petit que 1). Donc $(\mathbf{1}_U)_\mu^*$ est mesurable et vérifie l'inégalité maximale par le lemme A.2.1.

Maintenant, puisque B est négligeable, il existe une suite décroissante d'ouverts U_n contenant B , tels que $\lim_n m(U_n) = 0$. On a $(\mathbf{1}_B)_\mu^* \leq \lim_n (\mathbf{1}_{U_n})_\mu^*$ et

$$\|\lim_n (\mathbf{1}_{U_n})_\mu^*\|_p \leq \lim_n \|(\mathbf{1}_{U_n})_\mu^*\|_p \leq C \lim_n \|\mathbf{1}_{U_n}\|_p = 0,$$

d'où il résulte que $(\mathbf{1}_B)_\mu^*$ est négligeable ; c'est le résultat clé annoncé.

Étape 2. On considère le cas où f est mesurable positive et bornée. On choisit une suite bornée $(f_n)_{n \geq 1}$ de fonctions continues qui tend vers f presque partout et dans $L^p(X, m)$.

D'après l'étape 1, il existe un ensemble négligeable $E \subset X$ tel que, si $x \notin E$ et pour tout $r > 0$, la suite des fonctions $g \mapsto f_n(gx)$ converge vers la fonction $g \mapsto f(gx)$ pour μ_r -presque tout $g \in G$. Grâce au théorème de convergence dominée, on en déduit

alors que, pour $r > 0$ et $x \notin E$,

$$\lim_{n \rightarrow \infty} \mu_r \cdot f_n(x) = \mu_r \cdot f(x).$$

Par ailleurs, on peut supposer que $\|f_{n+1} - f_n\|_p \leq 2^{-n}$ pour $n \geq 1$. On en déduit, en utilisant l'inégalité maximale, que $\sum_n \|(f_{n+1} - f_n)_\mu^*\|_p < +\infty$, et donc aussi que $\sum_n (f_{n+1} - f_n)_\mu^*$ est finie presque partout. Par conséquent, on peut choisir E tel que si $x \notin E$, la convergence de $\mu_r \cdot f_n(x)$ vers $\mu_r \cdot f(x)$ soit uniforme en r sur $]0, +\infty[$. Comme $r \mapsto \mu_r \cdot f_n(x)$ est continue pour tout x , on voit que $r \mapsto \mu_r \cdot f(x)$ est continue si $x \notin E$. Il en résulte que f_μ^* est mesurable. L'inégalité $f_\mu^* \leq \liminf_n (f_n)_\mu^*$ et le lemme de Fatou montrent l'inégalité maximale pour f .

Étape 3. Si f est une fonction positive dans $\mathcal{L}^p(X, m)$, on l'approche par une suite croissante (f_n) de fonctions mesurables positives bornées. On conclut à l'aide de l'étape 2 et du lemme A.2.1. Enfin, pour tout $f \in \mathcal{L}^p(X, m)$, on utilise la décomposition classique de f comme différence de fonctions positives. \square

Remarque A.2.3. — En adaptant la démonstration de la proposition précédente, on voit que celle-ci reste vraie avec des inégalités maximales faibles.

Elle reste également vraie si on dispose seulement des inégalités maximales pour les fonctions appartenant à une partie \mathcal{D} de l'ensemble $C_c(X)$ des fonctions continues à support compact, dense dans $C_c(X)$ pour la topologie de la convergence uniforme. En effet, si $f \in C_c(X)$ et si $(f_n)_{n \geq 1}$ est une suite d'éléments de \mathcal{D} qui converge uniformément vers f , on a immédiatement $f_\mu^* \leq \liminf_n (f_n)_\mu^*$ et le lemme de Fatou montre que l'inégalité maximale s'étend à f .

En général, (X, \mathcal{B}, m) est seulement un espace mesuré standard sur lequel G agit mesurablement. Dans le cas où m est une mesure finie, on se ramène au cas de la proposition précédente à l'aide du théorème suivant.

Théorème A.2.4 (Theorem 5.7, [Var85]). — *Soit G un groupe localement compact et (X, \mathcal{B}) un G -espace mesurable standard. Il existe un espace métrique compact Y , une action continue de G sur Y , un borélien G -invariant $X' \subset Y$ et un isomorphisme G -équivariant Φ du G -espace mesurable X sur le G -espace mesurable X' , muni de la tribu borélienne induite.*

Si m est une mesure finie G -invariante sur X , il reviendra alors au même de travailler sur Y muni de la mesure image de m par Φ .

Observons que lorsque m n'est pas finie, cet argument ne permet pas de se ramener à la situation considérée dans la proposition A.2.2 car l'image de m par Φ n'est pas une mesure de Radon. Il ne semble pas exister de méthode générale pour traiter ce cas.

APPENDICE B

ÉLÉMENTS DE THÉORIE SPECTRALE

Nous rassemblons ici les principaux résultats nécessaires dans l'approche spectrale des théorèmes ergodiques. Pour plus de détails nous renvoyons aux ouvrages [Arv02] et [Rud91, Chap. 11].

On sait que toute matrice carrée M à coefficients complexes qui commute avec sa matrice adjointe M^* est diagonalisable. Plus précisément, si $\lambda_1, \dots, \lambda_k$ sont les valeurs propres de M (sans répétition), et si E_1, \dots, E_k sont les projecteurs orthogonaux sur les sous-espaces propres correspondants, on sait que ces projecteurs commutent deux à deux (autrement dit, les sous-espaces propres sont deux à deux orthogonaux), que $\sum_{i=1}^k E_i = I$ et que $M = \sum_{i=1}^k \lambda_i E_i$. Dans cet appendice, nous allons voir que ce résultat s'étend à toute famille d'opérateurs qui commutent entre eux et avec les adjoints des opérateurs de la famille. Le bon cadre pour énoncer le théorème spectral est celui des C^* -algèbres commutatives.

B.1. C^* -algèbres commutatives

Rappelons d'abord qu'une involution sur une algèbre complexe A est une application $a \mapsto a^*$ telle que $(a^*)^* = a$, $(a+b)^* = a^* + b^*$, $(ab)^* = b^*a^*$ et $(\lambda a)^* = \bar{\lambda}a^*$ quels que soient $a, b \in A$, $\lambda \in \mathbb{C}$. Une algèbre de Banach complexe A est dite *involutive* si elle est munie d'une involution telle que $\|a^*\| = \|a\|$ pour tout $a \in A$. Si de plus on a $\|a^*a\| = \|a\|^2$ pour tout $a \in A$, on dit que A est une C^* -algèbre. Si H est un espace de Hilbert et si $\mathcal{B}(H)$ désigne l'algèbre des opérateurs linéaires bornés sur H , observons que pour tout $T \in \mathcal{B}(H)$ on a $\|T^*T\| = \|T\|^2$. Par conséquent, toute sous-algèbre de $\mathcal{B}(H)$, fermée en norme et stable par passage à l'adjoint, est une C^* -algèbre. Un théorème fondamental de Gelfand-Naimark établit que toute C^* -algèbre peut être représentée de cette façon comme algèbre d'opérateurs.

L'algèbre de Banach $C(X)$ des fonctions continues complexes sur un espace compact X , munie de la norme de la convergence uniforme et de l'involution $f \mapsto \bar{f}$, est

bien sûr une C^* -algèbre commutative. Un autre théorème fondamental de Gelfand dit que toutes les C^* -algèbres commutatives avec unité sont de ce type. L'énoncé précis nécessite l'introduction de la transformation de Gelfand que nous définissons maintenant.

Soit A une algèbre de Banach complexe commutative avec unité. Un *caractère* de A est un homomorphisme non nul de A dans \mathbb{C} . Si $a \in A$, rappelons que le spectre de a (c'est-à-dire l'ensemble des $\lambda \in \mathbb{C}$ tels que $a - \lambda I$ ne soit pas inversible) est égal à l'ensemble des $\chi(a)$ où χ décrit l'ensemble $\text{Sp}(A)$ des caractères de A (voir [Rud91, Theorem 11.5]). On a donc $|\chi(a)| \leq \|a\|$ pour tout χ . Un caractère est donc automatiquement continu et de norme 1. De plus $\text{Sp}(A)$ est fermé dans la boule unité du dual topologique de A , munie de la topologie $*$ -faible, et la topologie induite fait de $\text{Sp}(A)$ un espace compact, appelé *spectre de Gelfand* de A . En particulier, pour tout $a \in A$ la fonction $\hat{a} : \chi \mapsto \hat{a}(\chi) = \chi(a)$ est continue sur $\text{Sp}(A)$. La *transformation de Gelfand* est l'homomorphisme de A dans l'algèbre $C(\text{Sp}(A))$ des fonctions continues complexes sur $\text{Sp}(A)$ qui associe à a la fonction \hat{a} . En général cette transformation n'est ni injective ni surjective, mais on a :

Théorème B.1.1. — *Soit A une C^* -algèbre commutative avec unité. La transformation de Gelfand est un $*$ -isomorphisme isométrique de A sur la C^* -algèbre $C(\text{Sp}(A))$.*

Pour la démonstration, on pourra se reporter à [Rud91, Theorem 11.20]. L'un des ingrédients essentiels est que tout caractère χ d'une C^* -algèbre commutative est *auto-adjoint*, c'est-à-dire que $\overline{\chi(a)} = \chi(a^*)$ pour tout $a \in A$. Le théorème de Stone-Weierstrass entraîne alors la densité de l'image de A dans $C(\text{Sp}(A))$ par la transformation de Gelfand. L'identité $\|a^*a\| = \|a\|^2$ assure que la transformation de Gelfand est isométrique.

B.2. Le théorème spectral

Définition B.2.1. — Une *résolution de l'identité* (aussi appelée *mesure spectrale*) sur un espace compact X est une fonction E sur les boréliens de X , à valeurs dans l'ensemble des projecteurs auto-adjoints d'un espace hilbertien H , telle que :

- (i) $E(\emptyset) = 0$,
- (ii) $E(X) = Id$,
- (iii) pour tous $v, v' \in H$, et toute suite (C_n) de boréliens deux à deux disjoints, la série $\sum \langle E(C_n)v, v' \rangle$ est convergente et sa somme vaut $\langle E(\cup C_n)v, v' \rangle$.

Ainsi, pour tous $v, v' \in H$, l'application $C \mapsto \langle E(C)v, v' \rangle$ est une mesure complexe.

Il résulte facilement de la définition que $E(C_1 \cap C_2) = E(C_1)E(C_2)$ quels que soient les boréliens C_1 et C_2 . Par conséquent les projecteurs de l'image de E commutent deux à deux.

Dans la suite nous considérons une sous- C^* -algèbre commutative B de l'algèbre $\mathcal{B}(H)$ des opérateurs bornés d'un espace de Hilbert H , et nous supposons que B contient l'opérateur identité Id de H . Fixons $v, v' \in H$. L'application $\hat{b} \mapsto \langle bv, v' \rangle$ est une forme linéaire continue sur $C(\text{Sp}(B))$ de norme inférieure ou égale à $\|v\| \|v'\|$. Par le théorème de représentation de Riesz, il existe une mesure complexe $m_{v,v'}$ sur $\text{Sp}(B)$ et une seule telle que pour tout $b \in B$ on ait

$$(B.2.1) \quad \langle bv, v' \rangle = \int \hat{b} \, dm_{v,v'} = \int \chi(b) \, dm_{v,v'}(\chi).$$

Cette mesure est appelée la *mesure spectrale associée* à (v, v') . On notera m_v la mesure positive $m_{v,v}$.

Clairement, pour v' fixé, l'application $v \mapsto m_{v,v'}$ est linéaire. De même les propriétés suivantes se déduisent immédiatement de la définition : quels que soient $v, v' \in H$ et $b \in B$ on a

$$\|m_{v,v'}\|_1 \leq \|v\| \|v'\|, \quad m_{v,v'} = \overline{m_{v',v}} \quad \text{et} \quad m_{bv,v'} = \hat{b} m_{v,v'}.$$

Ici, $\overline{m_{v',v}}$ est la mesure complexe conjuguée de $m_{v',v}$, et pour $f \in C(\text{Sp}(B))$, $f m_{v,v'}$ est la mesure de densité f par rapport à $m_{v,v'}$.

Enfin, si $T \in \mathcal{B}(H)$ commute avec B , on a, pour tous $v, v' \in H$,

$$(B.2.2) \quad m_{Tv,v'} = m_{v,T^*v'}.$$

Théorème B.2.2. — *Avec les notations précédentes, il existe une résolution de l'identité sur $\text{Sp}(B)$ et une seule telle que, pour tous $v, v' \in H$ et tout borélien C de $\text{Sp}(B)$, on ait*

$$\langle E(C)v, v' \rangle = \int \mathbf{1}_C \, dm_{v,v'}.$$

Démonstration. — L'unicité de E est évidente. Démontrons l'existence. Soit C un borélien de $\text{Sp}(B)$. L'application $(v, v') \mapsto \int \mathbf{1}_C \, dm_{v,v'}$ est une forme sesquilinéaire continue sur H , de sorte qu'il existe un unique opérateur $E(C)$ tel que

$$\langle E(C)v, v' \rangle = \int \mathbf{1}_C \, dm_{v,v'}$$

pour tous $v, v' \in H$. Montrons que E est une mesure spectrale sur $\text{Sp}(B)$. Les conditions (i), (ii) et (iii) de la définition B.2.1 sont évidemment réalisées. Il reste à vérifier que pour tout borélien C , l'opérateur $E(C)$ est un projecteur auto-adjoint. L'égalité $E(C)^* = E(C)$ résulte du fait que pour $v, v' \in H$, on a

$$\langle E(C)v, v' \rangle = \int \mathbf{1}_C \, dm_{v,v'} = \overline{\int \mathbf{1}_C \, dm_{v',v}} = \overline{\langle E(C)v', v \rangle}.$$

Assurons-nous aussi que $E(C)^2 = E(C)$. D'abord, pour tout $b \in B$, on voit que

$$\begin{aligned} \int \hat{b} \, d(\mathbf{1}_C m_{v,v'}) &= \int \hat{b} \mathbf{1}_C \, dm_{v,v'} = \int \mathbf{1}_C \, dm_{bv,v'} \\ &= \langle E(C)bv, v' \rangle = \langle bv, E(C)v' \rangle = \int \hat{b} \, dm_{v,E(C)v'}, \end{aligned}$$

d'où $\mathbf{1}_C m_{v,v'} = m_{v,E(C)v'}$. Finalement, pour tous $v, v' \in H$, on obtient

$$\begin{aligned} \langle E(C)^2 v, v' \rangle &= \langle E(C)v, E(C)v' \rangle = \int \mathbf{1}_C \, dm_{v,E(C)v'} \\ &= \int \mathbf{1}_C \, dm_{v,v'} = \langle E(C)v, v' \rangle. \end{aligned}$$

□

Une conséquence importante de (B.2.1) est que

$$\|bv\|^2 = \int |\chi(b)|^2 \, dm_v(\chi)$$

pour tout $b \in B$ et tout $v \in H$. Il suffit en effet de remarquer que $\chi(b^*b) = |\chi(b)|^2$.

Les résultats ci-dessous sont aussi utilisés à plusieurs reprises dans les chapitres précédents.

Lemme B.2.3. — *Gardons les notations ci-dessus.*

- (i) Si $T \in \mathcal{B}(H)$ commute avec B , il commute avec tous les projecteurs $E(C)$.
- (ii) Soit C un borélien de $Sp(B)$. Alors $v \in E(C)H$ si et seulement si

$$m_v(Sp(B) \setminus C) = 0.$$

En particulier, si $\chi' \in Sp(B)$, on a

$$E(\{\chi'\})H = \{v \in H; bv = \chi'(b)v, \forall b \in B\}.$$

- (iii) Soit (C_n) une suite croissante de boréliens dont la réunion vaut C . Alors la suite de projecteurs $(E(C_n))$ converge fortement vers le projecteur $E(C)$. En particulier on a

$$E(C)H = \overline{\cup_n E(C_n)H}.$$

Démonstration. — (i) est une conséquence immédiate de (B.2.2).

Pour (ii), il suffit de remarquer que $E(C)v = v$ si et seulement si $\langle E(C)v, v \rangle = \langle v, v \rangle$, c'est-à-dire si et seulement si $\int \mathbf{1}_C \, dm_v = \int dm_v$.

Supposons maintenant que $C = \{\chi'\}$. Si $bv = \chi'(b)v$ pour tout $b \in B$, alors $\int \hat{b} \, dm_v = \chi'(b)\|v\|^2$ et la mesure m_v est évidemment égale à $\|v\|^2 \delta_{\chi'}$. Réciproquement, si cette condition est réalisée, comme

$$\|bv - \chi'(b)v\|^2 = \int |\chi(b) - \chi'(b)|^2 \, dm_v(\chi),$$

on voit que $bv = \chi'(b)v$ pour tout $b \in B$.

L'assertion (iii) résulte immédiatement de l'égalité

$$\|E(C)v - E(C_n)v\|^2 = \int |\mathbf{1}_C - \mathbf{1}_{C_n}|^2 dm_v$$

et du théorème de convergence dominée. \square

Exemple B.2.4. — Soit $M \in M_n(\mathbb{C})$ une matrice qui commute avec son adjointe, et soit $B \subset M_n(\mathbb{C})$ la sous-algèbre engendrée par M, M^* et la matrice identité I . L'application $\chi \mapsto \chi(M)$ permet d'identifier $\text{Sp}(B)$ avec le spectre $\{\lambda_1, \dots, \lambda_k\}$ de M . Dans ce cas la résolution de l'identité est $E : C \mapsto \sum_{i=1}^k E_i \delta_{\lambda_i}(C)$, où E_i est le projecteur spectral relatif à λ_i .

Exemple B.2.5. — Pour terminer, explicitons la situation particulière rencontrée dans l'approche spectrale des théorèmes ergodiques. Soient G un groupe localement compact et $\mathcal{M}(G)$ l'algèbre de convolution formée par les mesures complexes μ sur G , de variation totale $\|\mu\|_1$ finie. Notons μ^* la mesure $C \mapsto \overline{\mu(C^{-1})}$, où C^{-1} désigne l'image du borélien C de G par l'application $g \mapsto g^{-1}$. L'application $\mu \mapsto \mu^*$ est une involution qui fait de l'algèbre normée $(\mathcal{M}(G), \|\cdot\|_1)$ une algèbre de Banach involutive. Nous avons déjà remarqué dans l'appendice A que toute représentation $\pi : g \mapsto \pi(g)$ dans les opérateurs unitaires d'un espace hilbertien H se prolonge en une représentation de $\mathcal{M}(G)$ dans $\mathcal{B}(H)$.

Considérons alors une sous-algèbre involutive commutative A de $\mathcal{M}(G)$, avec unité et fermée en norme. L'adhérence B de $\pi(A)$ dans $\mathcal{B}(H)$ est une C^* -algèbre commutative. Si χ est un caractère de B , alors $\chi \circ \pi$ est un caractère auto-adjoint de A , et on définit ainsi un homomorphisme injectif de $\text{Sp}(B)$ dans l'espace compact $\text{Sp}_h(A)$ des caractères auto-adjoints de A . La résolution de l'identité E associée à B s'étend en une résolution de l'identité sur $\text{Sp}_h(A)$, nulle en dehors de $\text{Sp}(B)$. Pour tout $v \in H$, la mesure $C \mapsto m_v(C) = \langle E(C)v, v \rangle$ sera aussi vue comme une mesure sur $\text{Sp}_h(A)$, ce qui permet d'écrire, pour $\mu \in A$,

$$(B.2.3) \quad \langle \pi(\mu)v, v \rangle = \int_{\text{Sp}_h(A)} \chi(\mu) dm_v(\chi).$$

et

$$(B.2.4) \quad \|\pi(\mu)v\|^2 = \int_{\text{Sp}_h(A)} |\chi(\mu)|^2 dm_v(\chi).$$

On dit que m_v est la *mesure spectrale associée à v et π* (ou $\pi(a)$ si A est engendrée par l'élément auto-adjoint a).

Dans certains cas l'espace des vecteurs G -invariants, c'est-à-dire des vecteurs $v \in H$ tels que $\pi(g)v = v$ pour tout $g \in G$, peut être déterminé grâce au lemme suivant.

Lemme B.2.6. — Soient $\nu \in \mathcal{M}(G)$ une mesure de probabilité, $v \in H$ et λ un nombre complexe de module 1 tels que $\pi(\nu)v = \lambda v$. Alors pour tout g dans le support de ν , on a $\pi(g)v = \lambda v$.

Démonstration. — On a en particulier $\|v\|^2 = \int \langle \pi(g)v, \lambda v \rangle d\nu(g)$, d'où

$$\int_G (\|v\|^2 - \Re \langle \pi(g)v, \lambda v \rangle) d\nu(g) = 0,$$

puisque ν est une mesure réelle. Comme $\|v\|^2 \geq \Re \langle \pi(g)v, \lambda v \rangle$, grâce à la positivité de ν on voit que l'ouvert $\{g \in G : \|v\|^2 > \Re \langle \pi(g)v, \lambda v \rangle\}$ ne rencontre pas le support de ν . La fonction $g \mapsto \Re \langle \pi(g)v, \lambda v \rangle$ est donc constante égale à $\|v\|^2$ sur le support de ν et par conséquent $\pi(g)v = \lambda v$ sur cet ensemble. \square

Corollaire B.2.7. — *Supposons l'existence d'une mesure positive $\nu \in A$ telle que le sous-groupe fermé engendré par son support soit G tout entier. Soit $\chi_0 \in \text{Sp}_h(A)$ le caractère trivial $\mu \mapsto \mu(G)$. Alors $E(\{\chi_0\})H$ est l'espace des vecteurs G -invariants.*

Démonstration. — D'après le lemme B.2.3, on a

$$E(\{\chi_0\})H = \{v \in H : \pi(\mu)v = \mu(G)v, \forall \mu \in A\}.$$

Si v est un vecteur G -invariant et si $\mu \in A$, on a bien sûr

$$\pi(\mu)v = \int \pi(g)v d\mu(g) = \mu(G)v.$$

Réciproquement, soit $v \in E(\{\chi_0\})H$. Grâce au lemme précédent on obtient $\pi(g)v = v$ si g appartient au support de ν . Pour conclure il suffit de remarquer que l'ensemble des g tels que $\pi(g)v = v$ est un sous-groupe fermé de G . \square

APPENDICE C

THÉORÈME DE ROTA POUR UNE SUITE D'OPÉRATEURS DE MARKOV

Soit (X, \mathcal{B}, m) un espace mesuré standard, avec $m(X) = 1$.

Définition C.1.8. — On dit qu'un opérateur $P : L^1(X, m) \rightarrow L^1(X, m)$

- (a) est une *contraction* $L^1 - L^\infty$ s'il préserve $L^\infty(X, m)$ avec $\|P\|_{L^1 \rightarrow L^1} \leq 1$ et $\|P\|_{L^\infty \rightarrow L^\infty} \leq 1$;
- (b) est *positif* s'il préserve le cône des fonctions positives;
- (c) *préserve* m si $\int_X P(\varphi) dm = \int_X \varphi dm$ pour tout $\varphi \in L^1(X, m)$.

Dans cet appendice, un opérateur possédant ces trois propriétés sera appelé *opérateur de Markov*.

Un tel opérateur P étant une contraction positive de $L^\infty(X, m)$ vérifie $P(\mathbf{1}_X) \geq 0$ et $\|P(\mathbf{1}_X)\|_\infty \leq 1$, ce qui, avec la condition (c), entraîne l'égalité $P(\mathbf{1}_X) = \mathbf{1}_X$. On voit aisément que l'opérateur adjoint $P^* : L^\infty(X, m) \rightarrow L^\infty(X, m)$ est une contraction positive de $L^\infty(X, m)$ qui laisse fixe la fonction $\mathbf{1}_X$. De plus, P^* se prolonge en une contraction positive de $L^1(X, m)$. Ce prolongement, qu'on notera encore P^* , est à son tour un opérateur de Markov et on a l'égalité $P^{**} = P$.

Rappelons aussi qu'on peut associer à un opérateur de Markov P un *noyau markovien*, c'est-à-dire une application $(x, B) \mapsto N(x, B)$ de $X \times \mathcal{B}$ dans $[0, 1]$ telle que pour tout $x \in X$, $N(x, \cdot)$ est une probabilité et, pour tout $B \in \mathcal{B}$, $N(\cdot, B)$ est une fonction mesurable, de telle sorte qu'on a

$$P(f)(x) = \int_X f(y)N(x, dy)$$

m -presque sûrement pour tout $f \in L^1(X, m)$ (voir [Nev70, Section V.4]).

Nous allons voir que pour toute fonction $f \in L \log L(X, m)$, la suite $P^n P^{*n}$ converge m -presque sûrement. C'est un cas particulier du théorème suivant, dû à Rota [Rot62] :

Théorème C.1.9. — Soit $(P_n)_{n \geq 1}$ une suite d'opérateurs de Markov sur (X, \mathcal{B}, m) . Alors pour toute fonction $f \in L^1(X, m)$, la suite $P_1 \dots P_n P_n^* \dots P_1^* f$ converge dans $L^1(X, m)$. Si de plus $f \in L \log L(X, m)$, alors elle converge presque sûrement.

Démonstration. — On considère l'espace produit $\Omega = X^{\mathbb{N}}$ muni de la tribu produit $\mathcal{B}^{\mathbb{N}}$, et on note $X_n : \Omega \rightarrow X$ les applications coordonnées. Pour tout $n \geq 0$, on note $\mathcal{B}_n = \sigma(X_0, \dots, X_n)$ la tribu des événements antérieurs à n et $\mathcal{C}_n = \sigma(X_n, X_{n+1}, \dots)$ la tribu des événements postérieurs à n .

La donnée de la mesure m sur X et de la suite (P_n) permet de construire une probabilité \mathbb{P} sur Ω vérifiant, pour toute fonction \mathcal{B}_n -mesurable F , la relation

$$\int F(\omega) \, d\mathbb{P}(\omega) = \int F(x_0, x_1, \dots, x_n) \, dm(x_0) N_1(x_0, dx_1) N_2(x_1, dx_2) \cdots N_n(x_{n-1}, dx_n),$$

où $(N_n)_{n \geq 1}$ est la suite des noyaux associés aux opérateurs P_n . Sous la probabilité \mathbb{P} , le processus (X_n) est une chaîne de Markov, inhomogène en temps.

Si on associe à toute fonction mesurable f définie sur X la variable aléatoire \tilde{f} définie sur Ω par

$$\tilde{f}(x_0, x_1, x_2, \dots) = f(x_0)$$

on définit un isomorphisme de $L^1(X, \mathcal{B}, m)$ sur $L^1(\Omega, \mathcal{B}_0, \mathbb{P})$ qui permet d'identifier l'opérateur d'espérance conditionnelle $\mathbb{E}^{\mathcal{B}_0}$ avec un opérateur de $L^1(\Omega, \mathcal{B}^{\mathbb{N}}, \mathbb{P})$ sur $L^1(X, \mathcal{B}, m)$. Nous noterons $\mathbb{E}^{\mathcal{B}}$ cet opérateur. Nous noterons aussi, pour alléger l'écriture, \mathbb{E}_n l'opérateur d'espérance conditionnelle $\mathbb{E}^{\mathcal{C}_n}$ par rapport à la tribu \mathcal{C}_n .

Le point de départ de la démonstration est l'égalité

$$(C.1.1) \quad P_1 \dots P_n P_n^* \dots P_1^* f = \mathbb{E}^{\mathcal{B}} \mathbb{E}_n \tilde{f}.$$

On montrera ensuite que $\mathbb{E}^{\mathcal{B}} \mathbb{E}_n \tilde{f}$ converge m -p.p. pour tout $f \in L \log L(X, m)$ en utilisant le théorème de convergence des martingales inverses.

Commençons par montrer (C.1.1), ce qui se fera en deux temps.

Lemme C.1.10. — Pour $f \in L^1(X, m)$ et $n \in \mathbb{N}^*$, on a

$$\mathbb{E}_n \tilde{f} = (P_n^* \dots P_1^* f) \circ X_n.$$

Démonstration. — Puisque la variable aléatoire $(P_n^* \dots P_1^* f) \circ X_n$ est \mathcal{C}_n mesurable, il suffit de démontrer qu'on a

$$\int_A \tilde{f} \, d\mathbb{P} = \int_A (P_n^* \dots P_1^* f) \circ X_n \, d\mathbb{P}$$

pour tout événement A de la forme

$$A = \{X_n \in B_n, X_{n+1} \in B_{n+1}, \dots, X_{n+p} \in B_{n+p}\}$$

avec $p \geq 1$ et des ensembles B_n, \dots, B_{n+p} appartenant à \mathcal{B} .

Par définition de \mathbb{P} , on a

$$\begin{aligned} \int_A \tilde{f}(\omega) \, d\mathbb{P}(\omega) &= \int_{\Omega} f(x_0) \mathbf{1}_{B_n}(x_n) \mathbf{1}_{B_{n+1}}(x_{n+1}) \cdots \mathbf{1}_{B_{n+p}}(x_{n+p}) \, d\mathbb{P}(x_0, \dots, x_n \dots) \\ &= \int_X f(x) \left(P_1 \dots P_n (\mathbf{1}_{B_n} P_{n+1} (\mathbf{1}_{B_{n+1}} \dots P_{n+p} \mathbf{1}_{B_{n+p}})) \right) (x) \, dm(x) \\ &= \int_X (P_n^* \dots P_1^* f)(x) (\mathbf{1}_{B_n} P_{n+1} (\mathbf{1}_{B_{n+1}} \dots P_{n+p} \mathbf{1}_{B_{n+p}})) (x) \, dm(x). \end{aligned}$$

L'invariance de la mesure m sous l'action des opérateurs P_n permet de remplacer cette dernière intégrale par

$$\int_X P_1 \dots P_n \left((P_n^* \dots P_1^* f) \mathbf{1}_{B_n} P_{n+1}(\dots) \right) (x) \, dm(x).$$

En écrivant chacun des opérateurs P_1, \dots, P_{n+p} sous la forme d'une intégrale à noyau cette intégrale devient

$$\begin{aligned} &\int_{X^{n+1}} (P_n^* \dots P_1^* f)(x_n) (\mathbf{1}_{B_n} P_{n+1}(\dots))(x_n) \, dm(x) N_1(x, dx_1) \dots N_n(x_{n-1}, dx_n) \\ &= \int_{X^n \times B_n \times \dots \times B_{n+p}} (P_n^* \dots P_1^* f)(x_n) \, dm(x) N_1(x, dx_1) \dots N_{n+p}(x_{n+p-1}, dx_{n+p}) \\ &= \int_A (P_n^* \dots P_1^* f) \circ X_n \, d\mathbb{P}. \quad \square \end{aligned}$$

Dans un deuxième temps, on vérifie le résultat suivant :

Lemme C.1.11. — *Pour une fonction $H \in L^1(\Omega, \mathbb{P})$ de la forme $H = h \circ X_n$ avec $h \in L^1(X, m)$, on a $\mathbb{E}^{\mathcal{B}} H = P_1 \dots P_n h$.*

Démonstration. — Il suffit de vérifier que pour tout événement B de \mathcal{B}_0 , on a

$$\int_B (P_1 \dots P_n h) \, d\mathbb{P} = \int_B h \circ X_n \, d\mathbb{P}$$

ce qui résulte de la définition même de \mathbb{P} . □

Ces deux lemmes démontrent donc l'égalité (C.1.1). Observons maintenant que la suite de tribus \mathcal{C}_n est décroissante, ce qui permet d'utiliser le théorème de convergence des martingales inverses. Pour sa démonstration nous renvoyons à [Nev72, Cor. V-3-12].

Théorème C.1.12 (Convergence des martingales inverses)

Soit $(\mathcal{C}_n, n \geq 0)$ une suite décroissante de sous-tribus dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Si $F \in L^1(\Omega, \mathbb{P})$, la suite $\mathbb{E}^{\mathcal{C}_n}(F)$ converge \mathbb{P} -p.p. et dans L^1 vers $\mathbb{E}^{\mathcal{C}_\infty}(F)$ où $\mathcal{C}_\infty = \bigcap_{n=1}^\infty \mathcal{C}_n$.

Puisque l'opérateur $\mathbb{E}^{\mathcal{B}}$ est une contraction de $L^1(\Omega, \mathbb{P})$ dans $L^1(X, m)$, on déduit du théorème C.1.12 que si f est intégrable, la suite $\mathbb{E}^{\mathcal{B}} \mathbb{E}_n \tilde{f}$ converge dans $L^1(\Omega, \mathbb{P})$. Pour obtenir la convergence presque sûre il suffit, puisque l'opérateur $\mathbb{E}^{\mathcal{B}}$ est positif, de pouvoir dominer la suite $\mathbb{E}_n \tilde{f}$ par une variable aléatoire intégrable. C'est ici que

l'hypothèse plus forte $f \in L \log L(X, m)$ est utile. Elle assure que la variable aléatoire $Z_0 = \mathbb{E}_0 \tilde{f} = \tilde{f}$ appartient à $L \log L(\Omega, \mathbb{P})$, et donc que $\sup_n |\mathbb{E}_n \tilde{f}|$ est intégrable, grâce au résultat suivant :

Proposition C.1.13. — *Pour toute martingale inverse $(Z_n)_{n \in \mathbb{N}}$ vérifiant la condition $\mathbb{E}(|Z_0| \log^+ |Z_0|) < \infty$, la variable aléatoire réelle $\sup_{n \in \mathbb{N}} |Z_n|$ est intégrable.*

Ce dernier fait permet donc de déduire la convergence presque sûre de la suite $\mathbb{E}^{\mathcal{B}} \mathbb{E}_n \tilde{f}$ de la convergence presque sûre de la suite $\mathbb{E}_n \tilde{f}$ et de conclure ainsi la démonstration du théorème de Rota. \square

Démonstration de la proposition C.1.13. — Rappelons tout d'abord qu'une martingale (M_n) vérifie l'inégalité maximale de Doob

$$a \mathbb{P}(\sup_{m \leq n} |M_m| > a) \leq \int_{\{\sup_{m \leq n} |M_m| > a\}} |M_n| \, d\mathbb{P}$$

pour tout $n \in \mathbb{N}$ et tout $a \in \mathbb{R}_+$.

On montre alors le résultat suivant ([Nev72, pp 70-71]).

Proposition C.1.14 (Inégalité maximale $L \log L$). — *Soient $n > 0$ et une martingale $(M_m)_{0 \leq m \leq n}$. Alors, pour tout $p > 1$ on a*

$$(1 - \frac{1}{e}) \mathbb{E}(\sup_{m \leq n} |M_m|) \leq 1 + \mathbb{E}(|M_n| \log^+ |M_n|).$$

Démonstration. — Partons de la majoration

$$\mathbb{E}(\sup_{m \leq n} |M_m|) = \int_0^\infty \mathbb{P}(\sup_{m \leq n} |M_m| > a) \, da \leq 1 + \int_1^\infty \mathbb{P}(\sup_{m \leq n} |M_m| > a) \, da.$$

L'inégalité maximale de Doob implique alors

$$\int_1^\infty \mathbb{P}(\sup_{m \leq n} |M_m| > a) \, da \leq \int_1^\infty \frac{da}{a} \int_{\{\sup_{m \leq n} |M_m| > a\}} |M_n| \, d\mathbb{P}$$

Or

$$\begin{aligned} \int_1^\infty \frac{da}{a} \int_{\{\sup_{m \leq n} |M_m| > a\}} |M_n| \, d\mathbb{P} &= \mathbb{E}[|M_n| \log^+(\sup_{m \leq n} |M_m|)] \\ &\leq \mathbb{E}(|M_n| \log^+ |M_n|) + \frac{1}{e} \mathbb{E}(\sup_{m \leq n} |M_m|), \end{aligned}$$

grâce à l'inégalité élémentaire $a \log^+ b \leq a \log^+ a + \frac{b}{e}$ vérifiée pour a et b dans \mathbb{R}_+ . On trouve donc

$$\mathbb{E}(\sup_{m \leq n} |M_m|) \leq 1 + \mathbb{E}(|M_n| \log^+ |M_n|) + \frac{1}{e} \mathbb{E}(\sup_{m \leq n} |M_m|),$$

ce qui donne l'inégalité annoncée. \square

Pour démontrer la proposition C.1.13, il suffit d'appliquer la proposition C.1.14 à la martingale $(Z_n, Z_{n-1}, \dots, Z_0)$ relativement à la filtration $\mathcal{C}_n \subset \mathcal{C}_{n-1} \subset \dots \subset \mathcal{C}_0$. On obtient alors

$$\left(1 - \frac{1}{e}\right) \mathbb{E}\left(\sup_{m \leq n} |Z_m|\right) \leq 1 + \mathbb{E}(|Z_0| \log^+ |Z_0|).$$

Ceci étant vrai pour tout $n \geq 0$, la variable $\sup_{m \geq 0} |Z_m|$ est intégrable, ce qui montre la proposition C.1.13. \square

Remarque C.1.15. — Pour un opérateur de Markov auto-adjoint P agissant sur $L^1(X, m)$, le théorème de Rota donne la convergence presque sûre de la suite $P^{2n} f$ pour toute fonction $f \in L \log L(X, m)$. Pour $f \in L^p(X, m)$, $p > 1$, ce dernier résultat a été démontré par Stein dans [Ste61]. Un contre-exemple dû à Ornstein montre que le théorème de Stein n'est pas vrai en général pour des fonctions $f \in L^1(X, m)$ [Orn69].

APPENDICE D

L'ESPACE HYPERBOLIQUE

Il existe (à isométrie près) une unique variété riemannienne complète, connexe, simplement connexe, de dimension $d + 1$ et de courbure sectionnelle constante égale à -1 . C'est le revêtement universel de toutes les variétés riemanniennes connexes de dimension $d + 1$ et de courbure sectionnelle constante égale à -1 . On appelle *espace hyperbolique* de dimension $d + 1$ n'importe lequel de ses modèles.

Dans cette annexe, nous allons en présenter différentes descriptions, puis nous décrirons son groupe d'isométries et nous finirons par des énoncés utiles dans les chapitres 4 et 7.

Pour plus de détails sur les résultats de cet appendice, nous renvoyons à [Rat94], où l'on trouvera aussi de nombreuses références ainsi que des notes historiques.

D.1. Les différents modèles

La référence de cette section est [Rat94, chap. 3 et 4]. Les deux premiers modèles ont été étudiés indépendamment par Beltrami (1868) et Poincaré (1882), et sont souvent attribués à ce dernier. Le *modèle du demi-espace* de l'espace hyperbolique de dimension $d + 1$ ($d \geq 1$) est l'ensemble $\mathbb{H}^{d+1} = \mathbb{R}^d \times \mathbb{R}_+^*$ muni de la métrique $ds^2 = \frac{dx^2 + dy^2}{y^2}$, où l'on note $z = (x, y) \in \mathbb{R}^d \times \mathbb{R}_+^*$ un élément de \mathbb{H}^{d+1} . Cette métrique est *hyperbolique*, *i.e.* de courbure sectionnelle constante égale à -1 . Si v est un vecteur tangent à \mathbb{H}^{d+1} au point $z = (x, y)$, nous noterons $\|v\|_z$ sa norme hyperbolique, et $|v|$ sa norme euclidienne. Alors on a $\|v\|_z = \frac{|v|}{y}$. Cette métrique permet de calculer la longueur d'un chemin $c : [a, b] \rightarrow \mathbb{H}^{d+1}$ de classe C^1 par

$$L(c) = \int_a^b \|\dot{c}(t)\| dt.$$

On définit alors une distance d sur \mathbb{H}^{d+1} en posant, pour tous points z et z' de \mathbb{H}^{d+1} ,

$$d(z, z') = \inf L(c),$$

où c parcourt l'ensemble des courbes de classe C^1 joignant z à z' .

Rappelons que l'*inversion* par rapport à la sphère $S(a, r)$, $a \in \mathbb{R}^{d+1}$, $r > 0$, de \mathbb{R}^{d+1} est l'application de \mathbb{R}^{d+1} dans \mathbb{R}^{d+1} définie par

$$x \in \mathbb{R}^{d+1} \mapsto a + \frac{r^2}{|x - a|^2} (x - a).$$

Introduisons le point $o \in \mathbb{H}^{d+1}$ de coordonnées $(0_{\mathbb{R}^d}, 1)$. La composée de la symétrie par rapport à l'hyperplan (de \mathbb{R}^{d+1}) d'équation $x_{d+1} = 0$ et de l'inversion par rapport à la sphère $S(o, \sqrt{2})$ induit un difféomorphisme de \mathbb{H}^{d+1} sur la boule unité \mathbb{B}^{d+1} de \mathbb{R}^{d+1} . Si \mathbb{B}^{d+1} est munie de la métrique $ds^2 = 4 \frac{dx^2}{(1 - |x|^2)^2}$, ce difféomorphisme est de plus une isométrie, et on parlera du *modèle de la boule unité* de l'espace hyperbolique. Dans \mathbb{B}^{d+1} , nous noterons encore o le centre de la boule, qui est l'image du point $o = (0_{\mathbb{R}^d}, 1) \in \mathbb{H}^{d+1}$ par l'application ci-dessus.

Pour finir, le *modèle de Minkowski*, ou *modèle de l'hyperboloïde* de l'espace hyperbolique le décrit comme l'ensemble

$$\mathcal{H}^{d+1} = \{(x_0, \dots, x_{d+1}) \in \mathbb{R}^{d+2}; q(x) = -1 \text{ et } x_0 > 0\},$$

où q est la forme quadratique définie par $q(x) = -x_0^2 + x_1^2 + \dots + x_{d+1}^2$. La métrique riemannienne est alors donnée par la restriction de q à l'espace tangent de \mathcal{H}^{d+1} , et le point o est le point de coordonnées $(1, 0, \dots, 0) \in \mathcal{H}^{d+1}$.

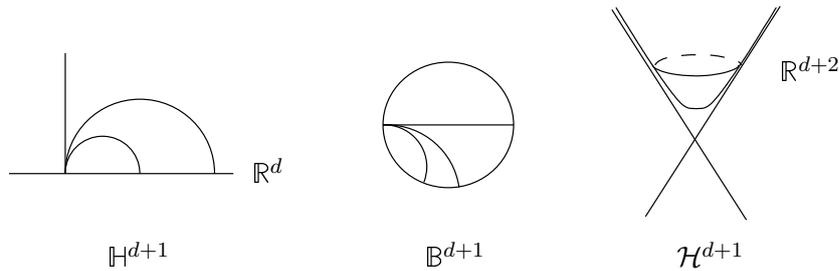


FIGURE 8. Les différents modèles de l'espace hyperbolique

Nous avons besoin de la forme explicite du volume riemannien dans l'espace hyperbolique (voir chapitre 4). La métrique g au point $x = (x_1, \dots, x_{d+1})$ est donnée

par une matrice $g(x) = (g_{ij}(x))_{1 \leq i, j \leq d+1}$. L'élément de volume riemannien est alors défini par

$$\mathrm{dvol}(x) = \sqrt{\det g(x)} \, dx.$$

Dans le modèle du demi-espace, la matrice $g(x, y)$ au point $(x, y) \in \mathbb{R}^d \times \mathbb{R}_+^*$ est donnée par $g(x, y) = \frac{1}{y^2} I_{d+1}$, de sorte que l'élément de volume hyperbolique vaut $\mathrm{dvol}(x, y) = \frac{1}{y^{d+1}} dx dy$. Dans le modèle de la boule unité, la matrice $g(x)$ au point $x \in \mathbb{B}^{d+1}$ est $g(x) = \frac{4}{(1 - |x|^2)^2} I_{d+1}$, et l'élément de volume hyperbolique vaut

$$\mathrm{dvol}(x) = 2^{d+1} \frac{dx}{(1 - |x|^2)^{d+1}}.$$

Intuitivement, cette formule correspond au fait suivant : l'élément de longueur dans la direction de la i -ème coordonnée vaut $\frac{2dx_i}{1 - |x|^2}$, et quand on fait le produit des $d+1$ éléments de longueur, on retrouve le volume ci-dessus.

Lemme D.1.1. — Dans \mathbb{B}^{d+1} , l'élément de volume en coordonnées polaires vaut

$$(D.1.1) \quad \mathrm{dvol}(\rho, \theta) = (\mathrm{sh} \rho)^d \, d\rho \, d\sigma(\theta),$$

où $\theta \in \mathbb{S}^d$, $\rho \in \mathbb{R}_+^*$ est la distance (hyperbolique) à l'origine o , et $d\sigma$ désigne l'élément de surface sur la sphère unité de \mathbb{R}^{d+1} .

Démonstration. — Notons $r = |x|$ la distance euclidienne de x à l'origine, et $\rho = d(o, x)$ la distance hyperbolique de x à l'origine. Par définition de la distance hyperbolique, on a

$$\rho = \int_0^r \frac{2}{(1-t^2)} dt.$$

d'où $r = \mathrm{th} \frac{\rho}{2}$, par un calcul élémentaire.

D'autre part, en coordonnées polaires (r, θ) , le volume euclidien s'écrit $dx_1 \dots dx_{d+1} = r^d dr d\sigma(\theta)$. L'élément de volume hyperbolique s'écrit donc

$$\mathrm{dvol}(r, \theta) = \frac{2^{d+1} r^d dr d\sigma(\theta)}{(1-r^2)^{d+1}}.$$

Le résultat voulu s'obtient en remplaçant r par $\mathrm{th} \frac{\rho}{2}$. □

D.2. Isométries de \mathbb{H}^{d+1}

La référence de ce paragraphe est [Rat94, chap. 4 et 5]. Toute isométrie de \mathbb{H}^{d+1} s'obtient comme *extension de Poincaré* d'une transformation conforme de \mathbb{R}^d . Rappelons que le groupe des transformations conformes de \mathbb{R}^d est engendré par les réflexions par rapport aux hyperplans de \mathbb{R}^d et les inversions par rapport aux sphères de \mathbb{R}^d .

Identifions maintenant \mathbb{R}^d à $\mathbb{R}^d \times \{0\}$ dans \mathbb{R}^{d+1} . L'extension de Poincaré d'une réflexion par rapport à un hyperplan \mathcal{P} de \mathbb{R}^d est la réflexion de \mathbb{R}^{d+1} par rapport à l'hyperplan vertical $\tilde{\mathcal{P}}$ de \mathbb{R}^{d+1} dont l'intersection avec \mathbb{R}^d est \mathcal{P} .

Si $a \in \mathbb{R}^d$, notons $\tilde{a} = (a, 0) \in \mathbb{R}^{d+1}$. L'extension de Poincaré à \mathbb{R}^{d+1} de l'inversion $I_{a,r}$ de \mathbb{R}^d est l'inversion $I_{\tilde{a},r}$ de \mathbb{R}^{d+1} par rapport à la sphère $S(\tilde{a}, r)$. Une vérification immédiate montre que l'extension de Poincaré d'une transformation conforme de \mathbb{R}^d préserve le demi-espace \mathbb{H}^{d+1} . Par restriction à \mathbb{H}^{d+1} , on obtient ainsi toutes les isométries de cet espace :

Lemme D.2.1. — *Le groupe \tilde{G} des isométries de l'espace hyperbolique \mathbb{H}^{d+1} est le groupe des extensions de Poincaré des transformations conformes de \mathbb{R}^d .*

Notons G le sous-groupe des isométries directes de \mathbb{H}^{d+1} . Avec la description géométrique ci-dessus des isométries, on démontre le résultat suivant :

Lemme D.2.2. — *Le groupe G agit transitivement sur \mathbb{H}^{d+1} , c'est-à-dire que pour tous $x, y \in \mathbb{H}^{d+1}$, il existe $g \in G$, tel que $gx = y$. De plus, il agit simplement transitivement sur le fibré unitaire tangent $T^1\mathbb{H}^{d+1}$, i.e. pour tous v, w de $T^1\mathbb{H}^{d+1}$, il existe un unique $g \in G$ tel que $gv = w$.*

Dans le modèle de Minkowski de l'espace hyperbolique, on voit immédiatement que le groupe \tilde{G} s'identifie au sous-groupe (d'indice 2) de $O(d+1, 1)$ formé des isométries qui laissent stable la nappe supérieure \mathcal{H}^{d+1} de l'hyperboloïde à deux nappes d'équation $q(x) = -1$, où $O(d+1, 1)$ est le groupe des transformations préservant la forme quadratique q . Le groupe G des isométries qui préservent l'orientation de \mathcal{H}^{d+1} s'identifie à la composante connexe $SO^o(d+1, 1)$ de l'identité dans $SO(d+1, 1)$.

Le modèle du demi-espace est toutefois plus agréable pour obtenir la décomposition d'Iwasawa de $G \simeq SO^o(d+1, 1)$. Le stabilisateur de o dans G est $K \simeq SO(d+1)$. Remarquons aussi que G contient le *groupe affine* $S_d = \{(b, a) \in \mathbb{R}^d \times \mathbb{R}_+^*\}$ (introduit au chapitre 4), qui agit simplement transitivement sur \mathbb{H}^{d+1} par

$$(b, a)(x, y) = (ax + b, ay).$$

On a donc

$$\mathbb{H}^{d+1} \simeq G/K \simeq S_d o$$

Soient $N = \{(b, 0); b \in \mathbb{R}^d\}$ et $A = \{(0, a); a \in \mathbb{R}_+^*\}$. Le groupe affine est le produit semi-direct $S_d = N \rtimes A$. Pour tout $g \in G$, il existe donc un unique élément $na \in NA$ tel que $go = (na)o$, d'où $g = nak$ avec $k \in K$. On obtient ainsi la décomposition d'Iwasawa $G \simeq N \times A \times K$.

D.3. Quelques formules

Le résultat ci-dessous (voir [Rat94, Th. 4.6.1]) donne une formule explicite pour la distance entre deux points z et z' de \mathbb{H}^{d+1} .

Proposition D.3.1. — *Pour tous points $z = (x, y)$ et $z' = (x', y')$ dans \mathbb{H}^{d+1} , on a*

$$\operatorname{ch} d(z, z') = 1 + \frac{|z - z'|^2}{2yy'}.$$

Démonstration. — En faisant agir les isométries de \mathbb{H}^{d+1} , on se ramène facilement à la situation où $x = x'$. En utilisant la définition de la métrique hyperbolique et le fait que les demi-droites verticales sont des géodésiques de \mathbb{H}^{d+1} , on obtient immédiatement $d(z, z') = \ln\left(\frac{y'}{y}\right)$. D'autre part, on a $|z - z'| = |y - y'|$ d'où

$$(D.3.1) \quad 1 + \frac{|z - z'|^2}{2yy'} = \frac{y^2 + y'^2}{2yy'} = \frac{1}{2} \left(\frac{y}{y'} + \frac{y'}{y} \right) = \operatorname{ch} d(z, z').$$

Ceci conclut la démonstration. □

Nous en déduisons le résultat suivant, également utilisé au chapitre 4.

Proposition D.3.2. — *Soit g une isométrie de \mathbb{H}^{d+1} de la forme $g = a_t n_x$, où $a_t = (0, e^t)$ et $n_x = (x, 1)$ avec $x \in \mathbb{R}^d$. On a alors*

$$(D.3.2) \quad \operatorname{ch} d(o, go) = \operatorname{ch} t + \frac{e^t}{2} |x|_d^2,$$

où $|x|_d$ désigne la norme euclidienne de x dans \mathbb{R}^d .

Démonstration. — Le point go a pour coordonnées $go = (e^t x, e^t)$ dans $\mathbb{H}^{d+1} = \mathbb{R}^d \times \mathbb{R}_+^*$. On en déduit que le carré de la distance euclidienne entre o et go vaut $|o - go|^2 = (e^t |x|_d)^2 + (e^t - 1)^2$. Le résultat voulu découle alors de la proposition D.3.1. □

BIBLIOGRAPHIE

- [Aar97] J. AARONSON – *An introduction to infinite ergodic theory*, Mathematical Surveys and Monographs, vol. 50, American Mathematical Society, Providence, RI, 1997.
- [Ank90] J.-P. ANKER – “ L_p Fourier multipliers on Riemannian symmetric spaces of the noncompact type”, *Ann. of Math. (2)* **132** (1990), no. 3, p. 597–628.
- [Ank92] ———, “Sharp estimates for some functions of the Laplacian on noncompact symmetric spaces”, *Duke Math. J.* **65** (1992), no. 2, p. 257–297.
- [ADY96] J.-P. ANKER, E. DAMEK & C. YACOUB – “Spherical analysis on harmonic an groups”, *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **23** (1996), no. 4, p. 643–679.
- [AJ99] J.-P. ANKER & L. JI – “Heat kernel and Green function estimates on noncompact symmetric spaces”, *Geom. Funct. Anal.* **9** (1999), no. 6, p. 1035–1091.
- [AK63] V. I. ARNOL'D & A. L. KRYLOV – “Uniform distribution of points on a sphere and certain ergodic properties of solutions of linear ordinary differential equations in a complex domain”, *Dokl. Akad. Nauk SSSR* **148** (1963), p. 9–12.
- [Arv02] W. ARVESON – *A short course on spectral theory*, Graduate Texts in Mathematics, vol. 209, Springer-Verlag, New York, 2002.
- [Ban26] S. BANACH – “Sur la convergence presque partout des fonctionnelles linéaires”, *Bull. Sci. Math.* **50** (1926), p. 27–32, 36–43.
- [Bar93] J. BARRIONUEVO – “Estimates for some Kakeya-type maximal operators”, *Trans. Amer. Math. Soc.* **335** (1993), no. 2, p. 667–682.
- [BG06] B. BEKKA & Y. GUIVARC'H – “A spectral gap property for random walks under unitary representations”, *Geom. Dedicata* **118** (2006), p. 141–155.

- [BdlHV08] B. BEKKA, P. DE LA HARPE & A. VALETTE – *Kazhdan’s property (T)*, New Mathematical Monographs, vol. 11, Cambridge University Press, Cambridge, 2008.
- [BL85] A. BELLOW & V. LOSERT – “The weighted pointwise ergodic theorem and the individual ergodic theorem along subsequences”, *Trans. Amer. Math. Soc.* **288** (1985), no. 1, p. 307–345.
- [Bes45] A. S. BESICOVITCH – “A general form of the covering principle and relative differentiation of additive functions”, *Proc. Cambridge Philos. Soc.* **41** (1945), p. 103–110.
- [Bes46] ———, “A general form of the covering principle and relative differentiation of additive functions. II”, *Proc. Cambridge Philos. Soc.* **42** (1946), p. 1–10.
- [Bes47] ———, “Corrigenda to the paper “A general form of the covering principle and relative differentiation of additive functions. II.””, *Proc. Cambridge Philos. Soc.* **43** (1947), p. 590.
- [Bir31] G. D. BIRKHOFF – “Proof of the ergodic theorem”, *Proc. Nat. Acad. Sci. USA* **17** (1931), p. 656–660.
- [Bou86] J. BOURGAIN – “Averages in the plane over convex curves and maximal operators”, *J. Analyse Math.* **47** (1986), p. 69–85.
- [Bou89] ———, “Pointwise ergodic theorems for arithmetic sets”, *Inst. Hautes Études Sci. Publ. Math.* **69** (1989), p. 5–45, With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein.
- [Bre07] E. BREUILLARD – “Geometry of locally compact groups of polynomial growth and shape of large balls”, Tech. report, arXiv:math.GR/0704.0095v1, 2007.
- [BH99] M. R. BRIDSON & A. HAEFLIGER – *Metric spaces of non-positive curvature*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 319, Springer-Verlag, Berlin, 1999.
- [Buf00] A. I. BUFETOV – “Operator ergodic theorems for actions of free semi-groups and groups”, *Funct. Anal. Appl.* **34** (2000), no. 4, p. 239–251.
- [Buf02] ———, “Convergence of spherical averages for actions of free groups”, *Ann. of Math. (2)* **155** (2002), no. 3, p. 929–944.
- [BF34] H. BUSEMANN & W. FELLER – “Zur Differentiation der lebesgueschen Integrale”, *Fundamenta Mathematicae* **22** (1934), p. 226–256.
- [Cal53] A.-P. CALDERÓN – “A general ergodic theorem”, *Ann. of Math. (2)* **58** (1953), p. 182–191.

- [Cal68] ———, “Ergodic theory and translation-invariant operators”, *Proc. Nat. Acad. Sci. U.S.A.* **59** (1968), p. 349–353.
- [Che05] I. CHEREDNIK – *Double affine Hecke algebras*, London Mathematical Society Lecture Note Series, vol. 319, Cambridge University Press, Cambridge, 2005.
- [CS74] J. L. CLERC & E. M. STEIN – “ L^p -multipliers for noncompact symmetric spaces”, *Proc. Nat. Acad. Sci. U.S.A.* **71** (1974), p. 3911–3912.
- [CW76] R. R. COIFMAN & G. WEISS – *Transference methods in analysis*, American Mathematical Society, Providence, R.I., 1976, Conference Board of the Mathematical Sciences Regional Conference Series in Mathematics, No. 31.
- [Cór79] A. CÓRDOBA – “Maximal functions, covering lemmas and Fourier multipliers”, in *Harmonic analysis in Euclidean spaces (Proc. Sympos. Pure Math., Williams Coll., Williamstown, Mass., 1978), Part 1*, Proc. Sympos. Pure Math., XXXV, Part, Amer. Math. Soc., Providence, R.I., 1979, p. 29–50.
- [CF75] A. CÓRDOBA & R. FEFFERMAN – “A geometric proof of the strong maximal theorem”, *Ann. of Math. (2)* **102** (1975), no. 1, p. 95–100.
- [CFS82] I. P. CORNFELD, S. V. FOMIN & Y. G. SINAIĬ – *Ergodic theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 245, Springer-Verlag, New York, 1982, Translated from the Russian by A. B. Sosinskiĭ.
- [Cow78] M. COWLING – “The Kunze-Stein phenomenon”, *Ann. Math. (2)* **107** (1978), no. 2, p. 209–234.
- [Cow79] ———, “Sur les coefficients des représentations unitaires des groupes de Lie simples”, in *Analyse harmonique sur les groupes de Lie (Sém., Nancy-Strasbourg 1976–1978), II*, Lecture Notes in Math., vol. 739, Springer, Berlin, 1979, p. 132–178.
- [Cow97] ———, “Herz’s “principe de majoration” and the Kunze-Stein phenomenon”, in *Harmonic analysis and number theory (Montreal, PQ, 1996)*, CMS Conf. Proc., vol. 21, Amer. Math. Soc., Providence, RI, 1997, p. 73–88.
- [Cow08] ———, “Applications of representation theory to harmonic analysis of Lie groups (and vice versa)”, in *Representation theory and complex analysis*, Lecture Notes in Math., vol. 1931, Springer, Berlin, 2008, p. 1–50.
- [CGCG02] M. COWLING, J. GARCÍA-CUERVA & H. GUNAWAN – “Weighted estimates for fractional maximal functions related to spherical means”, *Bull. Austral. Math. Soc.* **66** (2002), no. 1, p. 75–90.

- [CGGM91] M. COWLING, G. GAUDRY, S. GIULINI & G. MAUCERI – “Weak type $(1, 1)$ estimates for heat kernel maximal functions on Lie groups”, *Trans. Amer. Math. Soc.* **323** (1991), no. 2, p. 637–649.
- [CGHM94] M. COWLING, S. GIULINI, A. HULANICKI & G. MAUCERI – “Spectral multipliers for a distinguished Laplacian on certain groups of exponential growth”, *Studia Math.* **111** (1994), no. 2, p. 103–121.
- [CGM93] M. COWLING, S. GIULINI & S. MEDA – “ L^p - L^q estimates for functions of the Laplace-Beltrami operator on noncompact symmetric spaces. I”, *Duke Math. J.* **72** (1993), no. 1, p. 109–150.
- [CGM95] ———, “ L^p - L^q -estimates for functions of the Laplace-Beltrami operator on noncompact symmetric spaces. II”, *J. Lie Theory* **5** (1995), no. 1, p. 1–14.
- [CGM01] ———, “ L^p - L^q estimates for functions of the Laplace-Beltrami operator on noncompact symmetric spaces. III”, *Ann. Inst. Fourier (Grenoble)* **51** (2001), no. 4, p. 1047–1069.
- [CGM02] ———, “Oscillatory multipliers related to the wave equation on noncompact symmetric spaces”, *J. London Math. Soc. (2)* **66** (2002), no. 3, p. 691–709.
- [CN01] M. COWLING & A. NEVO – “Uniform estimates for spherical functions on complex semisimple Lie groups”, *Geom. Funct. Anal.* **11** (2001), no. 5, p. 900–932.
- [DM00] G. DAVID & P. MATTILA – “Removable sets for Lipschitz harmonic functions in the plane”, *Rev. Mat. Iberoamericana* **16** (2000), no. 1, p. 137–215.
- [DS56] N. DUNFORD & J. T. SCHWARTZ – “Convergence almost everywhere of operator averages”, *J. Rational Mech. Anal.* **5** (1956), p. 129–178.
- [DS88] ———, *Linear operators. part I*, Wiley Classics Library, John Wiley and Sons Inc., New York, 1988, General theory, With the assistance of William G. Bade and Robert G. Bartle, Reprint of the 1958 original, A Wiley-Interscience Publication.
- [Dun39] N. DUNFORD – “A mean ergodic theorem.”, *Duke Math. J.* **5** (1939), p. 635–646.
- [Ebe49] W. F. EBERLEIN – “Abstract ergodic theorems and weak almost periodic functions”, *Trans. Amer. Math. Soc.* **67** (1949), p. 217–240.
- [Eme74] W. R. EMERSON – “The pointwise ergodic theorem for amenable groups”, *Amer. J. Math.* **96** (1974), p. 472–487.

- [EMOT81] A. ERDÉLYI, W. MAGNUS, F. OBERHETTINGER & F. G. TRICOMI – *Higher transcendental functions. vol. I*, Robert E. Krieger Publishing Co. Inc., Melbourne, Fla., 1981.
- [Fel68] W. FELLER – *An introduction to probability theory and its applications. vol. i*, Third edition, John Wiley & Sons Inc., New York, 1968.
- [FTP83] A. FIGÀ-TALAMANCA & M. A. PICARDELLO – *Harmonic analysis on free groups*, Lecture Notes in Pure and Applied Mathematics, vol. 87, Marcel Dekker Inc., New York, 1983.
- [FN98] K. FUJIWARA & A. NEVO – “Maximal and pointwise ergodic theorems for word-hyperbolic groups”, *Ergodic Theory Dynam. Systems* **18** (1998), no. 4, p. 843–858.
- [Gar70] A. M. GARSIA – *Topics in almost everywhere convergence*, Lectures in Advanced Mathematics, vol. 4, Markham Publishing Co., Chicago, Ill., 1970.
- [GGHM88] G. GAUDRY, S. GIULINI, A. HULANICKI & A. M. MANTERO – “Hardy-Littlewood maximal functions on some solvable lie groups”, *J. Austral. Math. Soc. Ser. A* **45** (1988), no. 1, p. 78–82.
- [GGM90] G. GAUDRY, S. GIULINI & A. M. MANTERO – “Asymmetry of maximal functions on the affine group of the line”, *Tohoku Math. J. (2)* **42** (1990), no. 2, p. 195–203.
- [GdIH90] É. GHYS & P. DE LA HARPE (éds.) – *Sur les groupes hyperboliques d’après Mikhael Gromov*, Progress in Mathematics, vol. 83, Birkhäuser Boston Inc., Boston, MA, 1990, Papers from the Swiss Seminar on Hyperbolic Groups held in Bern, 1988.
- [GS90] S. GIULINI & P. SJÖGREN – “A note on maximal functions on a solvable Lie group”, *Arch. Math. (Basel)* **55** (1990), no. 2, p. 156–160.
- [Gla03] E. GLASNER – *Ergodic theory via joinings*, Mathematical Surveys and Monographs, vol. 101, American Mathematical Society, Providence, RI, 2003.
- [God03] R. GODEMENT – *Analyse mathématique. IV*, Intégration et théorie spectrale, analyse harmonique, le jardin des délices modulaires, Springer-Verlag, Berlin, 2003.
- [GNre] A. GORODNIK & A. NEVO – *The ergodic theory of lattice subgroups*, Annals of Mathematics Studies, vol. 172, Princeton University Press, Princeton, NJ, à paraître, arXiv:math.DS/0605596.
- [GMS97] L. GRAFAKOS & S. MONTGOMERY-SMITH – “Best constants for uncentred maximal functions”, *Bull. London Math. Soc.* **29** (1997), no. 1, p. 60–64.

- [Gre69] F. P. GREENLEAF – *Invariant means on topological groups and their applications*, Van Nostrand Mathematical Studies, No. 16, Van Nostrand Reinhold Co., New York, 1969.
- [Gre73] ———, “Ergodic theorems and the construction of summing sequences in amenable locally compact groups”, *Comm. Pure Appl. Math.* **26** (1973), p. 29–46.
- [Gri84] R. I. GRIGORCHUK – “Degrees of growth of finitely generated groups and the theory of invariant means”, *Izv. Akad. Nauk SSSR Ser. Mat.* **48** (1984), no. 5, p. 939–985.
- [Gri99] ———, “Ergodic theorems for the actions of a free group and a free semigroup”, *Math. Notes* **65** (1999), p. 654–657.
- [Gri00] ———, “An ergodic theorem for actions of a free semigroup”, *Proc. Steklov Inst.* **231** (2000), p. 113–127.
- [Gro87] M. GROMOV – “Hyperbolic groups”, in *Essays in group theory*, Math. Sci. Res. Inst. Publ., vol. 8, Springer, New York, 1987, p. 75–263.
- [Gro85] E. GROSSWALD – *Representations of integers as sums of squares*, Springer-Verlag, New York, 1985.
- [Gui69] Y. GUIVARC’H – “Généralisation d’un théorème de von Neumann”, *C. R. Acad. Sci. Paris Sér. A-B* **268** (1969), p. A1020–A1023.
- [Gui73] ———, “Croissance polynomiale et périodes des fonctions harmoniques”, *Bull. Soc. Math. France* **101** (1973), p. 333–379.
- [Hal51] P. R. HALMOS – *Introduction to Hilbert space and the theory of spectral multiplicity*, Chelsea Publishing Company, New York, N. Y., 1951.
- [HL30] G. H. HARDY & J. E. LITTLEWOOD – “A maximal theorem with function-theoretic applications”, *Acta Math.* **54** (1930), p. 81–116.
- [HW79] G. H. HARDY & E. M. WRIGHT – *An introduction to the theory of numbers*, fifth éd., The Clarendon Press Oxford University Press, New York, 1979.
- [Heb93] W. HEBISCH – “The subalgebra of $L^1(AN)$ generated by the Laplacian”, *Proc. Amer. Math. Soc.* **117** (1993), no. 2, p. 547–549.
- [Heb04] ———, “Calderón-Zygmund decomposition and singular integrals for NA groups of higher rank”, Announcement, 2004.
- [HS03] W. HEBISCH & T. STEGER – “Multipliers and singular integrals on exponential growth groups”, *Math. Z.* **245** (2003), no. 1, p. 37–61.
- [Hel84] S. HELGASON – *Groups and geometric analysis*, Integral geometry, invariant differential operators, and spherical functions, Pure and Applied Mathematics, vol. 113, Academic Press Inc., Orlando, FL, 1984. .

- [Her70] C. . HERZ – “Sur le phénomène de Kunze-Stein”, *C. R. Acad. Sci. Paris Sér. A-B* **271** (1970), p. A491–A493.
- [Hop54] E. HOPF – “The general temporally discrete Markoff process”, *J. Rational Mech. Anal.* **3** (1954), p. 13–45.
- [HM79] R. E. HOWE & C. C. MOORE – “Asymptotic properties of unitary representations”, *J. Funct. Anal.* **32** (1979), no. 1, p. 72–96.
- [Ion00a] A. D. IONESCU – “An endpoint estimate for the Kunze-Stein phenomenon and related maximal operators”, *Ann. of Math. (2)* **152** (2000), no. 1, p. 259–275.
- [Ion00b] ———, “A maximal operator and a covering lemma on non-compact symmetric spaces”, *Math. Res. Lett.* **7** (2000), no. 1, p. 83–93.
- [Ion02] ———, “Singular integrals on symmetric spaces of real rank one”, *Duke Math. J.* **114** (2002), no. 1, p. 101–122.
- [Ion03] ———, “Singular integrals on symmetric spaces. II”, *Trans. Amer. Math. Soc.* **355** (2003), no. 8, p. 3359–3378 (electronic).
- [Ion04] ———, “An endpoint estimate for the discrete spherical maximal function”, *Proc. Amer. Math. Soc.* **132** (2004), no. 5, p. 1411–1417 (electronic).
- [Ion05] ———, “Rearrangement inequalities on semisimple Lie groups”, *Math. Ann.* **332** (2005), no. 4, p. 739–758.
- [Jol90] P. JOLISSAINT – “Rapidly decreasing functions in reduced C^* -algebras of groups”, *Trans. Amer. Math. Soc.* **317** (1990), no. 1, p. 167–196.
- [Jon93] R. L. JONES – “Ergodic averages on spheres”, *J. Anal. Math.* **61** (1993), p. 29–45.
- [JRW03] R. L. JONES, J. M. ROSENBLATT & M. WIERDL – “Oscillation in ergodic theory: higher dimensional results”, *Israel J. Math.* **135** (2003), p. 1–27.
- [JSW08] R. L. JONES, A. SEEGER & J. WRIGHT – “Strong variational and jump inequalities in harmonic analysis”, *Trans. Amer. Math. Soc.* **360** (2008), no. 12, p. 6711 – 6742.
- [Jou86] J.-L. JOURNÉ – “A covering lemma for product spaces”, *Proc. Amer. Math. Soc.* **96** (1986), no. 4, p. 593–598.
- [Jou88] ———, “Two problems of Calderón-Zygmund theory on product-spaces”, *Ann. Inst. Fourier (Grenoble)* **38** (1988), no. 1, p. 111–132.
- [Kak51] S. KAKUTANI – “Random ergodic theorems and Markoff processes with a stable distribution”, in *Proceedings of the Second Berkeley Symposium*

- on Mathematical Statistics and Probability, 1950* (Berkeley and Los Angeles), University of California Press, 1951, p. 247–261.
- [Khi33] A. KHINTCHINE – “Zu Birkhoffs Lösung des Ergodenproblems”, *Math. Ann.* **107** (1933), p. 485–488.
- [Koo84] T. H. KOORNWINDER – “Jacobi functions and analysis on noncompact semisimple Lie groups”, in *Special functions: group theoretical aspects and applications*, Math. Appl., Reidel, Dordrecht, 1984, p. 1–85.
- [Kre85] U. KRENGEL – *Ergodic theorems*, de Gruyter Studies in Mathematics, vol. 6, Walter de Gruyter and Co., Berlin, 1985.
- [KB37] N. KRYLOFF & N. BOGOLIUBOFF – “La théorie générale de la mesure dans son application à l’étude des systèmes dynamiques de la mécanique non linéaire”, *Ann. of Math.* **38** (1937), p. 65 – 113.
- [KS60] R. A. KUNZE & E. M. STEIN – “Uniformly bounded representations and harmonic analysis of the 2×2 real unimodular group”, *Amer. J. Math.* **82** (1960), p. 1–62.
- [Lac95] M. T. LACEY – “Ergodic averages on circles”, *J. Anal. Math.* **67** (1995), p. 199–206.
- [Lac07] ———, “Commutators with Riesz potentials in one and several parameters”, *Hokkaido Math. J.* **36** (2007), no. 1, p. 175–191.
- [Leb72] N. N. LEBEDEV – *Special functions and their applications*, Dover Publications Inc., New York, 1972, Revised edition, translated from the Russian and edited by Richard A. Silverman.
- [Lin01] E. LINDENSTRAUSS – “Pointwise theorems for amenable groups”, *Invent. Math.* **146** (2001), no. 2, p. 259–295.
- [Mac03] I. G. MACDONALD – *Affine Hecke algebras and orthogonal polynomials*, Cambridge Tracts in Mathematics, vol. 157, Cambridge University Press, Cambridge, 2003.
- [Mag97] A. MAGYAR – “ L^p -bounds for spherical maximal operators on \mathbb{Z}^n ”, *Rev. Mat. Iberoamericana* **13** (1997), no. 2, p. 307–317.
- [Mag02] ———, “Diophantine equations and ergodic theorems”, *Amer. J. Math.* **124** (2002), no. 5, p. 921–953.
- [MSW02] A. MAGYAR, E. M. STEIN & S. WAINGER – “Discrete analogues in harmonic analysis: spherical averages”, *Ann. of Math. (2)* **155** (2002), no. 1, p. 189–208.
- [MNS00] G. A. MARGULIS, A. NEVO & E. M. STEIN – “Analogues of Wiener’s ergodic theorems for semisimple Lie groups. II”, *Duke Math. J.* **103** (2000), no. 2, p. 233–259.

- [MMNO00] J. MATEU, P. MATTILA, A. NICOLAU & J. OROBITG – “BMO for nondoubling measures”, *Duke Math. J.* **102** (2000), no. 3, p. 533–565.
- [MV08] S. MEDA & M. VALLARINO – “Weak type estimates for spherical multipliers on noncompact symmetric spaces”, Tech. report, arXiv:0802.1227, 2008.
- [Mel03] A. D. MELAS – “The best constant for the centered Hardy-Littlewood maximal inequality”, *Ann. of Math. (2)* **157** (2003), no. 2, p. 647–688.
- [MSS92] G. MOCKENHAUPT, A. SEEGER & C. D. SOGGE – “Wave front sets, local smoothing and Bourgain’s circular maximal theorem”, *Ann. of Math. (2)* **136** (1992), no. 1, p. 207–218.
- [Mor47] A. P. MORSE – “Perfect blankets”, *Trans. Amer. Math. Soc.* **61** (1947), p. 418–442.
- [MS04] D. MÜLLER & A. SEEGER – “Singular spherical maximal operators on a class of two step nilpotent Lie groups”, *Israel J. Math.* **141** (2004), p. 315–340.
- [MPTT06] C. MUSCALU, J. PIPHER, T. TAO & C. THIELE – “Multi-parameter paraproducts”, *Rev. Mat. Iberoam.* **22** (2006), no. 3, p. 963–976.
- [NT04] E. K. NARAYANAN & S. THANGAVELU – “An optimal theorem for the spherical maximal operator on the Heisenberg group”, *Israel J. Math.* **144** (2004), p. 211–219.
- [NTV97] F. NAZAROV, S. TREIL & A. VOLBERG – “Cauchy integral and Calderón-Zygmund operators on nonhomogeneous spaces”, *Internat. Math. Res. Notices* **15** (1997), p. 703–726.
- [NTV02] F. NAZAROV, S. TREIL & A. VOLBERG – “Accretive system Tb -theorems on nonhomogeneous spaces”, *Duke Math. J.* **113** (2002), no. 2, p. 259–312.
- [Nev70] J. NEVEU – *Bases mathématiques du calcul des probabilités*, Deuxième édition, revue et corrigée, Masson et Cie, Éditeurs, Paris, 1970.
- [Nev72] ———, *Martingales à temps discret*, Masson et Cie, éditeurs, Paris, 1972.
- [Nev94a] A. NEVO – “Harmonic analysis and pointwise ergodic theorems for noncommuting transformations”, *J. Amer. Math. Soc.* **7** (1994), no. 4, p. 875–902.
- [Nev94b] ———, “Pointwise ergodic theorems for radial averages on simple Lie groups. I”, *Duke Math. J.* **76** (1994), no. 1, p. 113–140.
- [Nev97] ———, “Pointwise ergodic theorems for radial averages on simple Lie groups. II”, *Duke Math. J.* **86** (1997), no. 2, p. 239–259.

- [Nev98] ———, “Spectral transfer and pointwise ergodic theorems for semi-simple Kazhdan groups”, *Math. Res. Lett.* **5** (1998), no. 3, p. 305–325.
- [Nev06] ———, “Pointwise ergodic theorems for actions of groups”, in *Handbook of dynamical systems. Vol. 1B*, Elsevier B. V., Amsterdam, 2006, p. 871–982.
- [NS94] A. NEVO & E. M. STEIN – “A generalization of Birkhoff’s pointwise ergodic theorem”, *Acta Math.* **173** (1994), no. 1, p. 135–154.
- [NS97] ———, “Analogues of Wiener’s ergodic theorems for semisimple groups. I”, *Ann. of Math. (2)* **145** (1997), no. 3, p. 565–595.
- [NT97] A. NEVO & S. THANGAVELU – “Pointwise ergodic theorems for radial averages on the Heisenberg group”, *Adv. Math.* **127** (1997), no. 2, p. 307–334.
- [Orn69] D. ORNSTEIN – “On the pointwise behavior of iterates of a self-adjoint operator”, *J. Math. Mech.* **18** (1968/1969), p. 473–477.
- [Oxt52] J. C. OXTOBY – “Ergodic sets”, *Bull. Amer. Math. Soc.* **58** (1952), p. 116–136.
- [Pat88] A. L. T. PATERSON – *Amenability*, Mathematical Surveys and Monographs, vol. 29, American Mathematical Society, Providence, RI, 1988.
- [Rat94] J. G. RATCLIFFE – *Foundations of hyperbolic manifolds*, Graduate Texts in Mathematics, vol. 149, Springer-Verlag, New York, 1994.
- [Rau76] G. RAUZY – *Propriétés statistiques de suites arithmétiques*, Presses Universitaires de France, Paris, 1976, Le Mathématicien, No. 15, Collection SUP.
- [Rie38] F. RIESZ – “Some mean ergodic theorems”, *J. Lond. Math. Soc.* **1** (1938), p. 274–278.
- [Rot62] G.-C. ROTA – “An “Alternierende Verfahren” for general positive operators”, *Bull. Amer. Math. Soc.* **68** (1962), p. 95–102.
- [Rub86] J. L. RUBIO DE FRANCIA – “Maximal functions and Fourier transforms”, *Duke Math. J.* **53** (1986), no. 2, p. 395–404.
- [Rud91] W. RUDIN – *Functional analysis*, second éd., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991.
- [Sch98] W. SCHLAG – “A geometric proof of the circular maximal theorem”, *Duke Math. J.* **93** (1998), no. 3, p. 505–533.
- [Sch81] K. SCHMIDT – “Amenability, Kazhdan’s property T , strong ergodicity and invariant means for ergodic group-actions”, *Ergodic Theory Dynam. Systems* **1** (1981), no. 2, p. 223–236.

- [STW03] A. SEEGER, T. TAO & J. WRIGHT – “Endpoints mapping properties of spherical maximal operators”, *J. Inst. Math. Jussieu* **2** (2003), no. 1, p. 109–144.
- [Sjö83] P. SJÖGREN – “A remark on the maximal function for measures in \mathbf{R}^n ”, *Amer. J. Math.* **105** (1983), no. 5, p. 1231–1233.
- [Ste61] E. M. STEIN – “On the maximal ergodic theorem”, *Proc. Nat. Acad. Sci. U.S.A.* **47** (1961), p. 1894–1897.
- [Ste70a] ———, *Singular integrals and differentiability properties of functions*, Princeton Mathematical Series, vol. 30, Princeton University Press, Princeton, N.J., 1970.
- [Ste70b] ———, *Topics in harmonic analysis related to the Littlewood-Paley theory.*, Annals of Mathematics Studies, vol. 63, Princeton University Press, Princeton, N.J., 1970.
- [Ste70c] ———, “Analytic continuation of group representations”, *Advances in Math.* **4** (1970), p. 172–207 (1970).
- [Ste76] ———, “Maximal functions. I. Spherical means”, *Proc. Nat. Acad. Sci. U.S.A.* **73** (1976), no. 7, p. 2174–2175.
- [Ste93] ———, *Harmonic analysis: real-variable methods, orthogonality, and oscillatory integrals*, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993.
- [SS83] E. M. STEIN & J.-O. STRÖMBERG – “Behavior of maximal functions in \mathbf{R}^n for large n ”, *Ark. Mat.* **21** (1983), no. 2, p. 259–269.
- [SW78] E. M. STEIN & S. WAINGER – “Problems in harmonic analysis related to curvature”, *Bull. Amer. Math. Soc.* **84** (1978), no. 6, p. 1239–1295.
- [SW99] ———, “Discrete analogues in harmonic analysis. I. ℓ^2 estimates for singular Radon transforms”, *Amer. J. Math.* **121** (1999), no. 6, p. 1291–1336.
- [SW00] ———, “Discrete analogues in harmonic analysis. II. Fractional integration”, *J. Anal. Math.* **80** (2000), p. 335–355.
- [SW71] E. M. STEIN & G. WEISS – *Introduction to Fourier analysis on Euclidean spaces*, Princeton Mathematical Series, vol. 32, Princeton University Press, Princeton, N.J., 1971, , No. 32.
- [Str81] J.-O. STRÖMBERG – “Weak type L^1 estimates for maximal functions on noncompact symmetric spaces”, *Ann. of Math. (2)* **114** (1981), no. 1, p. 115–126.

- [Tao01] T. TAO – “From rotating needles to stability of waves: emerging connections between combinatorics, analysis, and PDE”, *Notices Amer. Math. Soc.* **48** (2001), no. 3, p. 294–303.
- [Tem67] A. TEMPELMAN – “Ergodic theorems for general dynamical systems”, *Dokl. Akad. Nauk SSSR* **176** (1967), p. 790–793, (English translation: *Soviet Math. Dokl.*, **8** (1967), 1213–1216.).
- [Tes05] R. TESSERA – “Asymptotic volumes of spheres in metric measured spaces and in polynomial groups”, Tech. report, arXiv:math.GR/0506362v3, 2005.
- [Tol01a] X. TOLSA – “ BMO , H^1 , and Calderón-Zygmund operators for non doubling measures”, *Math. Ann.* **319** (2001), no. 1, p. 89–149.
- [Tol01b] ———, “A proof of the weak $(1, 1)$ inequality for singular integrals with non doubling measures based on a Calderón-Zygmund decomposition”, *Publ. Mat.* **45** (2001), no. 1, p. 163–174.
- [Val06] M. VALLARINO – “A maximal function on harmonic extensions of H -type groups”, *Ann. Math. Blaise Pascal* **13** (2006), no. 1, p. 87–101.
- [Val07] ———, “Spectral multipliers on Damek-Ricci spaces”, *J. Lie Theory* **17** (2007), no. 1, p. 163–189.
- [Var85] V. S. VARADARAJAN – *Geometry of quantum theory*, second éd., Springer-Verlag, New York, 1985.
- [vN32] J. VON NEUMANN – “Proof of the quasi-ergodic hypothesis”, *Proc. Nat. Acad. Sci. USA* **18** (1932), p. 70–82.
- [Wal57] A. WALFISZ – *Gitterpunkte in mehrdimensionalen Kugeln*, Monografie Matematyczne. Vol. 33, Państwowe Wydawnictwo Naukowe, Warsaw, 1957.
- [Wei03] B. WEISS – “Actions of amenable groups”, in *Topics in Dynamics and Ergodic Theory*, S. Bezuglyi, S. Kolyada (Eds), London Math. Soc. Lecture Notes, vol. 310, Cambridge University Press, 2003.
- [Wie39] N. WIENER – “The ergodic theorem”, *Duke Math. J.* **5** (1939), p. 1–18.
- [YK39] K. YOSIDA & S. KAKUTANI – “Birkhoff’s ergodic theorem and the maximal ergodic theorem”, *Proc. Imp. Acad., Tokyo* **15** (1939), p. 165–168.
- [YK41] ———, “Operator-theoretical treatment of Markoff’s process and mean ergodic theorem”, *Ann. of Math. (2)* **42** (1941), p. 188–228.
- [Zyg02] A. ZYGMUND – *Trigonometric series. Vol. I, II*, third ed., Cambridge Mathematical Library, Cambridge University Press, Cambridge, 2002.

INDEX

- A : sous-groupe des homothéties, 76
- A^* : opérateur maximal, 16, 45, 49
- A^{**} : opérateur maximal non centré, 49
- C^* -algèbre, 201
- G -espace mesuré, 195
- G -espace mesurable, 195
- M^* : opérateur maximal pour les moyennes sur les translatés à gauche, 82
- N : sous-groupe des translations, 76
- NAK : décomposition d'Iwasawa, 78, 216
- S_d : groupe des affinités, 76, 216
- $T(\mu)$: opérateur de convolution par $\tilde{\mu}$, 34, 197
- T_g : opérateur de translation, 34, 197
- $\mathbb{E}(f|\mathcal{I})$: espérance conditionnelle de f sachant \mathcal{I} , 14
- \mathbb{F}_d : groupe libre à d générateurs, 135
- \mathbb{H}^{d+1} : espace hyperbolique, 77, 213
- $|E|$: mesure de Haar de E , 28
- $|w|$: longueur du mot $w \in \mathbb{F}_d$, 135
- $|x|$: norme euclidienne de $x \in \mathbb{R}^d$, 86, 100
- β_r : probabilité uniforme sur la boule centrée en l'élément neutre, de rayon r , 46, 139, 168
- \mathcal{I} : tribu des invariants, 13
- $\text{Sp}_n(A)$: espace des caractères auto-adjoints de l'algèbre involutive A , 205
- $\text{Sp}(A)$: spectre de l'algèbre A , 202
- $\mathbf{1}_A$: fonction indicatrice de l'ensemble A , 18
- λ : mesure de Haar invariante à gauche, 28
- λ_d : mesure de Haar invariante à droite, 29
- Δ : fonction modulaire d'un groupe, 31
- Δ : laplacien, 171
- $\mu * f$: produit de convolution de μ par f , 34
- $\tilde{\mu}$: symétrique de la mesure μ , 197
- $\mu \cdot f$: produit de convolution de $\tilde{\mu}$ par f , 34, 197
- μ^* : adjoint de la mesure μ , 196
- $\pi(\mu)$: opérateur de convolution par μ , 197
- \mathbb{S}^{d-1} : sphère unité de \mathbb{R}^d , 100
- σ_r : probabilité uniforme sur la sphère centrée en l'élément neutre, de rayon r , 100, 139, 168
- $f_1 \asymp f_2$: rapport f_1/f_2 borné, 84
- $g(\varphi)$: fonction de Littlewood-Paley relative aux moyennes sphériques associée à φ , 103
- dg : mesure de Haar invariante à gauche, 28
- $\mathcal{M}(G)$: espace des mesures bornées sur G , 196
- $\mathcal{M}(G, K)$: espace des mesures bornées bi- K -invariantes sur G , 168
- action
 - ergodique, 37
 - topologiquement transitive, 39, 137
 - uniquement ergodique, 38
- algèbre de Banach involutive, 201
- boule hyperbolique, 78
- caractère
 - auto-adjoint d'une C^* -algèbre commutative, 202
 - d'une algèbre de Banach commutative, 202
- condition de Tempel'man, 58
- contraction $L^1 - L^\infty$, 20, 207
- croissance
 - exponentielle, 31
 - polynomiale, 31
 - sous-exponentielle, 31

- équidistribution des orbites, 138
- espace hyperbolique
 - modèle de la boule unité, 214
 - modèle de Minkowski, 214
 - modèle du demi-espace, 213
- espace mesurable standard, 195
- famille continue à un paramètre de mesures
 - de probabilité, 40, 198
- fonction de Littlewood-Paley, 101, 102
- fonction maximale de Hardy-Littlewood, 46
- fonction maximale dyadique, 56
- fonction maximale non centrée, 48
- fonction sphérique, 144, 186
- groupe moyennable, 29
- groupe unimodulaire, 31
- inégalité faible, 17
- inégalité forte, 17
- inégalité maximale sous-additive, 149
- lemme de recouvrement de Besicovitch, 50
- lemme de recouvrement de Vitali, 47
- lemme maximal de Doob, 56
- métrique hyperbolique, 77, 213
- mesure spectrale associée à un vecteur, 203
- moyenne de Birkhoff, 11
- opérateur
 - borné dans L^p , 17
 - de type faible (p, p) , 17
 - de type fort (p, p) , 17
- opérateur de Markov, 20, 207
- opérateur maximal, 16, 45
- opérateur semi-local, 42
- opérateur sous-additif, 17
- phénomène de Kunze-Stein, 85
- posséder presque des vecteurs invariants, 32
- principe de Banach, 16
- propriété (T) de Kazhdan, 33
- propriété de doublement de volume, 50
- réseau dyadique, 56
- résolution de l'identité, 202
- représentation unitaire continue, 197
- suite asymptotiquement invariante, 28
- suite de Følner, 28
 - à droite, 29
 - à gauche, 28
 - uniforme, 29
- suite sous-additive d'opérateurs de Markov, 148
- suite tempérée, 62
- transformation
 - ergodique, 24
 - uniquement ergodique, 25
- transformation de Gelfand, 202
- tribu des invariants, 13
- trou spectral, 33
- vecteurs de classe C^∞ , 197